University of Naples Federico II
Department of Electrical Engineering and Information Technology

Doctoral Thesis in Information and Communication Technology for Health

Optimizing Healthcare Logistics: Innovative Models, Methods, and Applications for Location, Scheduling, and Routing

Academic year 2023/2024

Tutor
Prof. Claudio Sterle

Coordinator
Prof. Daniele Riccio

Candidate
Andrea Mancuso

Abstract

This Ph.D. thesis explores the application of Information and Communication Technology (ICT) in healthcare logistics, assessing its potential to rationalize operations. At its core, this research primarily focuses on application of Operational Research methods to improve the efficiency and reliability of healthcare systems.

Before assessing specific problem arising from real world, this study offers an overview of the existing work in the field. Specifically, it highlights the key contributions and the potential for optimization enabled by technological progress.

Following this, the thesis progresses to its core scientific contribution, tackling a challenge identified within the context of Facility Location problems, specifically focusing on the strategic management of the Blood Supply Chain. This study introduces two modeling frameworks: a case-based and a scenario-based approach. Both are modeled using hard and soft constraints to represent real-world conditions that should be adhered to. The objective is to balance donor self-sufficiency with cost minimization in system management. The scenario-based formulation is presented in two variations: an average-case and a worst-case scenario, to address the sensitivity of the solution to perturbations in conditions. The effectiveness of this approach is confirmed by a broad experimentation on a real case study in the Italian regions of Campania, Puglia, and Lombardy.

After exploring Facility Location, the thesis shifts its attention to surgery scheduling, addressing the challenges of Integrated Operating Room Planning and Scheduling faced by hospitals. It introduces a tool based on two Integer Linear Programs (ILPs) designed to optimize surgical schedules. The tool aims to balance the maximization of performed elective surgeries with the management of emergency operations. The first ILP aims to enhance room occupancy, considering the potential arrival of emergencies that must be promptly integrated into the schedule, thereby providing an initial plan for the surgical department. The second ILP operates in real time, re-optimizing the schedule whenever there is a deviation from the planned activities, such as surgeries lasting longer than expected or the arrival of an emergency case. This ensures efficiency and responsiveness throughout the entire planning horizon. The effectiveness of these models is further underscored by their application to real world data from a hospital in Naples.

Additionally, the thesis investigates the field of drone logistics, focusing on potential application in healthcare delivery. It introduces an innovative Mixed-Integer Linear Programming model to address the Flying Sidekick Traveling Salesman Problem, employing a Branch and Cut approach. This method advances the field by representing the state of the art. This exploration not only provides a new methodology to the existing literature but also opens to further research in healthcare logistics, especially concerning the use of drones for efficient and timely delivery.

In conclusion, this thesis covers a wide range of Operational Research applications in healthcare. It exploits issues ranging from strategic planning and policy development to practical and logistical enhancements, providing a perspective on the potential and solutions in healthcare logistics influenced by ICT advancements. Each part of the study addresses specific issues while collectively offering a unified view of improving efficiency and effectiveness in healthcare logistics.

Acknowledgments

I wish to extend my deepest appreciation to my advisor, Professor Claudio Sterle, for providing me the guidance and resources fundamental to the success of my research. I am similarly grateful to Professor Maurizio Boccia and Adriano Masone from the Department of Electrical Engineering and Information Technology at the University of Naples Federico II for their invaluable inputs and assistance throughout this journey.

I extend my deepest and most heartfelt gratitude to the incredible individuals with whom I had the privilege of building friendships during this experience. To Imanol, Cristian T., Cristiano, Marta B., Maria, Ricardo, Bikey, Alan, Paulo, Hannah, Isabel, Sven, Miguel, Marta C., Jacob, Vittorio, Silvia B., Silvia C., Serena, Tom, Juan Manuel, Nicolas, Cristian D. M., Mikele, Martina C., Martina G., Giulia, Alberto, Andrea and Alice, among others. The shared experiences and memories have transformed this journey into an unforgettable adventure.

Lastly, I reserve my deepest and everlasting gratitude for the pillars of my life: my parents, family, and closest friends—all those who have always supported me. To Adele, Stefano, Fabrizio, Corrado, Emilio, Roberto, Andrea, Gianfranco, Michele, Maurizio, Ivan, Sergio, Valerio, Alessandra, Chiara, Marco, Eduardo, Cristina, Carlo, and Sabatino, I am profoundly grateful. Your constant support, understanding, and patience have been an endless source of inspiration. Thank you for standing by me, particularly during the most challenging moments, and for providing a constant encouragement when it was most essential.

Andrea Mancuso

Contents

Introduction

In recent years, the global healthcare landscape has undergone rapid transformations, largely driven by developments in Information and Communication Technology. This area includes new strategies in sensing, data analysis, logistics, and robotics, all with the goal of making healthcare systems more efficient and reliable. Leveraging the broad scope of the Ph.D. program in Information and Communication Technology for Healthcare (ICTH) at the University of Naples Federico II, which covers four key areas: Sensing for Health, Data for Health, Logistics for Health, and Robotics for Health, this thesis focuses specifically on the "Logistics for Health" aspect. The goal is to identify opportunities and create practical solutions in the field of healthcare logistics.

Recognizing the central role of health logistics in the healthcare system, we see that this area is rich of strategies and technologies designed to enhance the healthcare supply chain, optimize resource management, and ensure timely service delivery. To further improve its efficiency, reduce costs, and promote superior patient care, continuous refinement and innovation within this sector are essential. This involves leveraging organized, data-driven, and adaptable logistics systems. Taking a step in this direction, this thesis identifies three key focus areas within the Logistics for Health pillar — Blood Supply Chain Management, Integrated Operating Rooms Planning and Scheduling, and last-mile logistics, with the goals of developing a mathematical framework to address regional Blood Supply Chain inefficiencies considering various factors and constraints; creating a tool to optimize surgery scheduling by balancing operating room use and reactivity to unexpected events; and developing a new approach to solve the Flying Sidekick Traveling Salesman Problem in last-mile logistics, keeping in mind its possible future applications in healthcare.

This thesis is organized into four distinct parts, comprising five chapters. Each chapter addresses a crucial aspect of healthcare logistics, integrating valuable perspectives from ICTH.

Part 1, represented by Chapter 1, serves as an introduction to the Operational Research in healthcare. It provides an overview of the current landscape, helping readers understand the various tools and technologies spread in the industry today, and also offers insights into the application of Operational Research in health sector, highlighting significant advancements and implementations.

Part 2, detailed in Chapter 2, tackle the changes proposed by the Italian Healthcare Ministry for efficient the Blood Supply Chain management. The chapter propose a mathematical modeling methodology to work towards a system that is both efficient and adaptable to changing circumstances. It focus on the practical implications of these models, aiming to demonstrate how they can make the Blood Supply Chain more responsive and reliable.

Part 3 encompasses Chapters 3 and 4, address the Integrated Operating Room Planning and Scheduling problem faced by a hospital in Naples using a two-step strategy rooted into a Integer Linear Programming based optimization tool. Specifically, Chapter 3 aims to provide an optimal weekly schedule for the surgery department. This schedule focuses on elective surgeries, with the aim to maximize room occupancy. It also ensures timely accommodation of emergencies trying to minimize disruptions from unexpected events, leading to a more resilient and efficient surgery schedule. On the other hand, Chapter 4 address the dynamic aspects of the problem. If a surgery duration varies from the initial plan or an emergency arises, the schedule is promptly re-optimized. This ensures continued efficiency and adaptability, with a goal to adhere closely to the initial schedule. Both chapters leverage real data from the hospital for experimentation and analysis, offering insights that might benefit other hospitals facing similar challenges.

Part 4, detailed in Chapter 5, investigates the applicability of solutions from the Flying Sidekick Traveling Salesman Problem in healthcare. Although healthcare is not the primary focus, the chapter introduces an innovative approach to tackle vehicle synchronization issues, highlighting the potential of such solutions to enhance healthcare logistics. This is especially relevant for transportation and delivery systems within the healthcare domain.

Each chapter stands alone in addressing specific challenges but also comes together to provide a comprehensive view of the opportunities and solutions in healthcare logistics leveraging ICTH innovations.

The primary objective of this thesis is to enhance efficiency and effectiveness in healthcare logistics. It addresses challenges such as inefficiencies in the Blood Supply Chain and surgery scheduling, along with issues in last-mile delivery. The aim is to provide more than just theoretical understanding; it focuses on delivering practical, actionable solutions that can substantially improve logistics within the healthcare industry.

The research in these chapters has been done in collaboration with Professors Maurizio Boccia, Adriano Masone, and Claudio Sterle from the Optimization and Problem Solving Laboratory at the University of Naples Federico II.

In conclusion, we want to underline that the studies discussed in the chapters hold up well against or even lead in the current standards of the topics they address. Indeed, they have led to papers that are submitted to or published on some valuable journals in the Operations Research community. Specifically, the research found in chapter 2 has lead to a paper and submitted to *Transportation Research Part E: Logistics and Transportation Review* journal. The work from chapter 3 and 4 resulted in two papers that are submitted to *Soft Computing* and *Operations Research for Health Care* journals. Lastly, the study discussed in chapter 4 has already been published on *Networks* journal.

Chapter 1

Operational Research and healthcare optimization: an overview

Operational Research, traditionally rooted in decision-making and problem-solving, uses advanced analytical techniques to improve system efficiency. When applied to healthcare, it seeks to enhance patient outcomes, ensure timely care delivery, and optimize resource utilization. This chapter provides an introductory exploration into how these mathematical and analytical methods are applied in healthcare settings, highlighting their potential.

1.1 Introduction

In a world where societies are contending with demographic shifts marked by declining birth rates in developed regions and increasing average life expectancy globally, the need to manage and improve how we spend on public healthcare has become a pressing issue. Indeed, the demographic shift is putting a significant strain on healthcare systems around the world, requiring careful planning to ensure that resources adequately match the escalating demand while preventing costs from spiraling out of control. At the same time, the health of a nation citizens is closely linked with its economic stability, increasing the imperative for governments to identify and pursue a sustainable course of action. The situation is made even more complex by rapid advances in medical technology. While these advances bring promising new ways to treat patients, they also come with higher costs. Indeed, governments find themselves treading a delicate balance, needing to integrate the latest technology without letting expenses escalate to unsustainable levels. In this intricate scenario, the guidance from the Lancet Global Health Commission comes as a decisive directive [131]. The commission strongly advocates for nations to not only increase their investments in Primary Health Care but also to channel these investments more wisely through well-planned financial strategies. By doing so, it aims to assist governments in navigating the delicate balance of incorporating the latest medical technologies while averting unsustainable cost escalations, therefore encouraging a robust and economically stable healthcare landscape.

Operational Research could play a vital role in this effort, helping to find practical solutions to complex problems through mathematical analysis and strategic thinking. Originating from strategic military operations during World War II, Operational Research has evolved into an essential tool in various fields such as transportation, logistics, and finance [194]. Utilizing a combination of mathematical modeling, statistical analysis, and optimization techniques, Operational Research facilitates informed and data-driven decision-making, bringing a level of precision that is vital in today healthcare environment.

In more recent years, Operational Research evolved to expand its focus from primarily addressing resource allocation and strategic planning to exploring the operational issues, finding innovative solutions to systemic bottlenecks. This progression forecasts a future where data-driven strategies, mathematical programming, simulation, and heuristic algorithms are standard tools, empowering both researchers and practitioners to tackle the complex optimization challenges faced by the healthcare sector.

Indeed, Operational Research enables a thorough analysis of the diverse optimization opportunities within healthcare systems. It aids in the strategic placement of medical facilities and emergency response units, and guides the careful management of resources, including skilled personnel and necessary infrastructure. Furthermore, it aims to lead in the development of improved treatment methods for common diseases, and devises strategies to prevent the spread of infectious ones through proactive and controlled approaches [31].

Implementing Operational Research techniques in healthcare aims to substantially enhance the efficiency and effectiveness of healthcare systems. It fosters a holistic approach where resources are allocated more efficiently, schedules are fine-tuned to avoid bottlenecks, and treatment plans are developed with a high degree of precision, taking into account the specific needs of individual patients while ensuring a balanced system. The primary objective is to enhance the quality of patient care while fostering an cost-effective and high-performing environment.

In the remaining of the chapter, we break down the role of Operational Research in healthcare, covering several key areas: service planning, resource scheduling, logistics, medical therapeutics, disease diagnosis, and preventive care, following the structure proposed in [231]. This structure provides a clear understanding of how Operational Research is applied across various facets of the healthcare system.
In particular subsection 1.2, covers logistics and planning aspects, which lay the groundwork for healthcare strategies. Subsection 1.2.3 address healthcare management and its challenges. In subsection 1.3 we tackle the issue of healthcare practices. Finally, in subsection 1.4, we discuss specialized and preventive care methods.

1.2 Healthcare logistics and planning

Healthcare planning has become increasingly more important in the modern society. Indeed, as stated in Section 1.1 with an aging population and lower birth rates, many countries are feeling the pressure to find extra budget and resources to meet the healthcare needs of their citizens. This planning focuses on several main aspects: forecasting future healthcare needs, determining optimal locations for new healthcare centers and emergency vehicles, designing efficient routes for individual vehicles and entire fleets, and ensuring the optimal utilization of existing resources In the following sections, we explore these topics in more detail, using insights from extensive research and real-world case studies to provide an overview on the present state and upcoming trends in healthcare planning.

1.2.1 Demand forecasting

Predicting the future demand for healthcare services is a vital first step in planning. As shown in [281], demand forecasting is identified as a critical weak link in health supply chains. Accurate forecasting is essential for planning, resource allocation, and logistics management. While there are both qualitative and quantitative methods available for forecasting, the latter is often preferred due to its reliance on hard data, which tends to offer more accurate predictions.

Numerous researchers have explored this field, creating models to more precisely forecast future demands. In this context, in [111] is delineated a detailed nine-step process that underscores the importance of utilizing reliable historical data in forecasting healthcare service demand. Additionally, as discussed in [76], other scholars have explored various prevalent methods such as trend-line and seasonality approaches to anticipate healthcare service demands.

The reliability of different forecasting methods is examined in depth in [146]. They utilized data from daily patient arrivals at the emergency departments of three distinct hospitals, evaluating several methodologies including time series regression, exponential smoothing, seasonal auto-regressive integrated moving average, and artificial neural network models.

A study presented in [33], employed a wide array of data to define a market-based approach to healthcare service forecasting. This approach considered factors such as demographic variations in service-area populations, discharge utilization rates, and market dynamics to develop potential future demand scenarios based on explicit assumptions and trends.

In [201] is introduced a two-step strategy to forecast both future demand and capacity needs. Their strategy centered on crafting a comprehensive facility master plan that encompasses projected capacity and physician requirements.

In the study presented in [28], the authors addressed the complexities of demand forecasting and capacity management in emergency healthcare services. In particular, they introduced an integrated approach that combines advanced forecasting methods like neural networks and support vector regression with a capacity management logic and stochastic simulation. This approach empowers hospital managers to optimize resource allocation based on cost-benefit analysis.

The work presented in [308] analyzed the rising trend of end-stage renal disease population in the US. Utilizing historical data, the researchers forecasted developments up until 2010 through the application of step-wise auto-regressive and exponential smoothing models. This research underscores the pivotal role of data-driven strategies in predicting healthcare trends.

In [322] is faced the challenge of scheduling nursing shifts in advance when surgery demand is uncertain. Using seasonal auto-regressive integrated moving average modeling and data from a trauma center, the research developed a predictive model for short-term daily surgical case forecasts. Testing the proposed framework using long-term historical data, including holiday indicators, the results demonstrate the model effectiveness in estimating surgical case volumes 2-4 weeks in advance, enabling more reliable staff scheduling in healthcare facilities.

The study proposed in [269] addresses the challenge of managing blood supplies due to supply uncertainty and short shelf life of blood products, which often leads to wastage and shortages. The goal is to efficiently forecast blood component supply at blood facilities. Two types of forecasting techniques, time series (including ARIMA and seasonal ESM models) and machine learning (including ANN and multiple regression), were compared using five years of historical data from the Taiwan Blood Services Foundation. The study found that time series forecasting methods, particularly seasonal ESM and ARIMA models, outperformed machine learning algorithms.

In the study presented in [151], the authors centered their focus on healthcare supply chains, specifically managing blood demand and supplies. They presented a real case study where they forecast monthly demand for three vital blood components (red blood cells, plasma and platelets) using Artificial Neural Networks. They utilized real data from a transfusion center and a University Hospital. The study demonstrates that Artificial Neural Networks models outperform ARIMA models in blood demand forecasting, offering a promising approach to enhance Blood Supply Chain management.

The study detailed in [49] focuses on forecasting the demand for healthcare services. It gathered historical data on the demand for a specific drug over a span of 68 months from the University Health Centre. The primary objective of this research is to actively optimize inventory management within the healthcare setting. To achieve this goal, the study employed quantitative forecasting methods, with a specific emphasis on time series analysis, to actively identify underlying data trends. In particular, the research subjected ten distinct forecasting methods to rigorous testing and evaluation, showing that regression analysis yielded the most accurate results among these methods.

1.2.2 Location selection

Numerous studies have emphasized the importance of facility location decisions in the strategic planning of healthcare systems for both private and public organizations, as reviewed in [8]. The location and number of these facilities can significantly affect capital and inventory costs, as well as the quality of customer services. This is especially true in healthcare, where incorrect facility location decisions can have serious repercussions on community health, including increased morbidity and mortality rates.

Early studies in this area, such as the one proposed in [120], focused on locating hospitals and determining their capacities in specific regions.

Over the years, researchers have reviewed and classified various healthcare facilities (HCF) location studies from different angles, including the development planning of health services in developing nations, the evolution of ambulance location and relocation models.

Healthcare centers

Choosing the right locations for healthcare centers is a complex but crucial task, involving a careful consideration of various factors such as accessibility, population density, and existing healthcare infrastructure. For the sake of clarity, in the following, we divide the discussion on location contribution into several key areas including primary care facilities (PCFs) which are often the first point of contact for patients and play a crucial role in ensuring community health.

Next, we investigate the strategic positioning of blood banks which is vital in ensuring a steady and accessible supply of blood and blood products.

Furthermore, we address the locating healthcare facilities in rural areas and developing nations.

In rural regions, the sparse population and limited accessibility necessitate careful planning to guarantee that even individuals in the most remote areas can reach essential healthcare services.

In the context of developing nations, the scenario often involves a high demand for healthcare services coupled with limited resources. Here, the focus should be on building a healthcare infrastructure that is not only sustainable but also accessible to everyone, ensuring that each person, regardless of their location, has the opportunity to receive the necessary healthcare. This approach promotes inclusivity and equity in healthcare provision, aiming to leave no one behind in the pursuit of better health and well-being.

Primary care facilities

Ensuring accessibility to healthcare facilities for all population groups is a main goal for PCFs. Many of these facilities operate continuously, allowing individuals to find assistance at the nearest available center. Indeed, the strategic placement of PCFs has been a sustained focus of research during the last thirty years.

In [204], the authors investigated the selection of public service locations for populations exhibiting seasonal mobility. They developed a model to determine the optimal number and locations of primary healthcare units to accommodate these fluctuating needs. The model considers various factors like shifts in population density, geographical changes, and the availability of healthcare resources. The goal is to improve service delivery and reduce differences in healthcare access.

In the study conducted in [1], is developed a comprehensive tool for strategic planning of vaccination clinic locations during a flu outbreak. This tool aims to optimize the allocation of resources and ensure accessibility to vaccination services, thereby mitigating the impact of the outbreak and enhancing public health outcomes. The study provides insights into the various factors and considerations that influence the selection of clinic sites, including population density, geographic accessibility, and available healthcare infrastructure.

In the study proposed in [297], the focus was on determining optimal locations for placing preventive healthcare facilities to maximize utilization by the population. The researchers utilized data from Georgia, USA, and Montreal, Canada to conduct this analysis. The aim was to identify locations where the facilities would be most accessible and beneficial to the largest number of people, thereby promoting preventive healthcare practices and improving overall community health.

In [75], the authors focused on optimizing the accessibility of treatment for soldiers suffering from brain injuries. They developed a strategy to identify the most suitable locations for establishing new treatment centers, with an emphasis on minimizing costs. The objective was to ensure that affected soldiers could receive timely and effective treatment without imposing a significant financial burden on the healthcare system.

In the research conducted in [52], the authors aimed to enhance the efficiency of the transplant system in Italy. They developed a model to facilitate this improvement, taking into account crucial factors such as the time sensitivity of transplant procedures and the geographical locations of the transplant centers. The goal was to optimize the allocation of resources and ensure timely access to transplant services, thereby contributing to better patient outcomes and overall system effectiveness.

In the studies presented in [175] and [133], the focus was on enhancing healthcare accessibility for low-income individuals. Both research teams worked on strategizing service planning by considering the geographical locations of both the healthcare providers and the patients. The objective was to optimize the distribution of healthcare services, ensuring that people with limited financial resources could access the necessary medical care, thereby addressing disparities in healthcare accessibility and promoting equitable health outcomes for all.

In [280], the authors investigated strategies for planning the locations and sizes of hospital departments. They utilized real data sourced from German hospitals to conduct this analysis. The aim was to optimize the spatial arrangement and capacity of different departments within hospitals, thereby enhancing operational efficiency and ensuring that the varying needs of patients are effectively met.

In the study conducted in [72], the authors focused on hospital planning in Hong Kong. They developed a model aimed at aligning available healthcare resources with the actual needs of the population. The objective was to optimize the allocation of medical services and facilities, ensuring that the healthcare system is well-equipped to meet the diverse and evolving healthcare demands of the community.

In [318] the research team concentrated on developing a system aimed at enhancing the utilization of preventive healthcare facilities. The goal was to design a system that encourages more people to advantage of preventive healthcare services, thereby promoting early detection and management of health conditions and contributing to improved public health outcomes.

In the work presented in [190] is presented a novel approach to redesigning a healthcare facility network, involving location and service allocation. It combines system dynamics and mathematical programming. The process uses system dynamics simulations to forecast patient demand scenarios, then the information is passed to a two-stage stochastic facility location model. The first stage determines service types for health centers, while the second stage assigns patients based on scenarios. The study applied this approach to relocate health services in Southern Greece network of 32 primary care centers and 13 hospitals.

In the study conducted in [188] is addressed the Syrian crisis, emphasizing its complex and devastating nature. The authors primary focus is on identifying suitable locations for PCFs within Idleb Governorate, Syria, given the dire humanitarian situation. They draw upon data from 23 sub-districts and 338 communities to construct a mixed-integer weighted goal programming model, which is integrated with Geographic Information Systems (GIS) for optimization. Their objective was to establish 60 PCFs from a pool of 77 candidate locations.

In [183], the authors addressed the challenge of locating primary health care centers while integrating outpatient medical services and complementary services like nutrition counseling, dental care, mental health, clinical analysis, and imaging. Their goal is to optimize facility placement or upgrades while maximizing complementary service coverage and adhering to outpatient service travel distance constraints within a budget. They model the problem as a variant of the maximal covering location problem, introducing additional constraints. The study includes a case study involving the public health care system in the northern zone of Mexico, exploring the trade-off between travel distance and budget.

In [92], the authors integrated a GIS with mathematically weighted public demographic data to identify potential sites for primary healthcare facilities in Australia. They weighted the index of relative socio-economic advantage and disadvantage with data on the resident population and those not in the labour force to create an index of priority. Incorporating this index into GIS, they produced a map that ranked areas based on relative need. Their findings show that combining mathematical methods with GIS offers advantages over traditional site selection methods.

In the work proposed in [7], the authors investigated the critical role of health post networks in advancing public health by delivering essential primary healthcare services. They stressed that the efficacy of such networks is based on various factors, including the number and strategic placement of health posts. The quality of service can be amplified by managing congestion levels at these facilities. A significant policy recommendation proposed is the integration of a mixed workforce comprising both flexible and dedicated servers. The authors model the queuing system at each health post using multi-class M/G/ queues that provide multiple types of services. Moreover, they modeled the problem using an integer nonlinear programming formulation and subsequently linearize it for resolution. The applicability of their model is demonstrated through a hypothetical case study.

Blood banks location

The studies [218] and [219] provide a comprehensive review of the blood banking system. The authors explore the different facets of blood banking, highlighting the best practices for collecting, storing, and delivering blood. Their aim was to enhance the efficiency and reliability of the Blood Supply Chain (BSC), ensuring a consistent and safe supply of blood for medical needs and emergencies. In the following of the chapter, we will refer to ”blood centers”

(BCs) as facilities responsible for both collection and processing activities, while ”blood stations” (BSs) will denote facilities only engaged in collection activities.

In this context, further investigation on the topic makes it clear that location optimization is a recurrent and central theme in numerous studies, highlighting its critical role in enhancing the effectiveness of health-related systems, including blood banking. Specifically, the challenge of identifying optimal locations for blood banks, taking into account their proximity to hospitals and transportation logistics, emerges as a prominent area of research across various works.

In particular, in [66], the authors tackle the challenge of determining the best locations for blood banks by analyzing the proximity to hospitals and transportation logistics. Their main goal was to optimize the distribution system to hospitals, aiming to reduce both logistical and operational costs. To address this challenge, they formulated a deterministic, single-period, multi-objective model. This model is validated using small numerical examples to ensure its accuracy.

In [253], the authors tackled the practical challenge of organizing blood services across regions in Turkey. They focused on a hierarchical system that included top-tier facilities named Regional BCs and two secondary facilities: BSs, and mobile units. They approached this issue by applying a series of deterministic, single-period mathematical models.

In [99] is explored a distinct challenge and developed a mathematical model to determine the best locations for centrifuge centers, which handle blood separation, as well as the distribution of clinics responsible for blood collection. The main objective was to boost the overall profit for a health management organization in Israel.

In the study by [67], is formulated a deterministic model for a single period to identify the best locations for two kinds of collection facilities and determine their optimal distribution to processing facilities. The aim was to minimize transportation expenses and maximize the volume of blood both donated and processed, all while adhering to a budget. This model was tested using a case study from the Thai Red Cross, utilizing the 76 capitals of Thailand as potential sites, where the volume of donated blood at each potential site was assumed to be known in advance.

In [53] has been addressed the challenge of re-configuring an existing network of BCs, which handle both collection and processing activities. The potential actions considered were downgrading a BC (focused on collection and processing activities) to a BS (focused only on collection) or shutting it down entirely. The authors proposed a deterministic, single-period model which incorporated decisions about donor allocation and the distribution from BSs to BCs. The main goal was to reduce both location and transportation expenses. This approach was tested on a case study in Italy.

Instead, in [233], the authors addressed the problem from a uncertain perspective, developing a robust optimization model. Indeed, the model takes into account uncertainties related to blood demand at hospitals. Moreover, it also incorporates both inventory costs and the costs associated with blood shortages at hospitals. The model effectiveness was assessed using a case study centered around the city of Teheran, similarly to the one mentioned in [142]. This study highlighted 22 donor groups, each representing one of the city 22 districts. These districts were also considered as potential sites for Blood Facilities.

Continuing along the theme related to uncertainty, in the study proposed in [17], is introduced a two-stage stochastic programming model to determine the optimal locations for two facility types: BSs and BCs, coherently with the definition proposed above in this chapter. To model the problem the authors proposed a discrete multi-period planning time-frame. The first-stage strategic decisions pertained to the location of the facilities and the related allocation decisions. Meanwhile, the second-stage decisions, which is dependent on specific scenarios, considers the shipping the gathered blood units, accounting for the initial stage allocation determinations. The model efficacy, and the efficiency of the implemented accelerated Benders decomposition method were proved using a case from the East Azerbaijan Province.

The study presented in [254] introduced a robust fuzzy mathematical model designed to optimally locate BSs with capacity constraints across multiple periods. A main feature of this work was the incorporation of a multi-attribute group decision-making approach, which considered both qualitative and quantitative aspects in the network design challenge. For empirical validation, the authors applied their model to the compact case of Tabriz city in Azerbaijan, which comprised ten donors sites and fewer than ten prospective locations for blood facilities.

In [195] is presented a multi-period model designed to determine the optimal locations for BSs and BCs facilities, taking into account the unpredictable demand from hospitals. The model aimed to optimize two primary objectives: (i) minimizing the overall costs, and (ii) reducing the total time spent on activities like transportation and processing within the BSC. The primary solution methodologies is built upon an evolutionary genetic algorithms and the model effectiveness is tested through numerical analysis on randomly produced data sets.

In the study proposed in [13], the emphasis was on structuring a BSC network that integrated donors, BSs, BCs, and hospitals. Notably, the study did not addressed location-related decisions. Instead, the spotlight was on determining optimal allocation and inventory strategies. The objective was to concurrently enhance environmental, social, and economic outcomes, all while taking into account uncertainties in both the supply (from donations) and demand fronts. Building on a similar problem framework, subsequent research, as highlighted in [193], primarily emphasized the sustainability aspects. This was especially in connection to the closed-loop characteristic of the BSC under investigation.

Finally, in [285], the authors developed a single-period model for the BSC network design, accounting for uncertainties in donations, demand, and blood disposal rates. Their goal was to determine the optimal locations for three facility types: BSs, BCs, and regional hospitals. The primary objective was to reduce the cumulative costs associated with facility location, transhipments, and blood wastage. The spatial distribution of donors and their respective allocations were not a focus of their study. To tackle this challenge, the authors proposed an interactive possibilistic programming method. The model applicability was verified using a Teheran-based case study.

Healthcare facility location in rural areas and developing countries

Several studies have also focused on the specific needs of rural areas and developing countries, proposing solutions that take into account the unique challenges faced by these regions.

In the study by [230], the researchers addressed a comprehensive analysis to identify the most suitable locations for establishing healthcare facilities in rural settings, employing a data-driven approach. They utilized detailed data from Bangladesh, taking into account various factors such as population density, existing healthcare infrastructure, and transportation networks.

In the research presented by [274], the authors focused on the challenges of establishing community healthcare facilities in rural areas of developing countries. Using a combination of models and analytical methods, they assessed factors like population density, transportation access, and existing infrastructure. Their aim was to determine the best locations for healthcare facilities that would be both effective and sustainable over time.

In the study proposed in [198], the authors focused on enhancing the accessibility of health services by upgrading transportation routes to all-weather roads. This was particularly crucial in regions where adverse weather conditions often hindered access. Their research is based on a real case study in Ghana, highlighting the importance of reliable transportation infrastructure in ensuring consistent access to healthcare services.

In the study presented in [137], the focus moves to devising strategies for mobile health services. The goal was to create a system where these services were both easily accessible to the public and sustainable in the long run. To validate their proposed strategies and methodologies, the authors utilized data from Ghana as a real case study.

In the work addressed in [3], the authors investigate the under-utilization of PFCs services by rural populations in Nigeria, despite the uniform distribution of these services across local government areas (LGAs). The study exploit the overarching issues affecting PFCs uptake and puts forth strategies to increase the use of these services by rural individuals. The researchers highlight the need for PFCs policy makers and LGAs to acknowledge their role in the current low utilization rates and advocate for the creation of a new social order that emphasizes equity and human dignity.

The work proposed in [284] presents a quantitative method to determine optimal locations for new primary health care facilities to maximize population accessibility. Using a GIS, the authors assessed travel times for 26,000 homesteads in Hlabisa, KwaZulu-Natal, South Africa. This data was transformed into a map highlighting areas with the highest need for clinics. Their findings revealed that a clinic in the identified optimal location would improve accessibility by 3.6 times compared to the latest built clinic.

The research presented in [274] exploit the planning of community health schemes in rural areas of developing countries, specifically those initiated by non-governmental or faith-based organizations. The study evaluates planning from both overarching (top-down) and grassroots (ground level) perspectives, concluding that a blend of both approaches is essential for the sustainability of these health initiatives. Emphasizing the top-down strategy, the authors introduce hierarchical models developed specifically for positioning community health facilities, aiming for both efficiency and equitable distribution. Furthermore, the study showcases a real case study centered on determining optimal locations for self-sustaining primary healthcare workers in a rural part of India.

In this work proposed in [261], the primary focus is on enhancing healthcare accessibility in rural areas through the delivery of mobile healthcare services. These services involve doctors visiting remote villages lacking nearby healthcare facilities. The objective is to determine how doctors should be assigned to villages, create their monthly visit schedules, and establish the base hospitals from which their tours begin and end. The study frames this problem as a periodic location routing problem, drawing inspiration from the policies outlined by the Ministry of Health of Turkey. Additionally, the study introduces a heuristic algorithm that significantly improving the efficiency of solving larger instances of the problem.

In the study presented in [191], the authors address the unique challenges of locating facilities in emerging markets. They present a modified version of the set covering model tailored to these specific attributes. The model is built and verified using a real-world case study from the Moroccan Ministry of Health, which is revisiting its pharmaceutical product supply chain design. The study underscores that factors like road infrastructure and the spread of demand play crucial roles in strategic decisions about facility locations in emerging markets.

The collective findings from these studies underscore the critical significance of strategic placement and planning of healthcare services. Proper location and infrastructure can greatly enhance accessibility and efficiency, ensuring that communities, especially those in under-served areas, receive timely and effective healthcare.

Emergency vehicles

Emergency Medical Services (EMS) play a fundamental role in contemporary health systems, as highlighted [34]. These organizations are responsible for the crucial pre-hospital phase, including medical care and transportation from the moment an emergency call is received until the patient is transferred to a hospital. The efficiency with which they respond can have severe implications for patient outcomes.

To ensure they provide the population with an adequate level of service, EMS strategically positions a certain number of ambulances across their operational territory. This strategic positioning, known as the static ambulance location problem, involves determining the optimal standby sites and deciding the number of ambulances at each site, all while adhering to specific constraints. It is given that, after completing a mission, an ambulance would return to its designated standby site based on this pre-decided location plan.

The interest in the static emergency vehicle location problem has been significant, as highlighted in both [51] and [14]. Over time, the models addressing this problem have evolved and can be broadly categorized into three types: single coverage deterministic models, multiple coverage deterministic models, and probabilistic and stochastic models. These models have been refined to account for real-world challenges, such as unpredictable demand, availability of vehicles, and traffic congestion. Comprehensive surveys on these models have been presented in [239, 51, 29, 34].

In the context of single coverage deterministic models one of the foundational mathematical formulations in this domain, in particular the Maximal Covering Location Problem (MCLP) was introduced in [287]. In their work the authors proposed the concept of formulating the emergency vehicle location problem based on a coverage criterion. Their model was designed with an aim to minimize the number of vehicles while ensuring all demand zones received adequate coverage.

The MCLP, was further explored by various researchers.

In the study proposed in [98], the MCLP was employed to address a practical challenge: determining the optimal locations for ambulance deployment in Austin, Texas. This application underscored the real-world relevance of the model in emergency service logistics.

Meanwhile, in [113] the MCLP is exploited in a more abstract and methodological fashion. Indeed, the authors introduced a Lagrangean heuristic to refine and enhance the solutions generated by the MCLP.

Taking a different angle, [264] expanded upon the foundational MCLP model to accommodate EMS that operate with two distinct types of vehicles. This adaptation recognized the diverse needs and resources of EMS operations and aimed to provide a more versatile and comprehensive model for such systems.

Nevertheless, deterministic single coverage models work on the assumption that a vehicle is always ready when an emergency call comes in. But in reality, this is not always true. For instance, if two calls come in quick succession from areas covered by the same vehicle, the vehicle might still be attending to the first call when the second one arrives. Using just these single coverage models can lead to solutions that do not hold up well in real-world scenarios.

To address this issue, were developed multiple coverage models. Their goal is to increase the chances that an area has at least one available vehicle by increasing the total number of vehicles that can respond. These models are a step forward from the single coverage ones because they factor in the unpredictable nature of emergencies by focusing on vehicle availability.

In this context, in [80] introduced the hierarchical objective set covering problem (HOSC), which was the first to really dive into multiple coverage. The HOSC aims to use the fewest vehicles to ensure every area is covered at least once, and then to maximize the number of vehicles that can respond to an area. However, this model is not without its limitations. For instance, it may not be practically feasible to assign more than two vehicles to cover an area if the chances of both being busy at the same time are low. Indeed, in [98] and [138] both proposed models to address these issues.

On a related subject, in [115] is introduced the double standard model (DSM), which combined the idea of double coverage with varying coverage distances. This model was applied in real-world scenarios in cities like Vienna, Austria, and others in Canada, Belgium, and Shanghai, China.

In [166] and in [166] is expanded the DSM to accommodate different vehicle types and priority levels, testing their models with data from Chicago, USA.

In [279] the authors put forward a goal programming model, aiming to position a set number of vehicles to both minimize uncovered demand and maximize the areas covered by more than one vehicle.

While multi-coverage models offer advantages over single coverage ones, they come with their own set of challenges. For example, having double coverage in an area does not necessarily guarantee excellent service. Moreover, if the demand is low, such extensive coverage might be excessive. To tackle these challenges, researchers began exploring probabilistic and stochastic models that more directly account for uncertainties.

The first set of these models, called expected covering location models, aim to find vehicle locations that maximize expected coverage. This expected coverage often takes into account whether a vehicle is available.

As pointed out in [102], considering uncertainties in these models can offer clear benefits. Early models, like the one proposed in [78, 79], integrated vehicle availability. However, as observed in [30], there were instances where these models did not align well with actual real-world outcomes. To bridge this gap and enhance the accuracy of predictions, the authors proposed specific modifications and refinements to the existing models.

Most models so far assumed travel times were consistent. But in reality, things like traffic can make travel times vary. For instance, in [77, 118] the authors took this aspect into account in their models. In addition, they also considered different issues, as dispatch priorities. Later models, like the one proposed in [140], added more aspects to the problem, like the time necessary from getting a call to dispatching a vehicle.

Over the years, various models have been developed for EMS organizations that deploy a diverse range of vehicles. Notable examples include the studies presented in [171] and [182]. Furthermore, the research outlined in [238] introduces an innovative model designed to significantly reduce missed calls. This model employs the Erlang loss formula, a proven method, to accurately estimate the probability of missing a call. Building on this foundation, other researchers, as exemplified in [291], have refined and expanded upon these existing models, adapting them to encompass larger geographical areas with increased efficiency.

Chance-constrained programming offers an alternative methodology to address uncertainties in the location of emergency vehicles. Distinct from expected coverage models, this approach employs probabilistic constraints, prioritizing system reliability. Pioneering works in this domain, such as those presented in [240] and [241], laid the foundation by introducing this innovative approach. Subsequent models, notably those detailed in [173] and [174], integrated queueing theory, providing a more nuanced and precise understanding of system performance dynamics.

Additionally, certain models, such as the one presented in [22], adopted a probabilistic approach to encapsulate the unpredictable nature of emergency calls. Within these models, it is posited that each call adheres to a specific probability distribution. Building on this foundation, [22] subsequently proposed a refined variant of the original model.

1.2.3 Capacity planning

Hospital capacity planning refers to the strategic process of ensuring that a healthcare facility has the right amount of resources, including beds, staff, equipment, and infrastructure, to meet the current and future healthcare needs of its patients. This involves forecasting patient demand, analyzing current capabilities, and making informed decisions to optimize the utilization of resources. Effective capacity planning ensures that hospitals can provide timely care, reduce patient wait times, manage costs, and adapt to changing healthcare environments, such as sudden disease outbreaks or demographic shifts. It is a critical component in maintaining the efficiency, quality, and sustainability of healthcare services.

In this context, hospital capacity planning introduces a range of complex challenges. A study presented in [123] explored these challenges through Operational Research-focused analyses. They utilized a queuing model to address the urgent decisions associated with hospital capacity planning. Their research highlights the crucial role of Operational Research models in deriving essential insights and shaping effective operational strategies.

In the same year, a novel approach to address hospital capacity for emergency patients in the Netherlands was introduced in [123]. This approach was inspired by the overflow model commonly used in telecommunications for phone calls, leading to the development of a mathematical method. To validate the robustness of their findings were applied simulation techniques.

In another study proposed in [180], the focus was on emergency departments in the UK. Given the government directive, the goal was to ensure that 98% of patients received care and were discharged within four hours. The methodology adopted for this was a queuing model.

A distinct issue of hospital capacity planning was addressed in [167]. In this work, the authors investigated whether to increase capacity by building new operating rooms or to extend the working hours of the current ones. To such aim, they developed a multi-objective model with three main objectives: reducing scheduling wait times, enhancing the reliability of procedure start times, and maximizing hospital profits. Their research thoroughly examined the trade-offs among these objectives, and the model efficacy was confirmed through simulation techniques.

On a tactical level, a study proposed in [4] focus on optimizing resource utilization at a cardiothoracic surgery center. The challenge was approached by developing as a Mixed Integer Program (MIP) with variable lengths of stay. The results showed that to meet a particular patient throughput goal, master surgical schedules could be improved by integrating a stochastic variant of the deterministic challenge.

In the study presented in [208], the goal was to figure out the fixed hospital costs per unit of measurable output. The study further tried to identify the root causes of inefficiencies. To achieve this was devised a stochastic multilevel model and applied to hospitals in Portugal. The model highlighted various inefficiencies, encompassing challenges related to bed distributions, incentives, and economies of scale. Moreover, the model discerned the cost disparities across diverse hospital categories.

In [169], the authors emphasize the need for efficient utilization of care provider capacity and informed capacity planning to guarantee timely patient care. On this basis, they conducted a simulation and optimization study at a regional cancer center in British Columbia, Canada. This study assessed the combined effects of available consultation slots, appointment scheduling policies, oncologist specialization on patient access and physician workload. The paper primary contribution is its methodological framework, guiding specialty clinic managers in resource utilization and strategic capacity planning for incoming patient demand in the short to mid-term.

In the study presented in [295], is proposed a discrete event simulation method to examine the surgery waiting lists in Nova Scotia, Canada. Recognizing the challenges posed by increasing patient demand and limited resources, the authors created a model specifically for capacity planning. This model considered both the current state of hospital resources and anticipated future demands. Through a detailed analysis of various performance metrics, they identified some bottlenecks and areas for improvement. Their insights offered actionable recommendations to enhance patient flow and reduce waiting times. Their comprehensive approach aimed to optimize resource allocation, ensuring timely care for patients while maximizing hospital efficiency.

In the study presented in [16], the authors explore tactical capacity planning in healthcare, emphasizing the challenge of matching physician availability with unpredictable appointments demands. To this aim, they use the cardinality-constrained robust optimization to create plans that can handle uncertainties. The study takes into account both first-time and returning patients in outpatient settings, each having specific access time targets. The results show that clinics can ensure 100% feasible plans without being overly conservative, balancing costs and demand effectively. The model also identifies critical periods with potential maximum physician workload.

In [164], the authors challenge the conventional approach of treating hospital units in isolation for scheduling and capacity planning. They introduce a Markov decision process model for daily surgical patient scheduling, considering post-surgery stays and inpatient issues. They identify optimal patient admission strategies based on surgical wait-lists. Simulations using real data highlight the advantages of an integrated scheduling approach, especially when balancing demand with system capacities. The authors also demonstrate that traditional scheduling, focused solely on operating room usage, can lead to inefficiencies and increased costs.

In [162], the authors address the critical issue of capacity planning for long-term care services, essential for the elderly, disabled, care providers, and health policymakers. Recognizing that patients often transition between settings like nursing homes and home-based services, they model patient flows using an open migration network. This is combined with a newsvendor-type model to plan the required capacity. The study identifies structural properties and key factors, such as penalty costs for capacity shortages and transition rates between care settings, that influence capacity decisions. This systematic approach considers varying patient flow patterns and budgetary constraints.

In the study presented in [156], the authors address the critical role of nurses in neonatal intensive care units (NICUs) and the challenges of determining optimal staffing levels due to the unit stochastic nature. Factors such unpredictable patient arrivals, rejections, transfers due to capacity constraints, and varying lengths of stays necessitate a deep understanding of optimal nurse requirements. The paper introduces a stochastic approximation, grounded in the nurse-to-patient ratio and patient count in a teaching hospital NICU. The approach began with creating a model to simulate outcomes under different nurse counts. This simulated data then informed a statistical regression analysis to discern the relationship between inputs (nurse numbers) and performance metrics (admission rates, occupation rates, and satisfaction rates). The study also proposes several integer nonlinear mathematical models to identify optimal nurse capacities, considering multiple performance targets. The authors applied this approximation to a large hospital NICU, analyzing the results to validate their approach.

In [283], the authors address the challenge of allocating limited operating room capacity among hospital sub-specialties, especially when dealing with both elective and emergency surgeries. They introduce an adjustable robust model to manage operating room allocation under uncertain demand, aiming to minimize potential revenue losses from operating room shortages. The study evaluates the impact of the model conservativeness on revenue and provides insights for setting model parameters. Using an implementer-adversary algorithm, the authors compare their robust optimization approach with a scenario-based stochastic optimization. Their findings suggest that with appropriate adjustments, the robust solution can closely match the results of the stochastic approach while also limiting worst-case outcomes.

In the work proposed in [276], the authors addressed the challenges faced by the United States outpatient service sector, which is experiencing a rising demand for services but a projected decline in available physicians. In the study, are highlighted major inefficiencies, such as shortcomings in the appointment system and the issue of patient no-shows. In this context, the authors propose a prescriptive analytics framework, leveraging patient-related data, to create predictive models identifying the likelihood of a patient no-show. By integrating patient-specific no-show risks, they introduce new scheduling rules. The viability and efficiency of this framework are demonstrated using real data from a Pennsylvania Family Medicine Clinic. When compared to three other scheduling rules from existing literature, the proposed rules consistently showed superior performance across various clinic settings.

In the research presented in [268], the authors emphasize the challenges Home Health Care (HHC) companies face in addressing patient needs promptly due to uncertainties in caregivers schedules. Most of the literature focuses on deterministic models that overlook these uncertainties, potentially resulting in less resilient real-world solutions. To address this, the authors introduce a robust optimization model for HHC that considers uncertain travel and service times. The model is tested, using off-the-shelf solvers and a metaheuristic framework, such as: Simulated Annealing, Tabu Search, and Variable Neighborhood Search.

In [157], the authors address the challenge of scheduling medical appointments while considering varying patient no-show behavior throughout the day and unpredictable service times. Drawing from international datasets, they noticed a distinct time-of-day effect on patient attendance. To address this issue, they introduce a robust model to minimize the worst-case scenario of combined patient wait times and provider idling. The model complexity derive from the simultaneous uncertainties in its objective and constraints. The authors proposed both deterministic and varying no-show rates, with increasing level of complexity for the latter. By employing dual prices iteratively, the authors find an efficient scheduling solution.

In the work presented in [107], the authors explore resource planning strategies to improve health systems responses during epidemics and pandemics. They focus on integrated healthcare resource allocation, sharing, and patient transfers, aiming to ensure patient care access with minimal capacity expansion. Recognizing the time-sensitive nature of pandemics and the uncertainty surrounding patient numbers, they introduce a multi-stage stochastic program to optimize healthcare resource planning strategies. The study employs an agent-based continuous-time stochastic model to simulate uncertainties, which are then captured using a forward scenario tree construction approach. To aid real-time decision-making, is proposed a data-driven rolling horizon procedure, addressing some stochastic programming limitations and ensuring practical strategy implementation. Using two COVID-19 case studies, the authors validate their optimization and simulation tools.

Finally, in the work introduced in [50], an entire chapter was devoted to the topic of capacity planning and management in hospitals. This field has attracted considerable attention and research. Significant contributions include the studies in [275], [155], [134], and [242]. Each of these works has notably advanced the understanding and methodologies in bed capacity planning.

1.2.4 Vehicle routing

The integration of the Traveling Salesman Problem (TSP) into healthcare, particularly in HHC, has emerged as a crucial element in enhancing the efficiency and quality of health services provided to patients in their homes. Ineed, a considerable body of research has been conducted on this subject. The studies presented in [104], [87], and [73] offer detailed insights into this application, thoroughly examining the challenges, methodologies, and prospective paths for the future of this integration.

On this basis, in [73] is discusses the Home Health Care Routing and Scheduling Problem (HHCRSP) as an extension of the traditional vehicle routing problem (VRP). The HHCRSP involves designing routes for care workers to provide home treatments to patients in a specific geographic area. The paper investigate into the unique side-constraints that make this problem particularly challenging and provides an exhaustive overview of recent Operational Research models developed for HHCRSP. It categorizes the existing literature based on the formulation of constraints and objective functions, offering insights into the methods developed to solve HHCRSP and discussing potential future research avenues.

Building on this, the work presented in [87] offers an extensive literature survey on routing and scheduling problems in HHC. It underscores the importance of delivering a quality of service comparable to hospital care, while managing costs and improving patient living conditions. The paper focuses on the methods utilized in addressing these problems and emphasizes the constraints and objectives specific to HHC. Notably, it highlights on the uncertain and dynamic aspects prevalent in HHC, paving the way for identifying new research directions and providing guidance for both researchers and practitioners in the field.

Complementing these perspectives, in [104]is addressed the issue to optimize the daily routes and schedules of medical and social staff to ensure effective patient care at home. This optimization is not only crucial for maintaining the quality of care but also plays a significant role in controlling the associated costs. The review in this work presents a detailed taxonomy of the problem characteristics and constraints, summarizing decision-making solutions and highlighting related objectives.

In [172], is addressed the problem of daily planning for HHC. The model proposed considers the service requirements of each patient, the individual qualifications of the healthcare staff, and possible interdependencies among various service operations. These interdependencies may include time-sensitive tasks like administering medication before meals or activities requiring multiple staff members, such as assisting disabled patients. Additionally, the model incorporates the time preferences of patients through designated time windows. To tackle this problem, the paper proposed a planning approach that aims to optimize both economic and service-oriented performance metrics. It includes a detailed mathematical model formulation and introduces a powerful heuristic based on an advanced solution representation, designed to navigate the intricacies of home health care scheduling and resource allocation efficiently.

In [205], is introduced a multi-objective mathematical model to solve the Healthcare Inventory Routing Problem focusing on the distribution of medicinal drugs to healthcare facilities. The model addresses three primary objectives: minimizing total inventory and transportation costs, maximizing satisfaction by reducing forecast error due to drug shortages and expired drugs, and minimizing greenhouse gas emissions. To mitigate the risk of drug shortages, a demand forecast approach is integrated into the model. To address uncertainty, is utilized a hybridized possibilistic method. Furthermore, an interactive fuzzy approach is employed to resolve an auxiliary crisp multi-objective model.

In [296], the authors address a simultaneous facility location and vehicle routing problem in the healthcare sector of the Netherlands, focusing on optimizing medication delivery. The study explores the strategic use of lockers as delivery points and the planning of efficient delivery routes for home deliveries, with an aim to minimize both routing and locker opening costs. The problem is systematically defined and approached using a branch-and-bound algorithm and a novel fast hybrid heuristic. Extensive computational tests, including scenarios from Alliance Healthcare Netherlands, demonstrate the heuristic robust performance. The research offers valuable insights for optimizing healthcare logistics and improving medication delivery systems.

In [106], a bi-objective optimization methodology is proposed to address the multi-period and multi-depot home healthcare routing and scheduling problem in a fuzzy environment. This study incorporates patient satisfaction under uncertainty and employs the Jimenez fuzzy approach. It introduces the application of the social engineering optimizer (SEO) and a new modified version, the adaptive memory strategy-enhanced SEO (AMSEO), in healthcare logistics. In addition, the work proposes a comparative analysis of these algorithms based on multi-objective metrics and sensitivity analyses, highlighting the practicality and efficiency of AMSEO in this context.

In [20], is proposed a bi-objective mixed-integer nonlinear programming model, integrating the periodic inventory routing problem with location decisions for effective healthcare waste management. The model aims to minimize both operating costs and risk. To tackle the problem complexity, is introduced a two-step approach, including a MILP for scheduling and a bi-objective Adaptive Large Neighborhood Search Algorithm (BOALNS) for decision-making. Computational analyses demonstrate BOALNS superior performance in finding high-quality solutions.

In [103], is presented an approach to optimizing the pharmaceutical supply chain through a multi-objective optimization model. The model considers a range of pharmaceutical types, varied vehicle capacities, and multi-period planning, ensuring adaptability to new information and emerging demands. It aims to minimize unfulfilled prioritized requests for pharmaceuticals while also optimizing transportation activities and related costs. Employing a Lexicographic method, the model hierarchically optimizes these objectives, maintaining the integrity of each while progressing. The implementation of a hybrid algorithm, which combines non-dominated sorting genetic algorithm and multi-objective particle swarm optimization, effectively addresses the model complexities. Demonstrated through various case studies, the model shows adaptability and robustness, validated further by sensitivity analysis.

In [225],is proposed a model addressing the pollution routing problem, aiming to optimize the logistics of medical unions. This model considers the use of electric vehicles, soft time window expectations of customers, open path options, and carbon costs. The paper introduces a modified Differential Search Algorithm, enhanced with comprehensive learning and dynamic Cauchy variation strategies, to tackle the vehicle routing challenge. The results demonstrate the algorithm effectiveness in reducing carbon emissions and saving logistics and operating costs for medical device distribution, affirming the model practicality in improving resource allocation and environmental impact in healthcare logistics.

In [207], the paper presents a two-stage mathematical formulation for the HHCRSP spanning multiple days. This study introduces caregiver-patient compatibility, a factor that impacts their satisfaction, as a novel aspect of the problem. The first stage focuses on minimizing the Total Planning Distance (TPD), while the second stage maximizes overall compatibility and addresses unexpected events. A health center in Tehran serves as a case study to validate the model and perform sensitivity analysis on objectives against structural parameters, yielding practical managerial insights. Due to the NP-hard nature of the problem, the authors also proposed a matheuristic-based algorithm to reduce computational time, and the results confirm the algorithm effectiveness in achieving near-optimal solutions swiftly.

In [153], the study addresses the home health care routing and scheduling problem, focusing on enhancing nurse satisfaction by minimizing the gap between their actual and potential skills. This approach aims to mitigate the hidden costs associated with skill downgrading. A mathematical model is developed for this purpose, and an ϵ-constraint-based approach is proposed for solving this bi-objective problem. The paper also includes a sensitivity analysis on the Epsilon parameter and concludes with managerial insights and future research directions in home health care management.

In [309], the study introduces the multi-objective consistent home healthcare routing and scheduling problem for home health care companies. It focuses on optimizing routing cost, service consistency, and workload balance, considering uncertain travel and service times. To address the complexity of this NP-hard problem, an improved multi-objective artificial bee colony (IMOABC) metaheuristic is developed. The study numerical experiments demonstrate the effectiveness of IMOABC, with a trade-off analysis revealing possible balances between total costs, caregiver consistency, and workload balance.

In [266], the paper investigates a supply network configuration problem for a medical implant company, motivated by a real-world healthcare supply case. This problem is a variant of the classic location-inventory-routing problem (LIRP), encompassing elements like multi-product, multi-period, multi-type delivery, delivery time constraints, and vendor managed inventory (VMI). The study considers both deterministic and uncertain demand, with two delivery methods: scheduled bulk delivery to VMI warehouses and direct shipping to hospitals. The approach involves presenting a deterministic MILP model for the integrated LIRP. To tackle demand uncertainty, a robust optimization model is proposed and transformed into a linear equivalent formulation. Additionally, a new robust model is developed to account for the impacts of the COVID-19 pandemic on demand and delivery times. In addition, a real-world case involving medical implant supply configuration for 78 hospitals is successfully solved.

In the work presented in [319], is addressed a bi-objective optimization problem to manage routing and scheduling in home health care under uncertainty. The authors extend classical vehicle routing constraints to include caregiver qualifications and workload, using a MILP formulation. An adaptive large neighborhood search within a multi-directional local search framework is proposed for computational efficiency. The model incorporates uncertain service times for patients, handled by a scenario-based sample average approximation method. Using the Solomon data set, the authors validate their approach through various metrics and sensitivity analysis, and apply it to a real-life case to provide managerial insights.

1.3 Healthcare practice

Efficient healthcare management requires addressing numerous complex scheduling and logistical challenges, all crucial for ensuring cost-effective, high-quality care. In the following, we address the most critical themes in healthcare operations. We begin with a discussion on Patient Scheduling, emphasizing aligning patient needs with healthcare availability. Next, we move into Nurse Scheduling, focusing on balancing staffing needs with work-life considerations. Then we consider the specific issues of Resource Scheduling. This naturally lead into Operating Rooms and Surgeons Scheduling, highlighting the synergy needed between available operating rooms and surgeon availability. Finally, we exploit some applications of Operational Research Logistics, looking beyond schedules to the broader movement within healthcare, from equipment transport to patient transfers.

Together, these themes paint a comprehensive picture of the intricacies of healthcare management.

1.3.1 Patient scheduling

Achieving an optimized balance between patient-staff and patient-facility schedules can lead to significant cost savings and a marked improvement in service quality.

In [216] is tackled the challenge of scheduling surgeries for patients at a Swedish hospital, taking into account specific medical, financial, and time-related constraints. In adherence to Swedish healthcare guidelines, to avert prolonged wait times, patients can be scheduled at alternative hospitals. The researchers utilized a combination of simulation and Integer Programming (IP) to address this.

The work proposed in [206] address patient-staff scheduling problem, employing a hierarchical mathematical models to generate schedules. They identified three primary sub-problems: patient selection; patient assignment to staff; and daily patient scheduling. Their objectives included maximizing the number of patients scheduled, evenly distributing workload among physiotherapists, and reducing patient wait times on treatment days. Real-world hospital data was used to test their models.

The study in [215] modeled the dynamic generation of patient schedules, which have shifting priorities, to public healthcare facilities, leveraging a Markov decision process. They employed Approximate Dynamic Programming for solving the problem, evaluating the solution effectiveness via simulation. Moreover, the same authors, [214] also harnessed simulation to enhance resource allocation for diagnostic services, focusing on adaptable inpatient scheduling.

Addressing the issue of unexpected patient appointment cancellations, [123] approached the problem as a single-server queuing problem, in which customers who are about to enter service have a state-dependent probability of not being served and may rejoin the queue. The authors have analyzed the steady-state distributions of queue size, considering both deterministic and exponential service durations. Moreover, they presented some comprehensive computational findings, alongside suggestions for potential future research.

In [145] is investigated patient flow in outpatient settings, emphasizing parallel processes to reduce cycle times. They advanced existing multi-class, open queuing network models, confirming that parallelism indeed lowers cycle times for patients requiring multiple diagnostics or treatments. Moreover, for result comparison the authors use a simulation approach. The work presented in [148] focused on outpatient appointment scheduling, developing a local search procedure that aims for an optimal schedule, taking into account a weighted average of expected patient wait times, doctor idle periods, and tardiness. The model also accommodates the occurrence of no-shows. Notably, under specific parameter combinations, the well-know Bailey-Welch rule emerges as the optimal scheduling strategy.

The study presented in [70] propose a Genetic Algorithms for patient scheduling to enhance service quality. The researchers aim to cut down patient wait times and make the most of therapy equipment utilization. They frame this challenge as a hybrid shop scheduling problem. To confirm the robustness of their solution, they apply a Mixed Integer Program and support their approach with empirical data from a medical center in Taiwan.

The study presented in [237] introduces an interactive online game based on discrete event simulation. This game illustrates decision-making in hospitals from both financial and organizational perspectives. Hospital managers can use this platform to devise strategies for patients spanning various conditions and budgetary considerations. They can also assess capacity and make patient scheduling decisions through detailed analyses of resources and processes.

The research in [117] uses a queuing model to evaluate patient welfare across different management strategies for public health service waiting lists. The researchers investigate a variety of queue management rules, focusing on both patient well-being and the efficiency of the service. Under a first-come-first-served admission policy, the authors shown a notable welfare improvement when shifting from an implicit to an explicit rationing system for access to the waiting list. If patient waiting times and hospital admissions are determined based on clinical priority, an additional boost in welfare can be achieved. This is possible without resorting to explicit rationing, simply by redistributing the cumulative waiting time – reducing it for the more critically ill and increasing it for those with less severe conditions.

The study presented in [270] employs queuing models centered on the probability distribution of bed occupancy. This model focuses on optimizing admissions to the intensive care unit, with an emphasis on maximizing life-saving potential. Validated using data from an Israeli hospital, the research suggests that by refining admission criteria to prioritize those with significant potential survival benefits, up to 18 additional lives could be saved annually compared to traditional methods.

In the study presented in [255], the authors explore the synergistic use of predictive analytics, optimization, and overbooking to manage outpatient appointments, considering potential no-shows. The research focuses on optimal overbooking strategies based on no-show predictions that take into account both individual appointment details and the specific day. A novel heuristic emerges from their analysis, suggesting that likely attendees receive same-day appointments, while probable no-shows are scheduled for future dates. Validated through extensive simulations using a dataset of nearly 50,000 real-world appointments, their findings indicate a notable improvement in clinic performance, suggesting that this heuristic is more effective than open access in most scenarios.

In the work discussed in [91], the authors addressed the emerging service model of home health care, emphasizing its potential for delivering efficient door-to-door medical services. The study underscores the importance of strategically scheduling and routing nurses to minimize costs and enhance patient satisfaction. With a focus on known demand and service capabilities, the research introduces an IP model, catering to patient priorities and time window constraints. To address the optimization challenge, a genetic algorithm coupled with local search was employed. An empirical case study based in Shanghai, China, was undertaken to validate the proposed model. The results highlight the efficacy of the model and methodology, positioning it as a valuable tool for decision-making among home health care administrators.

1.3.2 Nurse scheduling

The central challenge in nurse scheduling is to assign shifts to nurses of varying skills, satisfying as many soft constraints and personal preferences as possible. This ensures adequate personnel coverage during a specific planning period, as illustrated in [161].

In [260], the authors introduce IP methods to address the International Nurse Rostering Competition problem. Initially employing a compact formulation that proved inefficient for current solvers, they propose enhanced cut generation strategies and primal heuristics. Their extensive computational tests led to the optimization of most instances, improvement in the best-known solutions by up to 15%, and robust dual bounds. To ensure reproducibility, all implementations were carried out using the Computational Infrastructure for Operational Research.

In the study presented by [158], the authors propose a two-stage approach that balances the individual preferences of nurses with the overarching operational needs of hospitals. In the initial stage, shifts are systematically allocated to nurses who place the highest bids, ensuring that hospital constraints and requirements are satisfied. Subsequently, the remaining shifts are distributed to complete the final roster. This method was effectively applied in a real-world context, specifically within an emergency department in Pennsylvania, USA.

Expanding on the theme of nurse rostering, the research outlined in [54] introduces a dynamic, real-world nurse-rostering methodology that stands in marked contrast to the more traditional fixed-period strategies commonly employed. This novel approach offers greater flexibility, accommodating the ever-shifting demands of healthcare environments.

In [288], the authors tackle the Nurse Rostering Problem (NRP). The paper introduces a hybrid solution algorithm merging MIP heuristics, specifically Fix-and-Relax (F&R) and Fix-and-Optimize (F&O), with Simulated Annealing (SA). The MIP-based heuristics break down the problem into smaller subproblems, optimizing each one iteratively. The hybrid strategy employs the F&R algorithm to produce high-quality starting solutions for the SA component. When improvements stall, the F&O method is incorporated into SA, enabling a diversified search space and enhanced solutions. In addition, the proposed framework has been tested on 24 publicly available instances showing results better or comparable with the state of the art.

In the work presented in [228], the NRP is characterized by the task of assigning a set of shift sequences to multiple nurses over a specified period while adhering to working constraints and preferences. To this aim, the authors introduced a hybrid algorithm, merging the benefits of IP and Constraint Programming (CP). The algorithm leverages IP capacity in deducing lower bounds and then individuate the optimal solutions exploiting CP capacity of identifying feasible solutions. To further refine the algorithm efficiency, innovative strategies are employed to discern crucial details, like the computational complexity of instances and constraints, subsequently adapting search parameters. This algorithm undergoes testing on two diverse datasets comprising various problem instances.

In [229], the authors address the NRP by introducing a hybrid algorithm that blends IP and the VNS. This algorithm starts with a greedy heuristic for the initial solution, enhances it via the Variable Neighbourhood Descent algorithm, and employs IP in a ruin-and-recreate strategy to aid the search. IP is used again for final solution improvements. Moreover, a novel scoring system evaluates the solution quality. Testing on 24 known instances, the algorithm outperforms two leading algorithms and Gurobi in most cases. Additionally, 11 newly generated challenging instances are introduced for benchmarking by the research community.

In the work in [21], the authors addressed the NSP, a critical concern in healthcare due to its role in efficiently assigning nurses to shifts while considering both hospital goals and nurse preferences. Although many NSP models have approached the issue deterministically, real-world applications are often marred by uncertainties regarding management objectives and nurse preferences. Recognizing this, the study introduces a stochastic optimization model for the Department of Heart Surgery in Razavi Hospital, which takes into account the unpredictable nature of patient demand and stay durations. The Sample Average Approximation method is employed to determine an optimal schedule aiming to minimize both regular and overtime assignment expenses. Numerical tests underscore the convergence of statistical bounds and the appropriateness of the sample size for specific experiments, thereby validating the model efficacy.

In the work addressed in [306], the authors propose a mathematical representation for the NRP. In addition, given the NRP complexity, is also introduced a particle swarm optimization method to address this issue. The problem constraint structure is analyzed and utilized to produce workstretch patterns, representing the basis for a rapid initial solutions, which are enhanced by the PSO algorithm. Additionally, an ad-hoc procedure is proposed for refining PSO-derived solutions. Emphasizing fair shift allocation as the main goal, the results indicate that the enhanced PSO algorithm consistently and efficiently yields optimal solutions in real-world test scenarios.

Within the same context, some of the same authors, in [55] merges heuristic ordering strategies with the robust Variable Neighbourhood Search (VNS) algorithm to enhance the nurse-rostering efficiency.

In a separate study presented in [282], the researchers leveraged a data-driven simulation approach for optimizing nurse-patient assignments. This simulation was based on real data sourced from hospitals in Texas, ensuring its relevance and applicability to real-world scenarios.

In certain specialized medical scenarios, the unique needs of patients necessitate the allocation of nurses with specific expertise. Recognizing this complicating aspect within the healthcare delivery context, both [224] and [197] developed a mathematical programming and an heuristic framework tailored for those assignments. Their research emphasizes the critical importance of such methodologies in ensuring optimal care for an especially vulnerable patient group: critically ill newborns. These methodologies aim to ensure that the most qualified nursing professionals are matched with newborns based on the severity of their condition and the specialized care they require.

The challenge of nurse rerostering emerges when unforeseen circumstances prevent a nurse from executing previously designated duties. Such disruptions necessitate swift and efficient strategies to ensure uninterrupted patient care. Addressing this complexity, in [213] is proposed a Genetic Algorithms to develop solutions that facilitate optimal rerostering. In a complementary approach, [196] conceptualizes the re-rostering problem as a bi-objective optimization problem. To tackle this, they employed Goal Programming.

Addressing the challenge of minimizing nurse-patient switches, [277] introduced a hybrid solution that combines Constraint Programming—a with a metaheuristic search, enhancing the robustness and efficiency of their methodology. This innovative strategy ensures continuity in patient care and nurtures the trust and rapport built between nurses and their assigned patients.

Taking a broader perspective on healthcare delivery, [40] addressed a combined problem of home healthcare and nursing. Their research extend the conventional scope of nurse scheduling to address the associated routing problem inherent to home healthcare services. By doing so, they aim to achieve a dual objective: reducing transportation costs, a critical factor in resource-limited environments, and increasing satisfaction levels for both patients and nurses.

In [154], the authors address integrated staffing and scheduling under demand uncertainty using a two-stage stochastic integer program. The initial phase sets staffing and schedules, while adjustments are made in the subsequent phase as demand becomes clearer. The research highlights the benefits of mixed-integer rounding inequalities in refining solutions. Using 3.5 years of patient data from Northwestern Memorial Hospital, the authors validate their approach across 20 instances. Their results underscore the efficiency of their methods and demonstrate that their stochastic model provides significant cost advantages over deterministic models, particularly with larger forecast inaccuracies.

In [24] is addressed the multifaceted challenge of nurse scheduling requires solutions that account for both non-negotiable (hard) and negotiable (soft) constraints. Addressing this complexity, the authors proposed two approaches: a integer-programming an heuristic ones. In addition, to further improve the nurse scheduling process addressing also cyclic scheduling, where patterns recur over a fixed number of days, and preference scheduling, where nurses individual preferences play a crucial role in assignment decisions.

Further deepening the exploration into nurse rostering, in [110] is highlighted the significance of staff satisfaction—a critical, yet often overlooked, component in healthcare scheduling. Through their novel polyhedral approach, the authors showed the efficacy of certain branching rules in enhancing the efficiency of the scheduling process.

Moreover, broadening the scope to encompass multiobjective rostering, in [39] is proposed a solution approach recognizing that nurse scheduling often necessitates balancing multiple, sometimes conflicting objectives. In light of these challenges, the authors exploited the utilization of various methodologies, underling their strengths and limitations.

Moreover, as widely known, metaheuristic approaches offer a range of solutions for complex optimization problems. Among these approaches, Evolutionary Algorithms have gained prominence in the Nurse Scheduling ([9, 10, 11]), as well as Tabu Search, which has also been extensively studied and implemented by ([41, 139, 90, 89]). Moreover, another proposed metaheuristic to address the problem is the Scatter Search, as highlighted in [56]. Finally, in some works has been proposed also Ant Colony Optimisation as a promising approach ([127]).

In terms of hybrid approaches, several scholars have explored the melding of different techniques. For instance, the works presented [38] and in [290] propose innovative combinations of Tabu Search with Local Search procedures, aiming to leverage the strengths of both methodologies.

Furthermore, some studies have taken on multi-dimensional challenges. As an example, the work in [36] where is not only addresses a nurse scheduling but also personnel and facility scheduling problem. In their study, they integrated nurse scheduling with operating room assignments, employing an approach based on Linear Programming coupled with Column Generation techniques.

1.3.3 Operating room and physician scheduling

Scheduling physicians and operating rooms in healthcare institutions is of primary importance, given its significant impact on patient care, operational efficiency, and resource allocation. Optimal scheduling not only allows medical professionals to provide timely and effective care but also reduces patient wait times, optimizes the use of costly medical equipment and facilities, and ensures a balanced workload for medical staff. The outcome is a marked improvement in patient results and the overall performance of the institution.

In the works presented in [42] and [43], are examined the complexities and challenges of operating room scheduling. Their approach, built on algorithmic foundations, effectively assign specific operation times to surgeons. However, while these studies capture the broader dynamics of scheduling, they lack the granularity to identify exact start and end times for procedures. This limitation can be consequential in high-demand healthcare settings, where precise timings can impact patient flow, resource allocation, and overall operational efficiency.

In the study proposed in [32], the authors explored the problem of physician scheduling in emergency departments. They proposed heuristics based on partial branch-and-bound techniques to address these challenges. The primary objective was to minimize penalties resulting from deviations from ideal scheduling, thus improving both patient wait times and workload distribution for the medical staff.

In the research by [211], the authors introduced a Goal Programming-based method to address the operating room planning problem. Their strategy aimed to minimize inefficiencies, notably by reducing idle times and overtime in operating rooms. Simultaneously, they emphasized the importance of enhancing satisfaction among surgeons, patients, and auxiliary staff. By balancing these objectives, the approach promises improved operational efficiency and a better satisfaction for all the actors involved in the surgical process.

In the research discussed in [259], is developed an IP model to allocate surgical blocks to different specialties. This model took into consideration both the availability of operating rooms and the necessary post-operative resources. Its practicality and efficiency were recognized and subsequently adopted by hospitals under the British Columbia Health Authority. Similarly, [316] took a combined approach, integrating both IP and Simulation methods. Their model was designed to distribute weekly operating room capacities among specialties, with a primary focus on reducing inpatient costs. Both studies highlight the evolving strategies in surgical scheduling and resource management.

In [122], the authors focused on optimizing staffing in emergency departments. They analyzed patient arrival patterns and harnessed queueing models to improve departmental efficiency. Their primary goal was to diminish patient waiting times, ensuring quicker care and better patient experiences. Both studies underscore the continuous efforts to improve healthcare efficiency and patient care through refined scheduling and resource allocation.

In [192], the authors investigate the issue of part-time employment for physicians, highlighting on its advantages. Through two illustrative case studies, they demonstrate how such employment arrangements can not only enhance service quality by potentially reducing burnout and increasing physician engagement, but also optimize the broader healthcare system design by offering more flexible staffing solutions and accommodating diverse workforce needs. This analysis underscores the importance of reconsidering traditional full-time roles in the ever-evolving landscape of healthcare.

In the study by [50], the authors highlight the role of simulation in healthcare decision-making and operational efficiency. Through a detailed analysis of four distinct scenarios, they illustrate the power of simulation in providing actionable insights. These scenarios encompass inpatient surgical scheduling complexities, tracking emergency room wait times to improve patient experience, understanding the importance of emergency room scheduling, and ensuring accuracy in medicine ordering processes. Their work underscores the potential of simulation tools to transform healthcare operations and enhance patient care.

Of particular significance in this field is the Integrated Operating Room Planning and Scheduling (*IORPS*). This entails determining the date, operating room (*OR*), and start time for each surgery while also considering the estimated completion time and allocation of medical resources.

Central to the *IORPS* are two key decision-making sub-problems: the allocation of patients to designated surgical blocks (defined by the *OR* and day) and the sequencing of surgeries within these allocated blocks. These sub-problems are traditionally referred to as the Advance Scheduling Problem (*ADP*) and the Allocation Scheduling Problem (*ALP*), as identified in [317]. While numerous studies have chosen to investigate each sub-problem in isolation, for those interested in a brief overview, [59, 310] cover the *ADP* and [95, 179] shed light on the *ALP*.

However, some research endeavors have embraced a more comprehensive version of the *IORPS*, aiming to concurrently solve both the *ADP* and the *ALP*. Two predominant strategies emerge from the literature: the sequential and the integrated approach. While the sequential methodology presupposes a clear hierarchy and tackles the sub-problems one after the other, the integrated method seeks to resolve them in tandem. The limitations of the sequential method, especially given the intrinsic interdependence of the sub-problems, are outlined in [62].

In the study in [159], the authors developed a stochastic programming approach for optimizing operating room planning. They categorized surgical demand into elective and emergency types. To address the challenges posed by these categorizations and the unpredictable nature of surgeries, they employed a hybrid approach, blending the Monte Carlo simulation technique with IP. This combination offers a more straightforward and efficient scheduling system, ensuring better resource allocation and timely patient care.

In the study proposed in [217], is presented an approach to address the *IORPS* by refining the conventional Job Shop scheduling problem. This enhancement aimed to provide more efficient and realistic scheduling solutions for surgical procedures.

In the work proposed in [130], is addressed the *IORPS*, along with incorporating planned slack times. The authors introduced heuristics with objectives to curtail overtime, prevent surgery cancellations, and enhance operating room efficiency. This approach underscores the importance of balancing precision in scheduling with the unpredictability of surgical operations.

In the work addressed in [299], the authors tackled the issue of Hospital Resident Scheduling Problem (RSP) harmonizing departmental requirements, individual expertise, and personal preferences. The RSP aims to optimally align residents preferences with departmental needs when assigning night shifts over a typical 4 to 5-week scheduling period. Utilizing Genetic Algorithms, they presented a responsive and flexible solution. Conversely, [267] approached this problem by developing a MIP, enriched with heuristic techniques.

In the research detailed in [35], the authors tackled the issue of bed occupancy in hospitals, with a specific emphasis on formulating balanced surgery schedules. Leveraging MIP-based heuristics that are further enhanced with metaheuristics procedures, their approach was designed to proactively mitigate the complexities arising from potential bed shortages.

In the study presented in [286],is examined the challenge of the *IORPS* for emergency medical residents. Their specific scheduling problem is particularly complex due to a dense set of rules and constraints, which range from residents training requirements to patient care demands. Recognizing that these guidelines can often pull in opposing directions. In light of thise, the authors utilizing a multi-objective perspective, sought to harmonize these conflicting demands, striving for a solution that effectively balances patient outcomes and and hospital operational efficiency. This method underscores the importance of a holistic view addressing challenges of medical scheduling.

In [244], the *IORPS* in a Norwegian hospital is addressed in a multi-objective fashion by prioritized minimizing patient waiting times, reducing surgeon overtime, and especially reducing waiting times for pediatric patients. Their approach, which encompassed both a mathematical formulation and a meta-heuristic strategy, hinged on the weighted sum linearization of the multi-objective function.

Further adding to the complexity, [176] introduced an *ILP* model targeting a week-long *IORPS* in a Lisbon hospital. Here, the goal transitioned to optimizing the use of surgical suites in light of the surgical department equipment constraints. This approach was later expanded in [177], where a dual-objective *IORPS* was considered. Their approach not only maximized surgical suite use but also the number of scheduled surgeries, backed by heuristics informed by the hospital's real data.

Other studies, such as [298] and [46], have addressed the intricacies of the *IORPS* across varied geographical landscapes and institutional frameworks. For instance, [298] navigated the challenges at a publicly-funded hospital in the Midwest United States, where factors like patient demographics, infrastructure, and funding mechanisms play a fundamental roles. On the other hand, [46] addressed the *IORPS* within the context of a Neapolitan hospital, where cultural, logistical, and regional healthcare policies introduced distinct challenges and objectives. Both studies underscore the multifaceted nature of surgical scheduling, reflecting the requirements intrinsic to each hospital environment.

The challenge of effectively scheduling surgeries, especially in the context of emergencies, is a topic that has been rigorously tackled in recent research.

In particular, in the work presented in [101], the authors propose a MILP to dynamically adjust the schedules for elective surgeries following the sudden arrival of emergency cases. Their model considers a wide range of factors spanning from the fiscal implications of operating room overtimes and potential postponements of elective procedures to the cost of declining emergency patients. Essentially, the model provides a strategic framework to make decisions on whether to accommodate an emergent case or, reroute it to an alternative medical facility. The overarching goal of this model is to minimize associated costs from reshuffling elective patients, declining emergency interventions, or prolonging operating room operations.

In the study proposed in [25], the focus is directed towards formulating an optimal strategy for integrating emergencies into an already-set surgical schedule. The approach is based on a stochastic programming framework, considering scenarios where elective procedures might face delays, yet are not moved across different operation theaters. Furthermore, the authors have incorporated a probabilistic dimension to emergency arrivals, modeling it as a non-homogeneous Poisson process. Their findings revealed a marked contrast in elective schedule responsiveness based on two different rules: the "shortest expected processing time first" and the "best insertion method".

The research proposed in [186] investigate the *IORPS*, with particular focus on effectively planning around unpredictable emergency cases. In their approach, the authors propose a MIP framework to develop an initial surgical schedule, with a primary aim to reduce the instances of unscheduled elective surgeries. Enhancing this approach, they introduce two heuristic methodologies, seamlessly integrated with their foundational MIP structure, designed to facilitate online adjustments on the actual day of the surgeries.

In the study by [271], is examined the *IORPS* addressing for both planned and unexpected surgeries in the same operating room setup. Recognizing the uncertainties in both surgery arrivals and duration, the authors proposed a stochastic processes to estimate these variables. To make scheduling more manageable, they divided the workday into specific time blocks. Using a dynamic programming algorithm, they developed a method to adjust the schedule for planned surgeries within a day schedule.

Finally, readers further interested into the complexities of operating room scheduling, are addressed to [223], [125], and [86]. These works offer detailed analyses and methodologies for refining surgical scheduling processes. Moreover, for a comprehensive overview that spans the entire spectrum of operating room planning and scheduling challenges and solutions, [63, 226, 321] stands as an essential reference.

1.4 Specialised and preventive healthcare

The increasing life expectancy, escalating healthcare costs, and the emergence of new diseases in recent years have underscored the need for a robust focus on preventive and specialized healthcare by medical professionals. To understand the spread of infectious diseases, assess the economic advantages of screening tests for at-risk populations, and tackle issues related to organ transplantation and donation, researchers have utilized both stochastic methods and discrete optimization techniques.

The intricacies of preventive care have deepened with challenges such as terrorism and the potential for bioterrorism. Within the domain of organ donation and transplantation, most research has primarily focused on developing policies for allocating donated organs to patients on the waiting list.

Kidneys and livers stand out as the most frequently transplanted organs, capturing significant attention in medical research and policy formulation. In this context, the studies conducted in [314] and [315] provide a deeper insight into the world of organ transplantation, particularly concerning kidneys. In these studies, the author gave a wide overview of exchanges, highlighting the often overlooked issues and ethical considerations in the process of kidney transplantation. Moreover, Their approach involved evaluating allocation strategies, examining queuing theories, and utilizing simulation models. By employing such comprehensive methods, they have made a lasting contribution to the refinement of organ allocation processes, ensuring that they are both efficient and equitable.

The works presented in [265] and [251] highlights advancements in the arena of organ transplantation, with a particular emphasis on kidneys. These researchers addressed the field of technological innovation and software solutions explicitly designed for kidney donation procedures. A central concern they tackled was the multifaceted problem of non-compatible patient-donor pairs — a recurrent challenge that introduces complexities in the organ transplantation. In addressing this, their solutions were grounded in the creation of robust frameworks that supports the prospects of successful transplantations.

In the research presented by [258], the focus shifts to the issue of liver transplants. The authors provided a holistic perspective on the role of timing within transplantation procedures. They recognized that the outcome of a transplant is often highly influenced to its timing. In addition to the temporal considerations, the researchers also addressed the issue of organ quality assessment. Indeed, the health and viability of a donated organ are crucial factors in determining the post-transplant trajectory of the recipient, influencing both their immediate recovery and long-term health.

In the study proposed in [119], are addressed the challenges of uncertainty in the organ transplantation supply chain. The study introduces a mathematical model to optimize the chain, considering uncertainties in shipment times and organ demand. A possibilistic programming model, combined with a simulation-based method, is proposed for decision-making in transplant center locations and organ distribution.

Within the domain of disease prevention, optimization research prominently addresses the issues of vaccine selection as shown in [81]. In this work are highlighted the disparities in immunization access worldwide, emphasizing the challenges faced by low- and middle-income countries in vaccine distribution. The paper presents a comprehensive literature review on vaccine distribution chains in these countries and is divided into two main sections. The first section address the distinct characteristics and challenges of these distribution chains, incorporating both quantitative and qualitative studies. The second section organizes the operations research and operations management literature based on seven criteria: decision level, methodology, parts of the vaccine distribution chain modeled, uncertainties and features addressed, performance metrics, real-world applicability, and the countries and vaccines examined. A comparative analysis between real-world issues and academic investigations is conducted.

Expanding on the topic of vaccines, the study presented in [143] introduces and refines a distinctive vaccine selection algorithm. The authors emphasize the economic repercussions of vaccine wastage and champion the strategic advantage of combination vaccines, highlighting on their cost-effectiveness and potential for broad-spectrum protection.

Complementing this, other studies, as highlighted in [128] and [305], tackle diverse challenges inherent to vaccine formulation and selection. These studies collectively contribute to a richer understanding of the decision-making processes behind vaccine development and distribution, ensuring more efficient and effective strategies in combating infectious diseases.

In [163], are discussed the inefficiencies of the vaccine distribution chain in many low- and middle-income countries, supported by the World Health Organization Expanded Program on Immunization. These vaccines need a temperature-controlled "cold" distribution chain, but the current systems are not cost-effective. The study suggests redesigning this chain by determining the locations of intermediate distribution centers, flow paths from central stores to health clinics, and the allocation of transport vehicles and cold storage devices. Unlike the traditional four-tiered structure, alternative network structures can be adopted. A MIP model is introduced for optimizing small-to-medium-sized networks, and a hybrid heuristic–MIP method is proposed for larger networks. The effectiveness of these models is demonstrated using data from sub-Saharan African countries.

In [236], are emphasized the potential strain on the healthcare system due to the convergence of COVID-19 and other respiratory illnesses like the flu. Concerns about COVID-19 have heightened the vaccine shortage in developing countries, which aim to widely vaccinate their populations to lessen the burden on healthcare services.The authors proposed an inventory-location MILP model tailored for fair distribution of influenza vaccines in developing countries during the pandemic. This model prioritizes critical healthcare workers, the elderly, pregnant women, and individuals with underlying health issues. The effectiveness of this optimization model is showcased through a broad case study.

In [116], is highlighted the global urgency for vaccine access due to the COVID-19 pandemic, emphasizing the challenges faced by developing countries. This study introduces a mathematical model for a sustainable supply chain for the COVID-19 vaccine, encompassing economic, environmental, and social dimensions, ensuring both domestic and international distribution. To address the global disparity in vaccine distribution, a robust data-driven model, using a polyhedral uncertainty set, is presented. The model is validated using a real case in Iran, providing to be less conservative than conventional models when dealing with uncertainties.

In the study in [100], the author focused on the optimal distribution of influenza vaccines across diverse population subgroups. A mathematical model is presented to minimize vaccine doses and control early-stage outbreaks. An equity constraint ensures fairness in vaccine distribution. The research introduces an exact solution approach and conducts sensitivity analyses on key epidemic factors. The model scalability is explored for subgroups based on age and location. The paper also discusses an alternative optimization method and emphasizes the importance of considering group-specific transmission dynamics in vaccine distribution decisions.

On a different vein, HIV prevention remains a significant area of study, as illustrated by the resource allocation model in [96]. Moreover, several scholars have addressed into resource allocation for epidemic containment, mixing non-linear optimization techniques with epidemic models. Furthermore, the incorporation of various computational models, such as Markov models as used in [257] and [135], has not only facilitated in-depth cost-benefit analyses but also enhanced the accuracy of projections for future patient demographics.

On a different vein, contact tracing, heralded as an essential mechanism in the battle against the spread of infectious diseases, has received significant attention in academic circles. Studies like [15] have taken a deeper dive into this area, introducing a meticulous simulation-driven assessment approach that evaluates various tracing strategies and their efficacy. On the other hand, influenza has been addressed in works such as [293]. This research investigate the potential fallout and response strategies for pandemic scenarios. Meanwhile, [69] offers insights into the challenges of supply chains, specifically addressing the logistics of vaccine distribution and resource allocation during flu outbreaks.

Bioterrorism preparedness stands as a critical area of exploration in contemporary research, especially given the potential global ramifications. In the study by [149], the authors addressed into a comprehensive analysis, examining various bioterrorist attack scenarios and meticulously formulating corresponding response strategies. These strategies encompass aspects like immediate containment, medical interventions, and public communication. Similarly, the research in [2] sheds light on the practical applications of advanced modeling techniques. Here, the authors employ discrete-event simulation, a powerful tool that allows for the realistic modeling of complex systems, to optimize the planning for emergency mass dispensing. This includes the logistics and efficiency of vaccination clinics, ensuring timely and effective responses during potential bioterrorism events.

In the study proposed in [234], is discussed the importance of urban water distribution systems, especially in the context of public health and industry. Despite their significance, current emergency response models for contamination are underdeveloped. The study introduces a dynamic decision support model that adapts to real-time emergency scenarios, determining optimal health protection measures. This model considers changing contamination data and integrates feedback from various stakeholders. Its effectiveness is demonstrated using a virtual city simulation. The ultimate goal is to incorporate this model into a comprehensive system for managing contamination threats.

In [301],is emphasized the significance of developing an emergency management plan to mitigate casualties from a bioterrorist anthrax attack. In particular, this studies introduces a multi-period model that integrates disease progression, medical interventions, and logistics deployment. The model takes into account the timeline of patients transitioning through various disease stages, the initiation of medical intervention, and the subsequent recovery rate fluctuations due to time lags. The proposed model aids in strategic decision-making during an actual anthrax attack and helps in assessing key factors influencing casualty numbers.

Chronic conditions, notably diabetes and hepatitis, continue to be significant areas of medical exploration. In the domain of chronic conditions, in [132] provides a comprehensive analysis on diabetic retinopathy, emphasizing the necessity and methodologies of early detection and preventive care. Advancing into diabetes management, [97] offers innovative models specifically crafted for Type-2 diabetes intervention, aiming to optimize treatment methodologies and improve patient outcomes.

Switching the focus to hepatitis, the economic aspects of Hepatitis C screening are extensively analyzed in the research community. In the study proposed in [273] addressed the cost-benefit aspects of screening, evaluating the immediate financial implications against the long-term health benefits of early detection and intervention. Complementing this, in [278] is further investigated the screening processes, providing valuable insights into the issues of cost-effective strategies and the broader implications for healthcare planning and policy-making.

1.5 Conclusions

In this Chapter, we highlighted the importance of Operational Research in healthcare. Through systematic and analytical approaches, Operational Research helps improve decision-making processes, aiming for better patient outcomes and more efficient healthcare systems.

Our analysis underscores that three fundamental problems of Operational Research literature — namely, Location Analysis, Scheduling Optimization, and Routing — show extensive relevance and application within the healthcare sector.

On this basis, in the following chapters we investigate into each of these areas, proposing for each one a methodology to solve a real-life problem. Our aim is to offer insights and contribute to understanding how operational research can continue to make healthcare better and more accessible for everyone.

Chapter 2

Reorganization of the Blood Supply Chain *(BSC)*: deterministic and uncertainty based facility location models

In this chapter, we introduce the complexities of the blood supply chain (BSC), with a focus on the Italian context undergoing significant reforms. By introducing advanced optimization models, we aim to enhance the BSC efficiency, addressing inherent challenges such as variable demand, logistical complications, and spatial distribution.

2.1 Introduction

Ensuring rapid and safe access to sufficient supplies of blood, as well as safe transfusion processes, is a fundamental component of any strong healthcare system worldwide [304]. Blood transfusions play a vital role in disaster situations, accidents, trauma cases, emergencies, and in aiding women suffering from bleeding related to pregnancy and childbirth. In countries with sufficiently developed healthcare systems, blood and blood products also provide critical support for medical procedures like transplantation and major surgeries. In addition, they help treat immune deficiency conditions and diseases related to blood and bone marrow [304].

Human blood is considered a scarce resource, and wasting it is unacceptable, as it can lead to postponed surgeries, untreated patients, or even death [303]. However, in many developing nations, the unavailability of safe blood contributes to a high number of deaths, even in some urban healthcare facilities. Voluntary donation remains the only source of blood; however, it has some drawbacks, including limited donor numbers, delays in testing, and a high risk of product perishability [220]. While artificial blood or blood components could serve as alternatives, they are not currently feasible options and remain an area for future research and development.

In this context, effective BSC management is a crucial issue for developed and developing countries. The BSC involves six processes related to blood, namely: collection, testing, blood processing, storage, distribution, and transfusion. These interrelated activities require coordination among five key players: donors, mobile collection sites (CSs), blood centers (BCs), demand points (such as hospitals or transfusion points), and patients [220]. Thus, managing the BSC can be a daunting task due to the complexity of the processes and the need for precise coordination among the various players involved. Nonetheless, it is crucial to ensure a safe and sufficient supply of blood and blood products for transfusions, as this measure plays a vital role in saving lives and enhancing health outcomes.

Therefore, optimal decision-making in BSC management is necessary to minimize shortages and waste, as well as to design a more efficient, effective, and robust system. Accordingly, as we discuss in the next section, the literature flourishes with contributions that, despite the shared problem setting, mainly differ for the specific application context and peculiarities of the national/regional case study.
This study focuses on the Italian BSC, which has been the subject of ongoing efficiency-oriented reforms and suffers from a relatively high fragility concerning blood demand satisfaction objectives (i.e., the so-called *self-sufficiency* goal).

In Italy, the management of the BSC falls under the responsibility of the Transfusion System, a public body that is part of the Italian National Health Service (NHS) established in 1978. The Italian BSC relies entirely on voluntary and unpaid donations, as in many other countries. In 2020, there were 2,893,788 donations, comprising 2,438,349 whole blood donations and 455,439 apheresis donations. These donations were contributed by 1,626,506 donors, including 1,352,162 periodic donors and over 355,174 occasional first-time donors, resulting in 2,822,504 transfusions [64]. However, the system is vulnerable and lacks robustness when faced with fluctuations in blood demand, as evidenced by the minimal difference of only 71,284 units (2.5%) between donations and transfusions.

In addition, in 2013, the Italian government published a national guideline for regional authorities related to blood and its components in compliance with the European Directive 2002/98/EC [114]. The guideline introduced efficiency measures that require collection processing activities, such as blood analysis, separation of blood components, and transformation into plasma-derived products, to be consolidated in fewer BCs. According to the recommendation, each BC should process a minimum of 40,000 units per year. However, this target is significantly higher than the average productivity of BCs in all Italian regions. Indeed, there are 278 BCs spread across the country, and considering the aforementioned 2,893,788 donations [189], this results in an average productivity of 10,409 units per BC. Therefore, the ongoing reorganization of the Italian regional BSC emphasizes the importance of providing stakeholders with a tool to rationalize and enhance its functioning, while ensuring an adequate supply of blood and blood products.

In this context, this work focuses on the first leg of the Italian BSC, i.e., the blood collection, which is the BSC activity mostly impacted by previously discussed reforms.

The regional authorities have identified two main strategic actions to enhance BC productivity. The first action involves reducing the number of employable BCs by selecting those to be shut among the existing ones. The second consists of resorting to Mobile Units (MUs) to reach isolated blood donation points and increase processing levels at each BC. In addition, as proposed in [53] and [88], we also consider the possibility of downgrading existing BCs to the so-called Blood Stations (BS), where only collection activities are carried out. Subsequently, the blood collected at the BSs is dispatched to the BCs for processing. Since the distance to travel in order to reach the collection facility deeply influences the aptitude to donate blood, the introduction of BSs should mitigate the distance negative effect on donors' donation propensity. In the following, when no distinction is needed, we refer to both BCs and BSs as Blood Facilities (BFs).

On this basis, in a nutshell, in line with the previous literature on the topic presented in [53] and [88], the main contribution of this work comprises the proposal of a mathematical modeling framework that can serve as a decision support tool for the strategic redesign of an existing regional BSC. More precisely, the proposed framework rationalizes the BSC collection structure by closing unnecessary BCs, downgrading BCs to BSs, and supporting collection processes through Mobile Units (MUs). Such facilities, together with donors, configure a multi-echelon network system. The overall objective is to minimize the blood transportation cost (from collection to processing facilities). Moreover, specific system tactical requirements are also considered, such as maximum and minimum capacities for collection and processing activities, respectively, as well as the need to satisfy the total regional demand for blood (i.e., the self-sufficiency goal). The latter are formulated as soft constraints, and their violation is accounted for in the objective function by adopting a penalty-based method. This method allows the evaluation of a range of different trade-off solutions between the minimization of BSC transportation costs and compliance with the above requirements. Finally, donors' accessibility is also envisaged. Specifically, we ensure that the average distance potential donors should travel to reach the facilities is below a maximum threshold. Based on this design choice, the rationalization of the existing network is expected to maintain the capillary presence of blood facilities over the territory.

The proposed modeling framework has been implemented in two different perspectives: case-based and scenario-based. The first one is designed to reorganize a BSC in situations characterized by known, steady, and regular donations. Conversely, the second formulation is conceived to address the uncertainty in yearly donations and their potential variations.

These two perspectives enable to perform a twofold investigation enriching the understanding of the problem and enhancing the quality of the decision-making process. On the one hand, the case-based formulation allows to perform a sensitivity analysis with respect to the donations' propensity and the system average accessibility, which represent the main factors affecting the strategic decision under investigation. On the other hand, the scenario-based formulation allows to handle inherent system uncertainty issues by designing BSC solutions which are robust with respect to the yearly donation propensity, accounting for expected value and worst-case scenarios. Both formulations, being extensions and modifications of the one presented in [88], represent a methodological advancement in the field. In particular, they introduce explicitly the accessibility issue and configure mixed-integer linear programming (MILP) models for multi-echelon facility location problems with soft constraints that allow for considering the system targets in a multi-objective fashion. Moreover, strengthening valid inequalities are also proposed to improve the computational performance when dealing with the solution of large size instances by off-the-shelf optimization software.

The proposed formulations are tested using real data from three Italian regions (i.e., Campania, Apulia, and Lombardy - representing mid-to-large-scale test cases and characterized by different topologies), evaluating two blood collection management strategies, namely:

1. a *centralized strategy* where the entire region is managed as a single unit, as previously shown in [88].

2. a *decentralized strategy* where the region is divided into different areas to prevent polarization caused by heavily populated cities, thereby ensuring a more significant presence of BFs throughout the territory for improved average accessibility.

The experimental results demonstrate the capability of the modeling framework in handling real-world instances, thus confirming its usability as a valuable decision-support tool. Indeed, on the one hand, it allows to derive useful managerial insights, and, on the other hand, it enables the simulation of different system configurations, operational conditions, and working scenarios. Notably, we show that all the instances are solved in acceptable computational times through an off-the-shelf solver, which makes it potentially interesting for practitioners seeking support in policy-making.

The rest of the chapter is organized as follows. Section 2.2 provides a comprehensive literature review on facility location approaches for BSC design problems to highlight the application and methodological contribution of this work. In Section 2.3, we present a formal description of the problem, while the developed mathematical models are given in Section 2.4. In Section 2.5, we report on the selected case studies and present an analysis of the results obtained from our experiments. Finally, Section 2.6 closes the chapter with some conclusions and potential avenues for future research.

2.2 Literature background and positioning of the current contribution

Operations Research is widely recognized as an essential tool for decision support in healthcare planning and management. The use of optimization techniques in healthcare has received significant attention over the past three decades, as evidenced by [231] and [8]. In this research domain, Blood Supply Chain management is an area of growing interest, dealing with a broad spectrum of strategic, tactical, and operational problems. For a comprehensive overview on the topic, the interested reader is addressed to [37], [209], [220], and [184]. In particular, the latter reference provides an insightful conceptualization of the different typologies of problems relevant to BSC management - classified by planning level - and the main decisions involved.

At the *tactical* and *operational* planning level, typical problems concern the definition of short-to-mid-term optimal collection, production, transportation, and administration policies of blood units and derivatives (and their daily scheduling) occurring at collection/production facilities or demand nodes (i.e., hospitals - see, e.g., [129, 165, 210], to name a few).

At the *strategic* level, instead, the focus is on the long-term (re-)design of BSC networks, which usually calls for the optimal siting of BFs and, eventually, the adequate definition of their sizes/capacities. Therefore, BSC network design problems are often rooted in the *Facility Location* literature (FLPs, see [160]).

As discussed in [184], BSC management problems concerning different planning levels have been rarely treated in a fully-integrated manner. Indeed, the complexity of the interconnections among the system components and operations may easily lead to intractability. However, it is widely recognized the significant impact of appropriate design strategic decisions on the tactical and operational performance of the whole BSC ([53]).

In this context, this work mainly focuses on the long-term planning problem related to the design and structure of the BSC under *ordinary conditions*, i.e., not motivated by emergencies and disaster scenarios (see [232] and the references therein). Thus, it falls within the strategic level problems. Nevertheless, as suggested in [184], in order to consider the inter-dependency among the levels, several tactical decision related to collection, productivity and transshipment policies are integrated within the proposed framework. Notably, various scholars investigated such problems by resorting to multi-echelon facility location/allocation models and solution methods. As discussed in [17], the literature can be classified on the basis of four main criteria: data availability, length of the time horizon, involved processes and structure of the system (in terms of number of echelons and type of facilities and their interactions). Accordingly, in order to better motivate the proposed case-based and scenario-based formulations, in the following, we provide a concise review of some recent contributions in the field, dividing them in two groups - namely deterministic and uncertainty based - highlighting for each of them the main aspects related to the these criteria.

Concerning deterministic approaches, in [66], is considered the problem of determining the optimal location of blood banks and their allocation to hospitals, seeking to minimize the related costs. To this end, a multi-objective, single-period, and deterministic model was developed, which was then tested on small-sized numerical examples. Thus, donors allocation is not explicitly considered

In [253], the real problem of regionalizing blood services in Turkey was addressed. The authors considered a hierarchical structure involving higher-level facilities, the so-called Regional Blood Centers, and three lower-level facilities, i.e., BCs, BSs and MUs. The authors tackled the problem by solving a series of subsequent deterministic, single-period mathematical models. In practice, the envisaged location-allocation decisions are not taken in an integrated fashion. Note also that no capacity restrictions were considered, nor the actual distribution of donors was explicitly taken into account.

A slightly different problem was investigated in [99]. In their work, the authors devised a mathematical model for the optimal location of centrifuge centers, responsible for blood separation, and the allocation of clinics, in charge of blood collection. The authors sought to maximize the total profit gained by a health management organization operating in Israel. As in [253], donors and capacity-related decisions were not involved.

In [67] the authors developed a single-period deterministic model to locate two types of collection facilities and decide their optimal allocation to processing facilities, seeking to optimize transportation costs and the amount of donated/processed blood under a budgetary constraint. The model was tested on the Thai Red Cross case study, using the 76 capitals of Thailand as candidate locations and assuming donated blood at each candidate location as known beforehand (that is, donors and capacities were not explicitly considered).

Finally, in [53], the problem of redesigning an existing network of BCs - responsible for collection and processing activities - was tackled, by means of the following decisions: downgrading a BC to a BS, performing only collection, or closing it. The authors devised a deterministic and single-period model, including donors and BSs-to-BCs allocation decisions, aimed at minimizing location and transportation costs. The method was extensively tested on a relatively large-sized case study corresponding to an entire Italian region. For the same problem, some of the same authors introduced several soft constraints to deal with capacity and demand satisfaction requirements - whose corresponding shortages/surplus were penalized in the objective function - and blood units transportation costs in a multi-objective-like fashion (see [88]).

Concerning instead uncertainty based approaches, in [313] is studied the problem of locating capacitated fixed and temporary BFs over a multi-period planning horizon, intending to satisfy the total demand for blood in each period at the minimum location and transportation cost. Donors' allocation decisions were also included. Specifically, a covering radius was considered for donors to reach a located facility. To model the problem, a robust possibilistic mathematical program accounting for the uncertain nature of some relevant parameters was proposed, which was tested on a relatively small-sized randomly generated instance. In a follow-up paper, the above work was extended by considering blood group compatibility [312].

A similar problem setting, both in terms of network structure and decisions involved, was explored in [233], where a robust optimization model motivated by uncertainty on blood demand at hospitals was developed. In the proposed model, inventory and shortage costs at hospitals were also considered. Experiments were run on a small case study from the city of Teheran (similar to that in [142]), characterized by 22 donor groups - corresponding to the 22 city districts - with the latter being also regarded as potential locations for BFs.

A two-stage stochastic programming model was proposed in [17] where two types of facilities, i.e., Blood Collection Centers and Blood Collection and Processing Centers, had to be located over a discrete multi-period planning horizon. First-stage strategic variables referred to the location of the above-mentioned facilities and appropriate allocation decisions. Second-stage (scenario-dependent) decisions, instead, were related to the shipment of collected blood units based on the first-stage allocation decisions. The East Azerbaijan Province case was used to illustrate the performance of the model and of the developed accelerated Benders' decomposition approach.

In [254] a robust fuzzy mathematical model for the optimal location of capacitated blood collection facilities in a multi-period setting was developed, with the main novelty of the work being in the multi-attribute group decision-making approach used to account for both qualitative and quantitative factors in the network design problem. The authors performed their empirical tests on the small case of the city of Tabriz in Azerbaijan (ten donors' locations and less than ten potential locations for blood facilities).

In [195], a multi-period model for the location of blood collection and blood processing facilities under stochastic hospital demands was proposed. The model optimized two objective functions: (i) the total costs and (ii) the total time (e.g., transportation/processing) in the BSC. Note that donors were not explicitly considered, as the number of blood units collected was regarded as an exogenously given parameter. The so-called "Interval Evidential Reasoning" approach was used to handle uncertainty, and evolutionary genetic algorithms were used as solution approaches. Numerical illustrative tests were performed on randomly generated instances.

In [13] the authors focused on the design of a BSC network comprising donors, blood collection facilities, blood centers, and hospitals. Notably, no location decisions were considered. Instead, optimal allocation and inventory policies were explored to simultaneously optimize environmental, social, and economic objectives under uncertainty on both supply (i.e., donations) and demand. Later on, a similar problem setting has been considered at the core of the developments presented in [193]. In particular, the latter reference mainly emphasized sustainability aspects related to the closed-loop nature of the investigated BSC.

Finally, a single-period model for BSC network design under uncertainty on donations, demand, and blood disposal rates was designed in [285], with the aim of optimally locating three types of facilities: Blood Donation Centers (BDCs), Blood Centers (BCs), and Regional Hospitals (RHs), seeking to minimize the total costs (location, transhipments, wastage). The investigated BSC also involved donors and final hospitals, although the geographical distribution of donors (and their allocations) was not considered. The authors employed an interactive possibilistic programming approach to solve the problem, tested on the case of Teheran (19 BDCs, two BCs, six RHs, and 23 hospitals).

Looking at the above discussed contributions, our work differentiates with previous literature review in two main perspectives, namely the application field and methodological approach.

Concerning the application perspective, it is easy to note that most of the contributions deal with very specific BSC design problem deriving from the national/regional context under investigation. Thus, also related solution approaches are highly application-dependent, and their usage cannot be straightly replicated and/or extended to other situations. The approach proposed in this work is specifically outlined for the Italian BSC. Nevertheless, since it is motivated by the transposition of a European Directive, contrarily to the previous ones, it could actually represent a baseline to be used and particularized in the reorganization of the BSCs of other European countries. Finally, it is also worth noting that in most of the previous literature, probably for the sake of tractability, the investigated case studies are relatively small in terms of cardinality of the involved sets, thus being only partially representative of real-world problems. Instead, our contribution deal with large scale test cases built upon a real dataset, thus providing an accurate system sketch and simulation of the BSC functioning.

Concerning instead the methodological perspective, consistently with the specific applications under investigation, this work differentiates from previous literature for one or more of the following features characterizing a BSC design problem, i.e.: donor allocation decisions; capacity, productivity and self-sufficiency constraints; number of echelons; functioning mechanism; targets to be optimized.

To the best of the authors' knowledge, no contribution in the literature deals with a multi-echelon BSC system envisaging simultaneously capacity, productivity, self-sufficiency and accessibility constraints, within a unique multi-objective framework in deterministic and uncertain settings. Some similarities can be found just in [53] and [88], which represent the first and only works specifically dealing with the problem under investigation. Nevertheless, this work provides the following advancements: it provides a modified and integrated MILP deterministic formulation, strengthened by valid inequalities, and envisaging accessibility issue; it extends the deterministic formulation introducing the scenario-based formulation in order to deal with uncertainty on donation propensity; finally, it presents an extensive experimental campaign aimed at simulating different collection strategies and performing both a sensitivity analysis and an evaluation for expected value and worst-case scenarios.

2.3 *BSC* multi-echelon facility location problem *(BSC-MFLP)*: description and assumptions

Referring to a generic region in Italy, we consider the existence of a discrete set of pre-existing BFs responsible for blood collection and processing activities. In our terminology, following the literature (see [253]), we refer to these facilities as BCs. Additionally, we assume the presence of a discrete set of points representing potential locations of blood donors. We posit that the propensity of donors to donate blood is denoted as the "*donation rate*", which remains constant across all points. The donation rate is expressed as the number of donations per 1,000 inhabitants, estimated using publicly available data. In addition, given that blood donations are voluntary and unpaid, we reasonably assume that distance plays a crucial role in influencing the propensity to donate. Hence, we postulate that donors are more likely to visit a BC if it is within a maximum distance from their geographical location, referred to as the "*reachability threshold*". The objective of the BSC is to ensure "*self-sufficiency*" within the healthcare system, meaning that the annual demand for blood units transfused at hospitals is adequately met. According to national government guidelines (see [64]), BCs should achieve a minimum productivity level of 40,000 blood units per year to maintain accreditation. However, current data indicates that the actual number of blood units processed per BC falls significantly below this threshold (refer to [189]). Consequently, strategies to enhance the efficiency of regional BSCs are being sought. In this regard, rationalizing the BSC by reducing the number of active BCs presents a viable possibility. However, it is crucial to ensure the widespread presence of BFs across the territory, as it strongly influences donation rates and, therefore, self-sufficiency. Thus, we consider the following potential decisions:

- Keep an existing BC active, holding both collection and processing activities;
- Downgrade a BC into a BS, performing collection activities only;
- Closing a BC and not using it anymore.

Clearly, donors have the option to make donations at both BSs and BCs within the reachability threshold. It is reasonable to assume that they would choose the closest facility (referred to as "*closest assignment*"). Furthermore, as discussed in Section 2.1, the data indicates that donations predomitly come from periodic and healthy donors. Therefore, in this chapter, we make the following assumptions:

- donors are defined as the percentage of the population in a given area who make donations, based on the donation rate;
- all donors, within the reachability threshold of a located BS or BC, will perform donations;
- all collected blood units will be processed.

Blood units collected at BSs need to be transferred and processed at an active BC. Since blood is a perishable product that degrades quickly, we assume that such transfers are allowed only between a BS and a BC located within a maximum given distance ("*degradation threshold*"). By adopting this approach, we aim to align the resulting system with regulatory frameworks by consolidating production in a reduced number of BCs. Together with BSs, these facilities ensure a widespread presence of collection points for users. Additionally, to prevent the consolidation of collection activities, we further assume that there is a maximum capacity for blood units at BCs and BSs due to limited personnel availability. Specifically, we assume a maximum capacity for blood collection only at BCs or BSs. This is necessary because medical staff is required to oversee and manage the donation process. In terms of productivity, there is no need to establish a fixed capacity since these activities rely primarily on specialized machinery and equipment, resulting in shorter processing times and uninterrupted operation. Furthermore, we consider the possibility of utilizing MUs to support collection activities. Specifically, we assume the availability of a limited fleet of MUs responsible for collecting blood units from specific donor locations and transporting them to a BC within the degradation threshold. It is important to note the significant implications of using MUs to ensure compliance with policy requirements within the BSC. Firstly, it enables reaching more isolated donor locations for self-sufficiency. Secondly, it helps meet the maximum collection capacity constraint at located BCs and BSs by reducing the number of blood units collected there. Lastly, as the allocation of MUs to BCs is a network design choice, it can enhance minimum productivity levels at BCs. These aspects are further clarified through a toy-example depicted in Figure 2.1 and presented in Section 2.4 to provide a clearer exposition of the introduced mathematical notation. Considering that the envisaged reorganization process involves a reduction in available BFs and recognizing the public utility of the BSC, we additionally assume that the decision-maker is interested in avoiding a significant deterioration in donors' accessibility to the service. To achieve this, we ensure that the average distance between donors and a located BF remains below a maximum threshold value (referred to as "*accessibility threshold*"), thus maintaining the presence of BFs spread across the territory. It is worth noting that overall accessibility considerations encompass all potential donors in the region. Although only a subset of donors is considered for self-sufficiency in our problem setting, even those located farther from a BS or BC than the reachability or degradation thresholds may theoretically donate blood.

In a nutshell, our problem involves a multi-echelon BSC network consisting of donors, MUs, BSs, and BCs. Location and allocation decisions need to be made to fulfill three primary requirements: (i) achieving self-sufficiency, (ii) ensuring maximum capacity collection at BSs, and (iii) maintaining minimum productivity levels at BCs. We assume that the decision-maker is interested in evaluating solutions that strike a balance between minimizing transportation costs for blood units (from BSs and MUs to BCs) and complying with the aforementioned requirements. This aligns with the ongoing process of reorganizing regional BSCs, where fully adhering to these requirements may not be feasible in practical terms due to existing vulnerabilities. To address this, these conditions are formulated as soft constraints, and any violations is accounted for in the objective function using a penalty-based approach. In the following, for the sake of readability, the problem is denoted as BSC-MFLP, i.e., Blood Supply Chain Multi-echelon Facility Location Problem.

2.4 Mathematical models for the *BSC-MFLP*

The objective of this chapter is to devise a mathematical modeling framework to tackle the BSC-MFLP at regional level in Italy. We recall that our modeling framework envisages two perspectives: case-based and scenario-based. In the following, these are referred to as *BSC-MFLP-CB* and *BSC-MFLP-SB*, respectively. Hereafter, we first describe the notation used to model the *BSC-MFLP-CB* in Section 2.4.1. The related mathematical formulation is given in Section 2.4.2. Then, Section 2.4.3 focuses on the scenario-based case, i.e., the *BSC-MFLP-SB* formulation. Finally, in Section 2.4.4, we present some model enhancements based on strengthened constraints re-formulations and valid inequalities.

2.4.1 Notation

To formulate the *BSC-MFLP-CB*, we introduce two primary sets: the set of potential BF locations, denoted by J, and the set of donor locations/points, represented by I, whose generic elements are denoted as i and j, respectively. The population of each location i, $i \in I$ is represented by p_i, and, being α the donation rate, the quantity of collectable blood units at point i is calculated as $a_i = \alpha \cdot p_i$ (expressed in thousands units). The donors-to-facility distances for each pair $(i, j), i \in I, j \in J$ are denoted by d_{ij}, while we use c_{jk} for the mutual distances between facilities $(j, k) \in J$. Besides, let r and c_{max} be the reachability and degradation thresholds, respectively.

On this basis, the following sets are defined:

- N_i: set of BF locations j, $j \in J$, such that their distance from the donation point i, $i \in I$, is at most equal to reachability threshold ($d_{ij} \leq r$);
- S_j: set of donation points i, $i \in I$, such that their distance from the BF location j, $j \in J$, is at most equal to to reachability threshold ($d_{ji} \leq r$);

- M_j: set of facilities k, $k \in J$, such that their distance from the BF location j, $j \in J$, is at most equal to the degradation threshold ($c_{jk} \leq c_{max}$). Moreover, we define also $M_j^+ = M_j \cup \{j\}$;
- O_i: set of BF locations j, $j \in J$, such that their distance from the donation point i, $i \in I$, is at most equal to the degradation threshold ($c_{ij} \leq c_{max}$);
- T_j: set of donation points i, $i \in I$, such that their distance from the BF location j, $j \in J$, is at most equal to the degradation threshold ($d_{ji} \leq c_{max}$).

Moreover, the following additional non-negative parameters are introduced: annual regional blood demand (D); blood collection minimum productivity (P_{min}); BF maximum collection capacity (C); number of available Mobile Units (n); penalty factors for productivity shortage (λ_1), capacity overrunning (λ_2), and blood self-sufficiency shortage (λ_3).

Given this setting the following binary decision variables are introduced:

- *Location variables*
  - $y_j^s \in \{0,1\}$, $j \in J$: each variable assumes value 1 if location j is selected for a BS, 0 otherwise.
  - $y_j^c \in \{0,1\}$, $j \in J$: each variable assumes value 1 if location j is selected for a BC, 0 otherwise.
- *Assignment variables*
  - $x_{ij} \in \{0,1\}$, $i \in I$ and $j \in N_i$: each variable assumes value 1 if donation point i is assigned to the BF in j, 0 otherwise.
  - $z_{ij} \in \{0,1\}$, $i \in I$ and $j \in O_i$: each variable assumes value 1 if donation point i is assigned to an MU which collects the blood and transfers it to a BC in j, 0 otherwise.
- *Blood transfer variables*
  - $q_{jj'} \in \{0,1\}$, $j \in J$ and $j' \in M_j^+$: each variable assumes value 1 if the blood collected at a BF in j is transferred to a BC in j', 0 otherwise.
  - $t_{ijj'} \in \{0,1\}$, $i \in I$, $j \in N_i$ and $j' \in O_i$: each variable assumes value 1 when donation point i, who could be assigned to an open BF in j, is served by an MU that collects the blood and transfers it to a BC in j', 0 otherwise.

On this basis, we can infer that, given a BSC solution, a donation point $i \in I$, falls in one of the following situations: (i) it is assigned to a BF in j (i.e.

$x_{ij} = 1$), thus donors of point i have to travel a distance $d_{ij} \leq r$ to reach the BF in j; (ii) it is assigned to a MU transferring the collected blood to a BF in j or in j' (i.e., either $z_{ij} = 1$ or $t_{ijj'} = 1$), thus collection occurs directly in i; (iii) it remains unassigned, thus donors in i should travel a distance $d_{ij} \geq r$ to reach a BF (and possibly donate blood). Hence, only donor locations falling in cases (i) and (iii) concur to the overall BSC accessibility. In view of this, we define additional auxiliary variables, denoted as $\overline{x}_{ij} = \{0, 1\}$, $i \in I$ and $j, j \in J$, to compute the actual accessibility of the system. Each of these variables assumes value 1 if the donation point i remains unassigned, but it is associated to a BF in j in the evaluation of the average accessibility.

Finally, the following four continuous non-negative decision variables are used:

- *Accessibility*
  - $l_i \geq 0$, $i \in I$: it computes the contribution of donors of location i to the average accessibility;

- *Self-sufficiency*
  - $\delta \geq 0$: it measures the scarcity of blood collection related to a given demand D;

- *Productivity and Capacity*
  - $\Phi_j \geq 0$ and $\Psi_j \geq 0$, $j \in J$: they represent the processed blood shortage and collection capacity overrun, respectively, at a BF location j.

To provide a clearer explanation of the introduced notation, we present an illustration of the BSC system under investigation and a feasible solution in Figure 2.1, highlighting all the possible instance occurrences. Each donation point i, $i \in I$ is represented by a black dot, labeled with a letter ($i \in \{a, \ldots, z\}$). Additionally, let us assume that each donation point is associated with the quantity of blood units given by the number in round brackets. Each potential facility location j, $j \in J$ is represented by gray geometrical figures, numbered from 1 to 5. Triangles represent facilities downgraded to BS (y_j^s variables equal to 1), squares represent facilities kept open as BC (y_j^c variables equal to 1), and circles represent closed facilities (y_j^c and y_j^s variables equal to 0). Two different covering areas - black and grey dashed line circles - are associated with each BF. These areas are built on the basis of the reachability threshold r and degradation threshold $c_{\max}$, respectively. Let us also consider the following parameter values: $D = 30$; $C = 10$ and $P_{\min} = 5$ for all the

BFs. Finally, let us assume that all the penalties λ_1, λ_2 and λ_3 are set to very high values, but $\lambda_3 >> \lambda_1$ and $\lambda_3 >> \lambda_2$. This setting pushes towards solutions guaranteeing primarily the satisfaction of the regional self-sufficiency, secondarily the productivity and capacity requirements and, finally, the minimization of the system blood transportation costs. We observe that 4 BFs have to be opened to satisfy the regional blood demand D, so avoiding incurring in the resulting penalty. The red and green lines connect donation points with open BFs and they illustrate the x_{ij} and the z_{ij} assignment variables, respectively. Indeed, red straight lines are associated with donation points whose distanced from an open BF is within the reachability threshold r, while green straight lines are associated with donation points served by an MU, whose distanced from an open BF is larger than reachability threshold r but within the degradation threshold c_{max}. Blue bold lines, connecting an open BS to an open BC whose mutual distances is within the degradation threshold c_{max}, illustrate the $q_{jj'}$ transfer variables. Then, purple bold lines stand for the $t_{ijj'}$ transfer variables, and connect a donation point (o in the Figure 2.1), which should be inherently assigned to the nearest BF (3 in Figure 2.1), to another open BF (4 in Figure 2.1) by using an MU. Finally, brown dotted lines are related to $\overline{x}_{ij}$ variables, thus they connect an unassigned donation point to an open BF (likely the nearest one), to concur in the computation of the average accessibility. On the basis of the introduced picture legend, we can now easily explain the structure of the solution represented in Figure 2.1. For the sake of comprehension, let us recall our assumption imposing that once a BF is open, every donation point within its reachability threshold must be assigned to the BF itself, or, in alternative, the blood of the donation point has to be collected by an MU and then transferred to another open BF within the degradation threshold. The BC located in 3 collects blood units from donation points h and i, so satisfying the imposed capacity limit and achieving the minimum productivity requirement. It also receives the blood collected from BS in 1, so increasing the amount of processed units. However, donation point j, which should be inherently assigned to BC 3, is instead assigned to BC 4 through the usage of an MU. The reason for this is that such blood transfer has a twofold positive effect. On the one hand, it avoids the violation of capacity constraints for BC 3 and, on the other hand, it allows the achievement of the productivity requirement for BC 4. Thus, no penalty is paid. Finally, BC in 5 meets the productivity requirement but it collects more blood than its capacity, thus incurring in the resulting penalty. The reason for this is that all the donation points within the reachability threshold of BC 5 can be assigned only to BC 5. To conclude, Figure 2.1

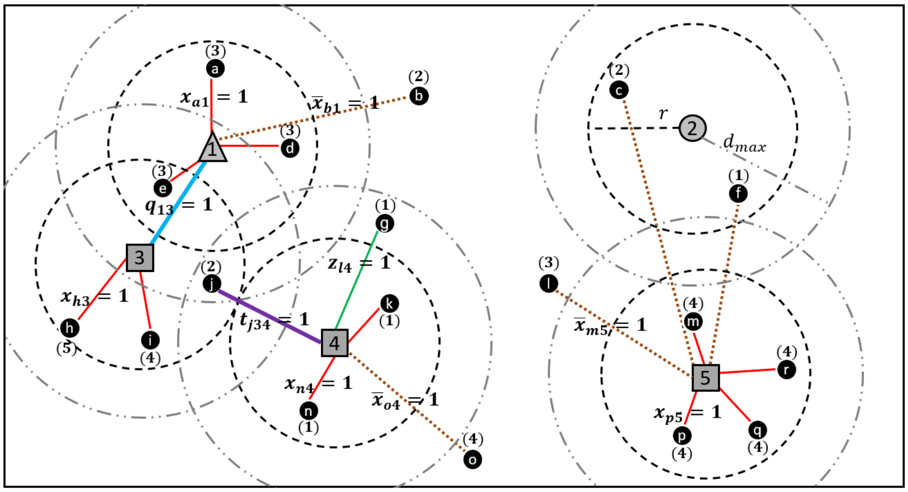


Figure 2.1: Illustration of a *BSC-MFLP-CB* solution, related setting and parameters

allows also to show the evaluation of the average accessibility of the system. Indeed, it is computed summing the distance values d_{ij} associated with all the x_{ij} and $\overline{x}_{ij}$ variables equal to 1 (red and brown straight lines), then divided by $|I|$. A feasible solution has an accessibility lower than or equal to the threshold β.

2.4.2 Case-based formulation (*BSC-MFLP-CB*)

Using the notation introduced in Section 2.4.1, the *BSC-MFLP-CB* can be formulated as follows:

$$\begin{aligned} Min \sum_{i\in I}\sum_{j\in N_i}\sum_{j'\in M_j} a_i d_{jj'} x_{ij} q_{jj'} + \sum_{i\in I}\sum_{j\in O_i} a_i d_{ij} z_{ij} + \\ + \sum_{i\in I}\sum_{j\in N_i}\sum_{j'\in O_i} a_i d_{ij'} t_{ijj'} + \sum_{j\in J}(\lambda_1 \Phi_j + \lambda_2 \Psi_j) + \lambda_3 \delta \end{aligned} \tag{2.1}$$

$s.t.$

$$y_j^s + y_j^c \leq 1 \qquad \forall j \in J \tag{2.2}$$

$$x_{ij} + \sum_{j'\in O_i} t_{ijj'} \leq y_j^s + y_j^c \qquad \forall i \in I, j \in N_i \tag{2.3}$$

$$\sum_{k\in N_i}(x_{ik}+(\sum_{l\in O_i}t_{ikl}))\geq y_j^s+y_j^c \qquad \forall j\in J, i\in S_j \quad (2.4)$$

$$\sum_{j'\in N_i}d_{ij'}(x_{ij'}+\sum_{k\in O_i}t_{ij'k})+(F-d_{ij})(y_j^s+y_j^c)\leq F \qquad \forall i\in I, j\in N_i \quad (2.5)$$

$$q_{jj'}\leq y_{j'}^c \qquad \forall j\in J, j^{'}\in M_j^+ \quad (2.6)$$

$$\sum_{j'\in M_j^+}q_{jj'}=y_j^s+y_j^c \qquad \forall j\in J \quad (2.7)$$

$$q_{jj}\geq y_j^c \qquad \forall j\in J \quad (2.8)$$

$$\sum_{j\in N_i}(x_{ij}+\sum_{j'\in O_i}t_{ijj'})+\sum_{j\in O_i}z_{ij}\leq 1 \qquad \forall i\in I \quad (2.9)$$

$$z_{ij}\leq y_j^c \qquad \forall i\in I, j\in O_i \quad (2.10)$$

$$t_{ijj'}\leq y_{j'}^c \qquad \forall i\in I, j\in N_i, j^{'}\in O_i \quad (2.11)$$

$$\sum_{i\in I}a_i[\sum_{j\in N_i}(x_{ij}+\sum_{j'\in O_i}t_{ijj'})+\sum_{j\in O_i}z_{ij}]+\delta\geq D \quad (2.12)$$

$$\sum_{i\in S_j}a_ix_{ij}-\Psi_j\leq C \qquad \forall j\in J \quad (2.13)$$

$$\sum_{k\in M_j^+}\sum_{i\in S_k}a_ix_{ik}q_{kj}+\sum_{i\in T_j}a_i(z_{ij}+\sum_{k\in N_i}t_{ikj})+\Phi_j\geq P_{\min}\,y_j^c \qquad \forall j\in J \quad (2.14)$$

$$\sum_{i\in I}\sum_{j\in O_i}(z_{ij}+\sum_{j'\in N_i}t_{ij'j})\leq n \quad (2.15)$$

$$\overline{x}_{ij}\leq y_j^s+y_j^c \qquad \forall i\in I, j\in J \quad (2.16)$$

$$\sum_{j\in J}\overline{x}_{ij}+\sum_{j\in N_i}(x_{ij}+\sum_{j'\in O_i}t_{ijj'})+\sum_{j\in O_i}z_{ij}=1 \qquad \forall i\in I \quad (2.17)$$

$$l_i\geq\sum_{j\in J}d_{ij}\overline{x}_{ij}+\sum_{j\in N_i}d_{ij}x_{ij} \qquad \forall i\in I \quad (2.18)$$

$$\frac{\sum_{i\in I}l_i}{|I|}\leq\beta \quad (2.19)$$

$$y_j^c\in\{0,1\}, y_j^s\in\{0,1\} \qquad \forall j\in J \quad (2.20)$$

$$x_{ij}\in\{0,1\} \qquad \forall i\in I, j\in N_i \quad (2.21)$$

$$z_{ij}\in\{0,1\} \qquad \forall i\in I, j\in O_i \quad (2.22)$$

$$q_{jj'} \in \{0,1\} \qquad \forall j \in J, j' \in M_j^+ \quad (2.23)$$
$$t_{ijj'} \in \{0,1\} \qquad \forall i \in I, j \in N_i, j' \in O_i \quad (2.24)$$
$$l_i \geq 0 \qquad \forall i \in I \quad (2.25)$$
$$\Phi_j \geq 0, \Psi_j \geq 0 \qquad \forall j \in J \quad (2.26)$$
$$\delta \geq 0 \qquad (2.27)$$

The objective function (2.1) represents the overall system costs. In particular, the first three terms represent the weighted blood transportation cost. The last two terms represents the penalty arising from the violation of the soft constraints. Constraints (2.2) ensure that a BF can be used either as a BC or as a BS. Constraints (2.3) are consistency constraints between assignment and location variables. They impose that donors can be assigned to a BF only if there is an open BS or BC. Constraints (2.4) ensure that if there is an open BF, the blood of all the donors located within the reachability threshold is collected. Constraints (2.5) guarantee that a donor is assigned to the nearest open BF within the reachability threshold. This constraint requires the definition of a *big M* value, denoted as F, $F = max_{i\in I, j\in J} d_{ij}$. Constraints (2.6) ensure that blood collected in a BF is processed in an open BC. Constraints (2.7) impose that blood collected in a BF has to be transferred to an open BC within the degradation threshold. Constraints (2.8) guarantee that blood collected in a BC is processed by itself. Constraints (2.9) ensure that blood of each donation point can be collected at most once. Constraints (2.10) impose that MUs can transfer blood only to open BC. Constraints (2.11) guarantee that blood re-allocations can occur only in correspondence with an active BC. The soft constraint (2.12) ensures that the total amount of blood collected is at least equal to the regional blood demand D (self-sufficiency goal). Soft constraints (2.13) define an upper limit to the capacity for the collection activities at each facility. Soft constraints (2.14) require that blood processed in each BC satisfies the minimum productivity target P_{min}. Constraints (2.15) ensure that the total number of the MUs does not exceed the fleet size. Constraints (2.16-2.18) guarantee the consistency between donor assignment variables (x, t, and z) and the auxiliary variables ($\overline{x}$ and l) used to compute the average donor distance. Constraint (2.19) ensures that the average donor distance to an open BC/BS is smaller than the threshold β. Finally, constraints (2.20) - (2.27) concern the nature of the variables.

The model just described is not linear due to the product between the two decision variables x_{ij} and $q_{jj'}$ in the objective function (2.1) and constraints (2.14). However, as widely known, it is possible to linearize both terms by introducing a new binary decision variable $w_{ijj'}$ ($w_{ijj'} = x_{ij}q_{jj'}$) and the following set of constraints:

$$w_{ijj'} \leq x_{ij} \qquad \forall i \in I, j \in N_i, j^{'} \in M_j^{+} \quad (2.28)$$

$$w_{ijj'} \leq q_{jj'} \qquad \forall i \in I, j \in N_i, j^{'} \in M_j^{+} \quad (2.29)$$

$$w_{ijj'} \geq x_{ij} + q_{jj'} - 1 \qquad \forall i \in I, j \in N_i, j^{'} \in M_j^{+} \quad (2.30)$$

Constraints (2.28), (2.29) and (2.30) ensure that a binary variable $w_{ijj'}$ is equal to 1 if and only if donors located in $i, i \in I$, are assigned to the BS located in $j, j \in N_i$, and the collected blood is transferred and then processed in $j^{'} \in M_j^{+}$. The objective function (2.1) and the constraints (2.14) are therefore rewritten in the following form:

$$\begin{aligned} \textit{Min} \sum_{i\in I}\sum_{j\in N_i}\sum_{j'\in M_j} a_i d_{jj'} w_{ijj'} + \sum_{i\in I}\sum_{j\in O_i} a_i d_{ij} z_{ij} + \\ + \sum_{i\in I}\sum_{j\in N_i}\sum_{j'\in O_i} a_i d_{ij'} t_{ijj'} + \sum_{j\in J}(\lambda_1 \Phi_j + \lambda_2 \Psi_j) + \lambda_3 \delta \end{aligned} \quad (2.31)$$

$s.t.$
$(2.2) - (2.13), (2.15) - (2.27)$

$$\sum_{k\in M_j^+}\sum_{i\in S_k} a_i w_{ikj} + \sum_{i\in T_j} a_i (z_{ij} + \sum_{k\in N_i} t_{ikj}) + \Phi_j \geq P_{min}\, y_j^c \qquad \forall j \in J \quad (2.32)$$

2.4.3 Scenario-based formulation (*BSC-MFLP-SB*)

The donation rate α is a crucial parameter for the problem under investigation and it has shown a quite stable trend over the last years ([64]). However, as discussed in Section 2.1, even small fluctuations can have significant impacts on the BSC operations. Thus, it represents an inherent source of uncertainty which deserves to be investigated by devising ad-hoc uncertainty based models which allow to design more robust BSC solutions.

Such situation configures a BSC-MFLP under uncertainty. We assume that uncertainty can be captured by a finite set of scenarios, say Π, where a "scenario" represents a full realization of the uncertain parameter α. Under this simplifying but widely known and used hypothesis (see, e.g., [74, 121]), the BSC-MFLP can be modelled by a deterministic scenario-based equivalent formulation (BSC-MFLP-SB), by indexing in Π the donation rate parameter ($\alpha_\pi, \pi \in \Pi$), and several decision-variables.

Specifically, we assume that only the strategic location-variables (i.e., y_j^c and y_j^s, $j \in J$) are scenario-independent decisions. Hence, a solution is sought which adapts to occurring scenarios by appropriately defining the donors-to-facilities allocations ($x_{ij\pi}$), the BSs-to-BCs assignments ($q'_{jj\pi}$), as well as the dispatching and use of MUs ($z_{ij\pi}$, $t_{ijj'\pi}$). Clearly, also the decision variables accounting for the violation of the main requirements (e.g., self-sufficiency, minimum productivity level, maximum collection capacity) are indexed in Π (δ_π, ψ_π, ϕ_π). The same applies to the auxiliary variables used for computing accessibility, which depend on scenario-dependent allocation decisions ($\overline{x}_{ij\pi}$, $l_{i\pi}$).

Finally, we also assume that the occurrence of scenarios $\pi \in \Pi$ can be measured probabilistically and we denote by ω_π such probabilities.

Two possible cases are considered, corresponding to two different decision-making attitudes towards risk: risk-neutral and risk-averse. Mathematically, we formulate them in terms of the minimization of the expected and maximum BSC's costs across the occurring scenarios, respectively. The goal is to provide additional evidence on the impact and the type of solutions that an uncertainty-aware decision-maker can evaluate as an alternative to classical sensitivity analysis. Using this additional notation (and that introduced in Section 2.4.1), the *BSC-MFLP-SB* model, in its linearized form, can be formulated as follows:

$$\begin{aligned} \mathit{Min} \sum_{\pi\in\Pi} \omega_\pi [& \sum_{i\in I} \sum_{j\in N_i} \sum_{j'\in M_j} a_{i\pi} d_{jj'} w_{ijj'\pi} + \sum_{i\in I} \sum_{j\in O_i} a_{i\pi} d_{ij} z_{ij\pi} + \\ & + \sum_{i\in I} \sum_{j\in N_i} \sum_{j'\in O_i} a_{i\pi} d_{ij'} t_{ijj'\pi} + \sum_{j\in J} (\lambda_1 \Phi_{j\pi} + \lambda_2 \Psi_{j\pi}) + \lambda_3 \delta_\pi] \end{aligned} \tag{2.33}$$

$s.t.$

$$y_j^s + y_j^c \leq 1 \qquad \forall j \in J \tag{2.34}$$

$$x_{ij\pi} + \sum_{j'\in O_i} t_{ijj'\pi} \leq y_j^s + y_j^c \qquad \forall i \in I, j \in N_i, \pi \in \Pi \tag{2.35}$$

$$\sum_{k\in N_i} x_{ik\pi} + (\sum_{l\in O_i} t_{ikl\pi}) \geq y_j^s + y_j^c \qquad \forall j \in J, i \in S_j, \pi \in \Pi \tag{2.36}$$

$$\sum_{j'\in N_i} d_{ij'} (x_{ij'\pi} + \sum_{k\in O_i} t_{ij'k\pi}) + (F - d_{ij})(y_j^s + y_j^c) \leq F \qquad \forall i \in I, \; j \in N_i, \pi \in \Pi \tag{2.37}$$

$$q_{jj'\pi} \leq y^c_{j'} \qquad \forall j \in J, j' \in M^+_j, \pi \in \Pi \quad (2.38)$$

$$\sum_{j' \in M^+_j} q_{jj'\pi} = y^s_j + y^c_j \qquad \forall j \in J, \pi \in \Pi \quad (2.39)$$

$$q_{jj\pi} \geq y^c_j \qquad \forall j \in J, \pi \in \Pi \quad (2.40)$$

$$\sum_{j \in N_i} (x_{ij\pi} + \sum_{j' \in O_i} t_{ijj'\pi}) + \sum_{j \in O_i} z_{ij\pi} \leq 1 \qquad \forall i \in I, \pi \in \Pi \quad (2.41)$$

$$z_{ij\pi} \leq y^c_j \qquad \forall i \in I, j \in O_i, \pi \in \Pi \quad (2.42)$$

$$t_{ijj'\pi} \leq y^c_{j'} \qquad \forall i \in I, j \in N_i, j' \in O_i, \pi \in \Pi \quad (2.43)$$

$$\sum_{i \in I} a_{i\pi} [\sum_{j \in N_i} (x_{ij\pi} + \sum_{j' \in O_i} t_{ijj'\pi}) + \sum_{j \in O_i} z_{ij\pi}] + \delta_\pi \geq D \qquad \forall \pi \in \Pi \quad (2.44)$$

$$\sum_{i \in S_j} a_{i\pi} x_{ij\pi} - \Psi_{j\pi} \leq C \qquad \forall j \in J, \pi \in \Pi \quad (2.45)$$

$$\sum_{k \in M^+_j} \sum_{i \in S_k} a_{i\pi} w_{ikj\pi} + \sum_{i \in T_j} a_{i\pi} (z_{ij\pi} + \sum_{k \in N_i} t_{ikj\pi}) + +\Phi_{j\pi} \geq P_{min}\, y^c_j \qquad \forall j \in J, \pi \in \Pi \quad (2.46)$$

$$\sum_{i \in I} \sum_{j \in O_i} (z_{ij\pi} + \sum_{j' \in N_i} t_{ij'j\pi}) \leq n \qquad \forall \pi \in \Pi \quad (2.47)$$

$$\overline{x}_{ij\pi} \leq y^s_j + y^c_j \qquad \forall i \in I, j \in J, \pi \in \Pi \quad (2.48)$$

$$\sum_{j \in J} \overline{x}_{ij\pi} + \sum_{j \in N_i} (x_{ij\pi} + \sum_{j' \in O_i} t_{ijj'\pi}) + \sum_{j \in O_i} z_{ij\pi} = 1 \qquad \forall i \in I, , \pi \in \Pi \quad (2.49)$$

$$l_{i\pi} \geq \sum_{j \in J} d_{ij} \overline{x}_{ij\pi} + \sum_{j \in N_i} d_{ij} x_{ij\pi} \qquad \forall i \in I, \pi \in \Pi \quad (2.50)$$

$$\frac{\sum_{i \in I} l_{i\pi}}{|I|} \leq \beta \qquad \forall \pi \in \Pi \quad (2.51)$$

$$y^c_j \in \{0,1\}, y^s_j \in \{0,1\} \qquad \forall j \in J \quad (2.52)$$

$$x_{ij\pi} \in \{0,1\} \qquad \forall i \in I, j \in N_i, \pi \in \Pi \quad (2.53)$$

$$z_{ij\pi} \in \{0,1\} \qquad \forall i \in I, j \in O_i, \pi \in \Pi \quad (2.54)$$

$$q_{jj'\pi} \in \{0,1\} \qquad \forall j \in J, j' \in M^+_j, \pi \in \Π \quad (2.55)$$

$$t_{ijj'\pi} \in \{0,1\} \qquad \forall i \in I, j \in N_i, j' \in O_i, \pi \in \Pi \quad (2.56)$$
$$l_{i\pi} \geq 0 \qquad \forall i \in I, \pi \in \Pi \quad (2.57)$$
$$\Phi_{j\pi} \geq 0, \Psi_{j\pi} \geq 0 \qquad \forall j \in J, \pi \in \Pi \quad (2.58)$$
$$\delta_\pi \geq 0 \qquad \forall \pi \in \Pi \quad (2.59)$$

For the sake of brevity, we do not provide a discussion of the model, since the explanation of the constraints follows that given for the *BSC-MFLP-CB* formulation. We just highlight that the objective function (2.33) minimizes the expected value of the system costs over all the scenarios weighted by the corresponding occurrence probability ω_π.
The proposed *BSC-MFLP-SB* formulation can be slightly modified to minimize the worst-case scenario by varying the objective function, as follows:

$$\begin{aligned} \textit{Min Max}_{\pi\in\Pi} \, [\sum_{i\in I}\sum_{j\in N_i}\sum_{j'\in M_j} a_{i\pi}d_{jj'}w_{ijj'\pi} + \sum_{i\in I}\sum_{j\in O_i} a_{i\pi}d_{ij}z_{ij\pi} + \\ + \sum_{i\in I}\sum_{j\in N_i}\sum_{j'\in O_i} a_{i\pi}d_{ij'}t_{ijj'\pi} + \sum_{j\in J}(\lambda_1\Phi_{j\pi} + \lambda_2\Psi_{j\pi}) + \lambda_3\delta_\pi] \end{aligned} \quad (2.60)$$

$s.t.$
$(2.34) - (2.59)$

Objective function (2.60) configures a $minmax$ problem that, as widely known, can be easily linearized replacing the objective function by an additional non-negative continuous decision variable and introducing bounding constraints. In the following, for the sake of readability, we denote the expected value and worst-case versions of the *BSC-MFLP-SB* as *BSC-MFLP-SB*$_{sum}$ and *BSC-MFLP-SB*$_{max}$, respectively.

2.4.4 Formulation enhancements

This section is devoted to present enhancements of the proposed MILP formulations in order to strengthen the linear relaxation and improve the overall computational efficiency. Upon examining constraints (4.7) and (2.37) for the *BSC-MFLP-CB* and *BSC-MFLP-SB* formulations, respectively, we observe that the usage of F configure a set of $bigM$ constraints in each model. As widely known, this can result in numerical instability and weaker formulations. Such drawback can be partially overcome by a constraint reformulation. For the sake of comprehension, since the extension to the *BSC-MFLP-SB* formulation is straightforward, our discussion is focused only on the *BSC-MFLP-CB* formulation. Constraints (4.7) are conceived as closest assignment constraints where the introduction of a $bigM$ is needed to make the constraint ineffective in case no BF is open either within the reachability threshold or within the degradation distance of a donation point. We can reformulate these constraints as follows:

$$\sum_{j \in N_i | d_{ij} > d_{ik}} x_{ij} + \sum_{h \in O_i | d_{il} > d_{ik}} t_{ilh} \leq 1 - (y_k^s + y_k^c) \qquad \begin{array}{r} \forall i \in I, k \in N_i, \\ l \in N_i \end{array} \quad (2.61)$$

Given a donation point i, $i \in I$, the constraints (2.61) still allow to obtain a closest assignment by introducing a distance-based ordering of the BFs belonging to N_i and the ones belonging to O_i. Specifically, they set to zero all the assignment variables related to BFs that are not the closest to a donation point. Such reformulation leads to a strengthened linear relaxation of the *BSC-MFLP-CB* model and are used in the following computational analysis. Finally, the following set of valid inequalities can be derived from the problem structure:

$$y_j^s \leq \sum_{i \in M_j} y_i^c \qquad \forall j \in J \quad (2.62)$$

These inequalities impose that, given a BF location j, $j \in J$, we cannot have a BS in j if no BC is open within the degradation threshold, i.e., the set M_j is empty. The proof of these inequalities is straightforward. Indeed, let us consider a location $\hat{j}$ whose set $M_{\hat{j}}$ is empty. We can easily verify that a fractional solution with both values $\hat{y_j^s}$ and $\hat{y_j^c}$ higher than 0, satisfy all the constraints linking these variables each other and with the transfer variables $q_{jj'}$, namely constraints (2.6), (2.7) and (2.8), but it does not satisfy the constraints (2.62). The constraints (2.62), involving only location variables, can be directly added also to the *BSC-MFLP-SB* formulation. Instead, the constraints (2.61) can be tailored for the *BSC-MFLP-SB* model by extending them to all the scenarios in π, $\pi \in \Pi$, as follows:

$$\sum_{j \in N_i | d_{ij} > d_{ik}} x_{ij\pi} + \sum_{h \in O_i | d_{il} > d_{ik}} t_{ilh\pi} \leq 1 - (y_k^s + y_k^c) \qquad \begin{array}{r} \forall i \in I, k \in N_i, \\ l \in N_i, \pi \in \Pi \end{array} \quad (2.63)$$

2.5 Computational results

In this section, we test and validate the proposed MILP formulations on real cases related to three Italian regions, namely Campania, Lombardy and Apulia. We selected these regions because they present different geographical/topological and demographic features. Specifically:

1. Campania has a very compact shape; Lombardy is one of the most the widely extended region in Italy; Apulia develops mostly longitudinally;
2. Lombardy and Campania are the first two Italian regions for population density, mainly concentrated in densely inhabited urban areas; the population of Apulia, instead, even if high, is more evenly distributed between urban and rural areas.

Section 2.5.1 describes the instance generation and parameter setting. Sections 2.5.2, 2.5.3 and 2.5.4 are devoted to the presentation and discussion of the experimental results obtained for each region.

2.5.1 Instance generation and parameter setting

The *BSC-MFLP* requires the definition of the donor locations (I) and potential blood facility (J) sets. Concerning the first set, we exploit the pre-existing municipalities constituting the regional territory, which have known populations (p_i), obtained using the latest consolidated census data according to the Italian National Statistics Institute—ISTAT. The number of donor locations (i.e., the cardinality of I) is 559, 257, and 1,513 for the Campania, Apulia, and Lombardy regions, respectively. Their positions are assumed to correspond to the centroids of the municipalities and are extracted using the open-source GIS software QGIS. Without loss of generality, we assume a uniformly distributed donation rate (α) across each municipality, with all donors located at their respective centroids. For regions with larger populations, such as Campania and Lombardy, the municipalities of Naples and Milan are each subdivided into ten sub-areas to effectively manage the high number of inhabitants. The donation rate is assumed to be equal across all the municipalities and its value is determined based on historical data provided by the regional authority. On average, it has been observed that nearly 50 donations per 1,000 inhabitants have been made in Italy over the past five years [64], thus setting $\alpha = 0.05$. Moreover, in order to simulate the effects of potential regional campaigns of awareness aimed at increasing the propensity of the population to blood donation, or the decrease of this rate aimed at simulating a stressful situation for regional blood management system, we also considered the cases with $\alpha = 0.06$ and $\alpha = 0.04$, respectively. We assume that each donor performs one donation per year, thus, the number of donated units for each areas is given by αp_i. The *BSC-MFLP-CB* formulation has been evaluated considering singularly the three chosen donation rate values. Instead, the *BSC-MFLP-SB* formulation has been evaluated considering simultaneously the three donation rate values, i.e., $\alpha_1 = 0.04$, $\alpha_2 = 0.05$ and $\alpha_3 = 0.06$, imposing $\omega_1 = 0.25$, $\omega_2 = 0.5$ and $\omega_3 = 0.25$, respectively.

The values of the annual regional blood demands (D), have been set equal to 161,360 for Campania, 163,881 for Apulia and 470,770 for Lombardy [64].

Concerning the potential BF set, we consider the number and the position of the existing blood centers in each region, i.e., 22, 21 and 33 BCs in Campania, Apulia and Lombardy, respectively. The P_{min} value is fixed to 40.000 units for all the BFs, according to the imposed EU and Italian regulation. We highlight that all the three regions have an average level of productivity significantly lower than the minimum productivity, in particular: 7,335 units for Campania (161,360 donations in 22 BCs); 7,804 units for Apulia (163,881 donations in 21 BCs); 14,265 for Lombardy (470,770 donations in 33 BCs). Consistently with the productivity value, for all the BFs, the capacity value C is assumed to be equal to 50.000 units.

All the required distances, d_{ij} and $c_{jk}, i \in I, j, k \in J$, were calculated as the shortest paths on the road network. Concerning the other used modelling parameters, r and c_{max} are set equal to $20km$ and $50km$, respectively. We recall that c_{max} is the maximum distance value between two fixed facilities or a mobile unit and a fixed facility. The number of MUs is equal to 20 (see [19]).

Finally, a discussion is needed about the setting of the average accessibility threshold and the penalty parameters in the objective function. In particular:

1. Four different values have been used for β: 15, 40, 45 and 60 km. These values allow to simulate different accessibility configurations.

2. Four different values of the penalty linked to the non-compliance with the efficiency requirements (minimum productivity), λ_1: 0, 1, 10 and 100.

3. Four different values of the penalty linked with the overrunning of the maximum collection capacity, λ_2: 0, 1, 10 and 100.

4. A single very large value for the penalty linked to the self-sufficiency

requirement, $\lambda_3 = 10^6$.

The main idea behind the chosen setting of the penalty values is that the reorganization process of the BSC cannot be an abrupt and binary shift from the current state to the final solution because of work force management and process re-engineering. Instead, it needs to happen gradually over a period of time. This entails initially closing or reconverting certain BCs and progressively moving towards the desired end goal. According to the Italian guidelines, the primary objective is to ensure self-sufficiency, while it is acceptable to temporarily violate the capacity and productivity constraints, considering that a complete system reconfiguration cannot be achieved within a very short time-frame. On this basis, we fixed a very high value for λ_3 and lower but increasing values for λ_1 and λ_2, which have been empirically defined. The chosen penalty values, in combination with the other calibration parameters, enable the evaluation of both the individual and combined effects on the attainable solutions. Additionally, they provide managerial insights to the decision-maker during the gradual reorganization of the regional BSC.

Finally, we highlight that the proposed model can be used with a twofold perspective:

1. A *centralized strategy*, which considers a region as a single block;
2. A *decentralized strategy* which divides a region in areas, each of them including one or more cities and related municipalities.

The decentralized strategy is advisable in cases where there is a city which acts as a gravity center polarizing the solution, as it occurs in Campania region where the inhabitants of the area of Naples constitute about 50% of the whole population. Thus, for the Campania region, we experience also the decentralized strategy dividing the municipalities in two blocks, the first including the areas of Naples and Caserta, the second including the areas of Salerno, Benevento and Avellino.

Therefore, summarizing, considering the combinations of all the tuning parameters, we performed the following experimentation:

- 48, 16 and 16 instances of the *BSC-MFLP-CB*, *BSC-MFLP-MS*$_{sum}$ and *BSC-MFLP-SB*$_{max}$, respectively, with centralized strategy for each region;
- 48 instances of the *BSC-MFLP-CB* with decentralized strategy in Campania.

Each instance is defined by the combination of the parameters considered and by the following encode: $Region_{ID}\_\lambda_1\_\lambda_2\_\beta$. The used $Region_{ID}$ are: C for Campania; P for Apulia; L for Lombardy.

All the instances have been run on an Intel Core i7, 2.60 GHz, 16.00 GB RAM, using Gurobi 9.0.1 as MILP solver. All the *BSC-MFLP-CB* instances have been solved to optimality within 600 seconds. The two variants of the *BSC-MFLP-SB* require a higher computational effort, between 2 and 4 hours on average. For the sake of brevity, given the focus of the work on the specific real problem under investigation, we do not go into details concerning the computational aspects. However, we highlight that, the optimal solution is typically found within approximately 20% of the total computational time. Then, a considerable portion of the computational effort, approximately 80%, is expended mainly in narrowing the optimality gap.

The results for *BSC-MFLP-CB* are consolidated in tables that provide detailed information including: the instance ID (column 1); the donation rate (α - column 2); the transportation component of the objective function (column 3), calculated excluding the penalty component of the objective function and denoted as *Obj*; the values of Φ_{tot}, Ψ_{tot} (columns 4 and 5); δ (column 6), which represents the difference between the quantity of blood collected and the annual regional blood demand (D); the count of open BCs and BSs (#BCs, #BSs - columns 7 and 8); the total number of donor locations allocated to the BFs ($X_{tot} = \sum_{i \in I} \sum_{j \in J} x_{ij}$ - column 9); the number of donors allocated to MUs ($Z_{tot} = \sum_{i \in I} \sum_{j \in J} z_{ij}$ - column 10); the count of donor locations transferred to other centers, represented by transfer variables t, ($T_{tot} = \sum_{i \in I} \sum_{j \in J} \sum_{j' \in J} t_{ijj'}$ - column 11); and lastly, the computation time (in seconds - column 12).
A more comprehensive discussion is warranted for indicators Φ_{tot} and Ψ_{tot}. Specifically, they are computed as: $\Phi_{tot} = \sum_{j \in J} \Phi_j$ and $\Psi_{tot} = \sum_{j \in J} \Psi_j$. Hence, Φ_{tot} represents the cumulative difference between the minimum productivity level P_{min} and the actual number of blood units processed by each center. Meanwhile, Ψ_{tot} reflects the aggregate of blood units exceeding the maximum capacity (C) set for each center. In practice, these indicators inform on the compliance of the obtained solutions w.r.t. productivity and collection capacity requirements.

Similar tables are used for the results of the *BSC-MFLP-SB* instances. In particular, they report: the instance ID (column 1); the objective function used, i.e., expressions (2.33) and (2.60), corresponding to formulations *BSC-MFLP-SB*$_{sum}$ and *BSC-MFLP-SB*$_{max}$, respectively (column 2); the transportation cost component of the objective function (*Obj* - column 3); then, the rest of the indicators already presented for the *BSC-MFLP-CB* tables follow (i.e., Φ_{tot}, Ψ_{tot}, δ, #BCs, #BSs, X_{tot}, Z_{tot}, T_{tot}). Note that, for the *BSC-MFLP-SB*$_{sum}$ case (i.e., the upper half of the tables), those indicators are calculated as the average across the set of scenarios (e.g., $\Phi_{tot} = \frac{1}{|\Pi|}\sum_{\pi\in\Pi}\sum_{j\in J}\omega_{pi}\Phi_{jw}$). Instead, the same indicators are expressed by their worst-case values for the *BSC-MFLP-SB*$_{max}$.

We report that, to enhance readability, the values in columns 3, 4, 5, and 6 have been scaled down by a factor of 1,000 for both the *BSC-MFLP-CB* and the *BSC-MFLP-SB*.

Moreover, for some selected instances, we also provide a graphical representation of the obtained solutions, using the following graphical objects:

1. red dots for the centroids of each municipality;
2. yellow dots and blue squares for the open BCs and BSs, respectively, in the obtained solutions;
3. black, blue and red links for the donor-to-BF assignments (x-variables), donor-to-MU assignments (z-variables), and BS-to-BC assignments (q-variables), respectively;
4. yellow links for the re-allocation of a donor location between two different facilities.

Such representations allow to easily catch the effect of the topology on the number and type of open BFs (i.e., BS or BC).

2.5.2 Experimentation on Campania region

This section is devoted to Campania region and it is structured in three subsections: Subsection (2.5.2) containing the results of the *BSC-MFLP-CB* with centralized strategy; Subsection (2.5.2), devoted to the results of the *BSC-MFLP-CB* with decentralized strategy; finally, Subsection (2.5.2) concerning the results of the *BSC-MFLP-SB* variants.

Results of *BSC-MFLP-CB*-centralized strategy

In Table 2.1 we report the results of the experimentation for the centralized strategy of the Campania region varying the value of α, the value of β and the penalty values λ_1, λ_2. Two main trends emerge, the first related to productivity and capacity requirements, the second related to the accessibility.

Table 2.1: Results of *BSC-MFLP-CB* for Campania region: centralized strategy

Instance_id	α	Obj	Φ_{tot}	Ψ_{tot}	δ	#BCs	#BSs	X_{tot}	Z_{tot}	T_{tot}	time (s)
C_0_0_15		0.00	513.87	0.00	0	18	0	390	0	0	34
C_0_0_30		0.00	443.75	0.00	0	16	0	343	0	0	30
C_0_0_45		0.00	188.03	11.47	0	8	0	209	0	0	29
C_0_0_60		0.00	375.10	0.00	0	14	0	326	0	0	26
C_1_1_15		432.68	181.38	0.00	0	9	0	335	20	0	70
C_1_1_30		0.00	57.21	0.00	0	5	0	197	0	0	130
C_1_1_45		0.00	18.83	0.00	0	4	0	168	0	0	119
C_1_1_60	0.04	0.00	18.83	0.00	0	4	0	168	0	0	86
C_10_10_15		432.68	66.46	0.00	0	6	3	323	20	0	98
C_10_10_30		78.63	29.22	1.06	0	4	1	168	20	0	206
C_10_10_45		93.32	0.00	0.32	0	3	1	141	7	0	123
C_10_10_60		93.32	0.00	0.32	0	3	1	141	7	0	111
C_100_100_15		2492.18	0.00	0.00	0	4	8	340	19	1	327
C_100_100_30		687.36	0.00	0.00	0	3	4	192	3	12	567
C_100_100_45		111.80	0.00	0.00	0	3	1	142	7	0	595
C_100_100_60		111.80	0.00	0.00	0	3	1	142	7	0	364
C_0_0_15		0.00	337.35	60.35	0	13	0	383	0	0	31
C_0_0_30		0.00	167.97	69.97	0	6	0	211	0	0	106
C_0_0_45		0.00	167.97	69.97	0	6	0	211	0	0	101
C_0_0_60		0.00	167.97	69.97	0	6	0	211	0	0	102
C_1_1_15		520.00	169.56	0.12	0	10	0	335	19	1	79
C_1_1_30		5.73	36.18	0.00	0	6	0	205	7	0	121
C_1_1_45		0.00	0.00	0.59	0	4	0	141	0	0	90
C_1_1_60	0.05	0.00	0.00	0.59	0	4	0	141	0	0	80
C_10_10_15		520.00	60.66	0.00	0	7	3	321	17	3	239
C_10_10_30		77.78	21.31	0.00	0	5	1	171	20	0	194
C_10_10_45		0.00	0.59	0.00	0	4	0	141	0	0	113
C_10_10_60		0.00	0.59	0.00	0	4	0	141	0	0	134
C_100_100_15		2349.45	0.01	0.00	0	6	7	346	19	1	340
C_100_100_30		536.47	0.00	0.00	0	5	2	183	20	0	338
C_100_100_45		15.88	0.00	0.00	0	4	1	142	0	0	115
C_100_100_60		15.88	0.00	0.00	0	4	1	142	0	0	179
C_0_0_15		0.00	233.07	79.97	0	11	0	376	0	0	28
C_0_0_30		0.00	433.27	13.53	0	17	0	294	0	0	32
C_0_0_45		0.00	221.25	87.80	0	9	0	278	0	0	30
C_0_0_60		0.00	293.84	12.32	0	13	1	200	0	0	36
C_1_1_15		393.27	136.52	5.44	0	10	0	337	20	0	84
C_1_1_30		0.00	39.19	0.79	0	7	0	225	0	0	124
C_1_1_45		0.00	0.00	6.69	0	5	0	172	0	0	123
C_1_1_60	0.06	0.00	0.00	6.69	0	5	0	172	0	0	93
C_10_10_15		393.27	61.88	0.00	0	8	3	333	16	4	190
C_10_10_30		65.44	17.57	1.58	0	6	2	201	8	0	145
C_10_10_45		13.77	0.00	0.00	0	5	2	173	0	0	93
C_10_10_60		13.77	0.00	0.00	0	5	2	173	0	0	126
C_100_100_15		1980.41	0.01	0.00	0	7	7	337	19	1	270
C_100_100_30		471.96	0.02	0.00	0	6	3	198	20	0	185
C_100_100_45		13.77	0.00	0.00	0	5	2	173	0	0	148
C_100_100_60		13.77	0.00	0.00	0	5	2	173	0	0	128

Concerning the first trend, we observe that the number of BFs on the territory is largely influenced by the penalty parameters λ_1, λ_2 and the accessibility threshold β. Indeed, at equal values of β, the number of located BFs decreases as λ_1 and λ_2 increase. Recall that the latter parameters are representative of the decision-maker sensitivity to minimum productivity and maximum collection capacity requirements. Therefore, to ensure compliance, the model pushes for a lower number of active BFs, which results in increased productivity values per BC. Notably, to this end, the possibility to downgrade active BCs in BSs is also successfully exploited. For instance, for $\alpha = 0.04$, the solution obtained for the C_0_0_15 instance (the first row in the table) has 18 BCs, which reduce to 12 for the C_100_100_15 case. The latter, however, are distinguished in four BCs and eight BSs, so that the productivity requirements are met ($\Phi_{tot} = 0$). Similar considerations apply for the other values of α. Besides, in terms of collection capacity, two main aspects emerge. First, the resulting overruns (reported under the Φ_{tot} column) are generally lower w.r.t. to capacity shortages, even when the corresponding penalty λ_2 equals zero. This is clear since the the located BFs perform collection, which naturally makes this goal easier to accomplish. Second, resorting to MUs becomes a viable option when penalty increases. Indeed, if we look at the T_{tot} column in the table, we can see that, sometimes, a non-negligible number of donors' locations is served by MUs for collection, although a located BF is within their reachability threshold r (e.g., 12 in the case of the C_100_100_30 solution for $\alpha = 0.04$).

In addition, the joint effect of penalties and accessibility on the number of located BFs is significant. In particular, for $\alpha = 0.04$, we can observe that, increasing the penalty values, the number of open BFs on the regional territory varies from 18 BCs (solution of C_0_0_15 characterized by large Φ_{tot} value) to just 4 open BFs, one BS and three BCs (solution of C_100_100_60, characterized by no productivity penalty).

Interestingly, the impact of the donation rate α also yields some intriguing findings. In general, as evidenced in previous works (e.g., [53]), higher donation rates allow for stronger rationalization actions, that is, massive reduction of located BFs. This is only partially true in our case. Indeed, we often see that higher values of α (especially 0.06) lead to more facilities kept open (all other conditions being equal). Although apparently counter-intuitive, this strategic choice becomes necessary to comply with the maximum collection capacity constraints.

Concerning the second trend, we observe that higher accessibility thresholds β (that is, worse accessibility conditions for potential donors), yields a lower number of located BFs. We underline that accessibility is not explicitly discussed in the Italian decree. However, it should be seriously taken into account since it can discourage donations because of the high resulting traveling distances to reach an open BF. From our results, it is also possible to note that $\beta = 45km$ is a critical accessibility threshold beyond which the solution does not vary almost for all the instances.

We highlight that using maximum values for λ_1 and λ_2 and fixing $\beta = 15km$, regardless of the donation rate (C_100_100_15) the obtained solutions respect the EU and Italian regulation in terms of productivity. Such solutions, in practice, guarantee the same average accessibility threshold than the current situation, by resorting to the use of MUs (see the Z_{tot} column in the table).

The above observations highlight, once more, the inherent multi-objective nature underlying our problem. Nevertheless, looking at the table, we observe that, using penalty values equal to 10, accepting accessibility values higher than 30, and regardless of the donation rate, we can achieve many solutions which open no more than 7 BFs (exploiting both BCs and BSs) providing very small and acceptable values for both Φ_{tot} and Ψ_{tot}. Thus, the obtained BSC configurations represent good trade-off solutions with respect to the conflicting objectives of the *BSC-MFLP*.

In other words, if accessibility is not an issue, compliance with the system requirements (self-sufficiency, minimum productivity, and maximum collection) can be obtained together with significant savings in terms of open BFs. Nevertheless, as our results show, this requires taking ad-hoc decisions regarding BCs downgrading to BSs and a fairly limited dispatching of MUs.

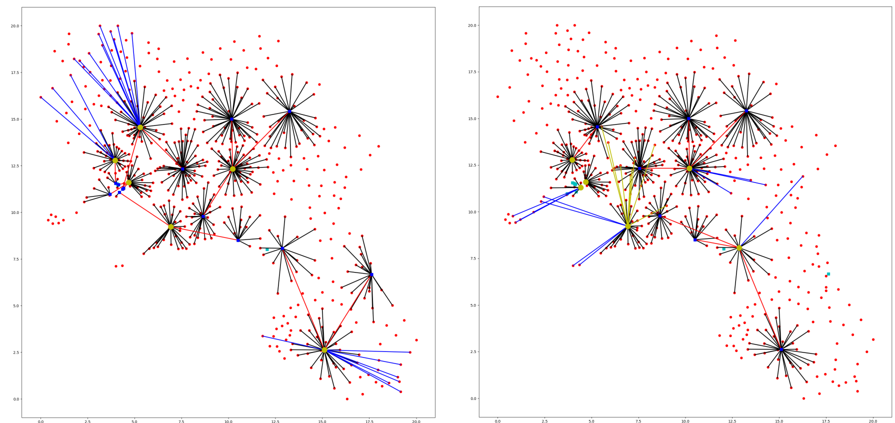
Instead, if accessibility is also envisaged, compliance may not be always guaranteed. Moreover, the achievement of very strict accessibility thresholds would result in a relatively high number of located BFs (that is, less cost-effective rationalization actions) and a more intense involvement of MUs.

Results of *BSC-MFLP-CB*-decentralized strategy

The results of the decentralized strategy are reported in Table 2.2. The obtained trends are similar to those observed in the centralized strategy, but, as expected, the number of open BFs, on average, is higher (a minimum of 4 BFs for the centralized strategy and 5 BFs for the decentralized strategy).

In order to better show the difference between the strategies, in the following we provide the graphical representation of some achieved solutions.

In particular, we show the solutions of the C_100_100_15 instance for both the strategies (centralized vs. decentralized) in Figures 2.2a and 2.2b,



(a) Centralized solution

(b) Decentralized solution

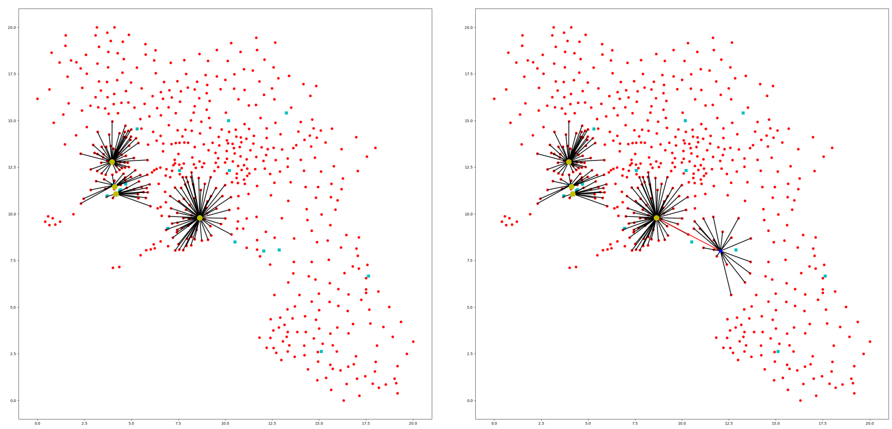
Figure 2.2: Comparison of centralized and decentralized strategy *BSC-MFLP-CB* solutions for the Campania region (C_100_100_15 instance, $\alpha = 0.05$)

respectively. These instances consider the same donation rate of the "as-is" situation ($\alpha = 0.05$) and guarantee the same average accessibility threshold ($\beta = 15km$). We chose the highest values of the penalty parameters to comply with the decree. In this case it is easy to see that, with a low value of average accessibility threshold, the centralized and decentralized strategies converge to similar solutions. Indeed, in both instances, we have 13 BFs left open (6 BCs and 7 BSs) out of 22 on the regional territory.

Fig. 2.3a and Fig. 2.3b show the solutions of the instance C_100_100_45, with the centralized and decentralized strategies, respectively. These instances are characterized by $\beta = 45km$ (i.e., a higher accessibility threshold), and the same parameter settings as above. We observe that in the centralized strategy, 5 BFs are open (4 BCs and 1 BS), while in the decentralized strategy, an additional BS is used. Moreover, in the centralized strategy all the facilities are mainly located in the area of Naples which acts as "gravity center". This solution, even though respecting the decree, is a good solution in terms of cost saving but it turns out to be penalizing for all the population in the southern and eastern areas of the region. The decentralized strategy mitigates this effect by providing a slightly more expensive - but less polarized - solution.

Similar considerations can be drawn by looking at the whole set of performed experiments.

In summary, we can conclude that if accessibility is sought (that is, low



(a) Centralized solution (b) Decentralized solution

Figure 2.3: Comparison of centralized and decentralized strategy *BSC-MFLP-CB* solutions for Campania region (C_100_100_45 instance, $\alpha = 0.05$)

threshold values for parameter β are imposed), no significant differences are devised depending on the chosen strategy (centralized or decentralized). Instead, if accessibility is not a primary issue (that is, higher threshold values are set), decentralized strategies can be winning, as they yield solutions characterized by a higher and more widespread distribution of BFs over the territory. In a nutshell, we can say that the decentralized strategy is, in some sense, more *equity-oriented*, since higher "coverage" of the population is obtained by reducing the concentration effect determined by large and populated cities.

In our opinion, this is an interesting point which stimulates a reflection on effective BSC management taking into account geographical inequalities in access. However, it has to be noted that the strategic organization of regional BSCs in Italy is not delegated to local authorities (e.g., at the provincial or health districts level), and is not driven by explicit accessibility and equity considerations. For these reasons, in the following, we validate the *BSC-MFLP-SB* focusing only on the centralized strategy.

Results of *BSC-MFLP-SB*

Table 2.3 summarizes the results of *BSC-MFLP-SB*$_{sum}$ and *BSC-MFLP-SB*$_{max}$, respectively, with the centralized strategy. For the sake of brevity, we do not provide a discussion on the effects of the penalty and average ac-

Table 2.2: Results of *BSC-MFLP-CB* for Campania region: decentralized strategy

Instance_id	α	Obj	Φ_{tot}	Ψ_{tot}	δ	#BCs	#BSs	X_{tot}	Z_{tot}	T_{tot}	time (s)
C_0_0_15	0.04	1.94	557.91	0.00	0	19	0	370	3	1	11
C_0_0_30		0.00	367.06	0.00	0	14	0	357	0	0	11
C_0_0_45		0.00	391.95	0.00	0	14	0	293	0	0	11
C_0_0_60		0.00	327.74	0.00	0	13	0	335	0	0	11
C_1_1_15		765.51	246.20	0.00	0	11	0	361	15	5	11
C_1_1_30		2.17	99.82	0.82	0	6	1	226	0	0	11
C_1_1_45		2.17	99.46	0.82	0	6	1	227	0	0	11
C_1_1_60		2.17	99.46	0.82	0	6	1	227	0	0	11
C_10_10_15		765.51	89.38	0.00	0	7	4	343	0	20	11
C_10_10_30		386.95	27.78	0.82	0	4	3	213	0	20	11
C_10_10_45		281.28	28.86	0.82	0	4	2	188	0	15	11
C_10_10_60		281.28	28.86	0.82	0	4	2	188	0	15	11
C_100_100_15		2894.33	2.61	0.00	0	5	7	367	17	3	11
C_100_100_30		955.98	0.00	0.00	0	3	7	223	0	7	10
C_100_100_45		666.63	0.00	0.00	0	3	5	199	0	4	11
C_100_100_60		666.63	0.00	0.00	0	3	5	199	0	4	11
C_0_0_15	0.05	5.19	473.48	0.00	0	18	0	364	6	4	11
C_0_0_30		0.00	367.25	5.97	0	15	0	375	0	0	11
C_0_0_45		0.00	238.58	3.40	0	11	0	248	0	0	11
C_0_0_60		0.00	373.66	0.12	0	15	0	311	0	0	10
C_1_1_15		552.96	217.74	11.94	0	11	0	361	15	5	11
C_1_1_30		0.00	88.27	0.59	0	7	0	239	0	0	11
C_1_1_45		0.00	54.51	0.59	0	6	0	216	0	0	11
C_1_1_60		0.00	54.51	0.59	0	6	0	216	0	0	11
C_10_10_15		552.96	82.28	0.12	0	8	4	355	17	3	11
C_10_10_30		480.98	24.73	0.59	0	5	2	226	0	20	11
C_10_10_45		0.00	27.88	0.59	0	5	0	159	0	0	11
C_10_10_60		0.00	27.88	0.59	0	5	0	159	0	0	11
C_100_100_15		2969.80	0.17	0.00	0	6	7	361	12	8	12
C_100_100_30		888.58	0.00	0.00	0	4	4	204	4	14	11
C_100_100_45		481.16	0.00	0.00	0	4	2	160	0	0	11
C_100_100_60		354.09	0.00	0.00	0	4	2	195	0	4	10
C_0_0_15	0.06	0.00	426.75	4.99	0	18	0	386	0	0	10
C_0_0_30		0.00	323.56	12.32	0	14	0	283	0	0	11
C_0_0_45		0.00	167.94	19.51	0	10	0	288	0	0	11
C_0_0_60		0.00	311.54	17.17	0	14	0	306	0	0	11
C_1_1_15		632.50	193.70	5.44	0	12	0	361	14	6	11
C_1_1_30		0.00	65.73	5.39	0	7	0	258	0	0	11
C_1_1_45		0.00	25.46	9.74	0	5	1	143	0	0	11
C_1_1_60		0.00	25.46	9.74	0	5	1	143	0	0	11
C_10_10_15		632.50	71.41	0.79	0	9	3	363	18	2	10
C_10_10_30		18.26	60.27	1.39	0	6	2	213	0	1	10
C_10_10_45		18.25	25.46	4.53	0	5	2	143	0	1	10
C_10_10_60		18.25	25.46	4.53	0	5	2	143	0	1	10
C_100_100_15		2755.59	1.38	0.00	0	7	8	363	12	8	14
C_100_100_30		1018.57	0.00	0.00	0	4	5	188	0	14	12
C_100_100_45		553.19	0.00	4.53	0	4	3	125	0	1	12
C_100_100_60		18.27	0.00	4.53	0	4	2	125	0	1	13

cessibility threshold values, since it would be consistent with the case-based formulation. Thus, we focus only on the differentiating aspects between the two variants of the *BSC-MFLP-SB* and between the case-based and scenario-based solutions.

Concerning the two variants of the *BSC-MFLP-SB*, we can observe that, except for the cases with λ_1 and λ_2 equal to 0, we obtain similar (or even the same) solutions in terms of the number of located BFs. As expected, for some particular instances, the number of BFs open using the *BSC-MFLP-SB*$_{max}$, as well as the transportation cost component of the objective function, are higher. This is consistent with the proposed formulation, which returns a conservative solution by looking at worst-case scenario occurrence. Nevertheless, for most of the tested instances, no significant differences are devised. This finding denotes the robustness of the proposed solutions to the variation of the donation rate α, regardless of the decision-maker aptitude towards risk.

Concerning, instead, the comparison with the case-based formulation, as we expected, the *BSC-MFLP-SB*$_{sum}$ formulation provides good compromise solutions with respect to those obtained by the *BSC-MFLP-CB* for the three considered donation rates. Indeed, all conditions being equal (i.e., with same values of all the relevant parameters), the number of BCs and BSs open in the *BSC-MFLP-SB*$_{sum}$ solution is always between the maximum and the minimum values returned by the *BSC-MFLP-CB* formulation. For instance, let us consider the results of the instance C_100_100_15. We can observe that the *BSC-MFLP-SB*$_{sum}$ formulation locates 13 BFs (5 BCs and 8 BSs), while the *BSC-MFLP-CB* provides solutions ranging from 12 open BFs (4 BCs and 8 BSs, for $\alpha = 0.04$) to 14 open BFs (7 BCs and 7 BSs, for $\alpha = 0.06$). Instead, in relation to the *BSC-MFLP-SB*$_{max}$ formulation, the objectives are similar but not identical to the worst-case solution obtained via the *BSC-MFLP-CB*. Indeed, we often see that the *BSC-MFLP-SB*$_{max}$ "dominates" the single-scenario ones. This is due to the fact that in the *BSC-MFLP-SB*$_{max}$ formulation, although our focus is on the worst-case cost minimization, all the scenarios are optimized simultaneously. Of course, this is not the case for the *BSC-MFLP-CB*. Hence, the single-scenario solution can be seen as a poor approximation of an uncertainty-aware one, if a robust solution is sought.

Therefore, in conclusion, our results prove that both the proposed approaches provide almost indistinguishable solutions for a decision-maker having a neutral aptitude towards risk (i.e., interested in minimizing the expected value cost). Instead, looking into a scenario-based solution is suggested if a risk-averse behavior is assumed.

Table 2.3: Results of *BSC-MFLP-SB*$_{sum}$ and *BSC-MFLP-SB*$_{max}$ for Campania region

Instance_id	ObjType	Obj	Φ_{tot}	Ψ_{tot}	δ	#BCs	#BSs	X_{tot}	Z_{tot}	T_{tot}	time (s)
C_0_0_15	SUM	0.00	293.92	60.91	0	12	0	388	0	0	205
C_0_0_30		0.00	348.11	1.66	0	14	0	300	0	0	223
C_0_0_45		0.00	472.70	1.66	0	18	0	356	0	0	220
C_0_0_60		0.00	293.06	5.99	0	13	0	315	0	0	191
C_1_1_15		21.89	171.64	2.77	0	10	0	337	20	0	878
C_1_1_30		5.74	42.91	3.81	0	6	0	205	7	0	2504
C_1_1_45		0.00	14.88	3.55	0	5	0	187	0	0	1345
C_1_1_60		0.00	14.88	3.55	0	5	0	187	0	0	1799
C_10_10_15		550.44	56.46	7.87	0	6	4	323	19	1	7438
C_10_10_30		91.88	27.41	0.81	0	5	2	171	20	0	3684
C_10_10_45		30.17	3.66	3.42	0	4	1	141.67	3.33	0.33	2134
C_10_10_60		30.17	3.66	3.42	0	4	1	141.67	3.33	0.33	2624
C_100_100_15		2658.40	0.45	0.00	0	5	11	341	18	2	57127
C_100_100_30		713.66	0.00	0.00	0	4	4	208.67	9.67	10.33	13224
C_100_100_45		129.96	0.00	1.51	0	3	2	125.67	14	0.33	3822
C_100_100_60		167.87	0.00	0.00	0	3	2	132.67	7.33	1.33	3460
C_0_0_15	MAX	50.89	584.59	4.99	0	22	0	390	20	0	293
C_0_0_30		0.00	601.38	0.00	0	20	0	365	0	0	394
C_0_0_45		0.00	302.59	45.07	0	11	0	333	0	0	276
C_0_0_60		0.00	368.25	0.00	0	14	0	378	0	0	334
C_1_1_15		67.33	360.79	0.00	0	14	1	390	20	0	648
C_1_1_30		4.59	66.67	0.00	0	6	0	205	7	0	4731
C_1_1_45		0.00	36.64	0.00	0	5	0	187	0	0	2734
C_1_1_60		0.00	36.64	0.00	0	5	0	187	0	0	2220
C_10_10_15		923.71	153.22	0.00	0	9	7	390	20	0	1471
C_10_10_30		73.14	44.45	0.00	0	5	2	171	20	0	5046
C_10_10_45		76.31	10.00	0.00	0	4	1	141	10	1	2633
C_10_10_60		76.31	10.00	0.00	0	4	1	141	10	1	2547
C_100_100_15		2719.40	31.72	0.00	0	6	14	390	20	0	1849
C_100_100_30		1071.01	0.03	0.00	0	4	4	181	18	2	271019
C_100_100_45		315.14	0.00	0.00	0	4	1	141	19	1	4916
C_100_100_60		240.37	0.00	0.00	0	3	2	118	0	9	3281

2.5.3 Experimentation on Apulia region

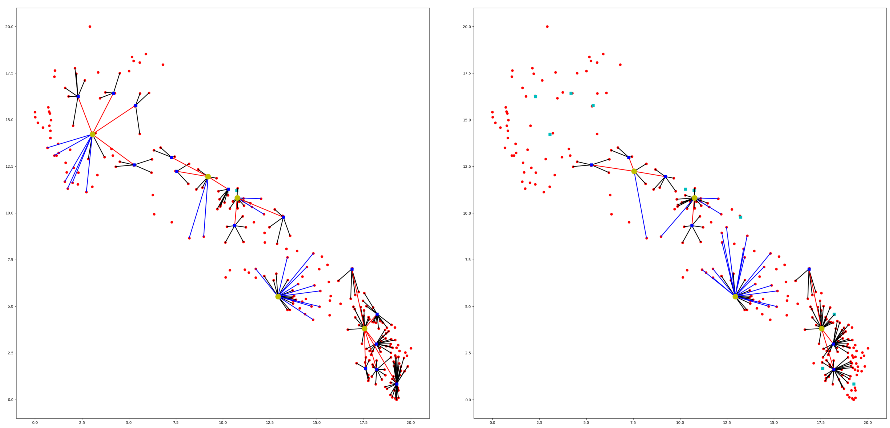
In this section, we present the results obtained for the Apulia region. In what follows, we start by focusing on the experiments performed for the case-based model (*BSC-MFLP-CB*) - Section 2.5.3. Afterwards, we present those obtained with the case-based formulations (*BSC-MFLP-SB*) in Section 2.5.3. In both cases, only a centralized strategy is investigated.

Results of *BSC-MFLP-CB*

The detailed results are displayed in Table 2.4. Of course, the interpretation of such results follows that presented for the Campania region (see Section 2.5.2). However, it is interesting to note that, differently from the previous case, for the Apulia region, considering the donation rate of the "as-is situation"($\alpha = 0.05$), the number of open BFs is not highly affected by the increase of the penalty values and of the accessibility threshold. Indeed, the overall number remains close to 21, but we can observe the downgrade of several existing BCs to BSs. At the same time, we can also note that all the solutions are characterized by quite large Φ_{tot} values, due to fact that most of the BFs do not satisfy the productivity requirement imposed by the decree. Significant reductions of the number of BFs are achievable just in the case of $\alpha = 0.06$ and high values of the average accessibility threshold.

In order to understand this fact, we look into two detailed solutions, i.e., those obtained for the P_100_100_15 and P_100_100_45 instances, displayed in Figures 2.4a and 2.4b, respectively. The former, which is obtained with an accessibility threshold $\beta = 15$ km (the "as-is" one), is characterized by the presence of 20 BFs (5 BCs and 15 BSs), spread throughout the regional territory. This number reduced to 11 BFs (4 BCs and 7 BSs), for $\beta = 45$ km. Note that, also in this case, the located facilities are well distributed in the region.

Such results allow us to make the following considerations. First, the elongated shape of the Apulia, the more sparse distribution of the centroids of the municipalities (used, we recall, as representative points where donors' are located) and the more homogeneous distribution of the population among the cities allows to have more distributed solutions in terms of BFs. Second, in most cases, the obtained solutions are not significantly different with respect to the "as-is" configuration for $\alpha = 0.04, 0.05$. This is due to the fact that, considering the number of donors, the Apulia region has a low margin with respect to the self-sufficiency requirement. Thus, relevant savings in terms of



(a) P_100_100_15 solution (b) P_100_100_45 solution

Figure 2.4: Comparison of *BSC-MFLP-CB* solutions for Apulia instances ($\alpha = 0.05$)

located BFs could be obtained only by increasing the number of donors, which is testified by the results obtained for $\alpha = 0.06$.

Results of *BSC-MFLP-SB*

The results for the scenario-based formulations, i.e., *BSC-MFLP-SB*$_{sum}$ and *BSC-MFLP-SB*$_{max}$ are in Table 2.5. Also in this case, we focus only on the differentiating aspects between the two variants of the *BSC-MFLP-SB* and between the case-based and scenario-based solutions.

Concerning the two variants of the *BSC-MFLP-SB*, we can observe that, regardless of the parameters' setting, they return very similar solutions in terms of both number and type of BFs, and objective function components. This is mainly due to the fact that, in the scenario with $\alpha = 0.04$, the number of donations does not meet the regional demand requirement. This occurrence, together with the very high value imposed for λ_3, hijacks the solution, keeping open all the existing BFs in order to collect the maximum amount of donations. The comparison of the *BSC-MFLP-SB* with *BSC-MFLP-CB* suffers from the same drawback. This is confirmed by the fact that the solutions of the *BSC-MFLP-SB* are very similar to those the *BSC-MFLP-CB* with $\alpha = 0.04$. Thus, the performed experimentation highlights that Apulia region is much more sensitive to blood donation fluctuations.

Table 2.4: Results of the *BSC-MFLP-CB* for Apulia region

Instance_id	α	Obj	Φ_{tot}	Ψ_{tot}	δ	#BCs	#BSs	X_{tot}	Z_{tot}	T_{tot}	time (s)
P_0_0_15		54.96	695.45	0	5.44	21	0	171	20	0	7
P_0_0_30		52.67	652.44	0	2.43	20	0	171	20	0	7
P_0_0_45		52.67	652.44	0	2.43	20	0	171	20	0	7
P_0_0_60		52.67	652.44	0	2.43	20	0	171	20	0	7
P_1_1_15		59.21	615.44	0	5.44	19	2	171	20	0	8
P_1_1_30		58.34	612.43	0	2.43	19	1	171	20	0	8
P_1_1_45		58.34	612.43	0	2.43	19	1	171	20	0	8
P_1_1_60	0.04	58.34	612.43	0	2.43	19	1	171	20	0	8
P_10_10_15		2237.98	175.44	0	5.44	8	13	171	20	0	8
P_10_10_30		2037.98	132.43	0	2.43	7	14	171	20	0	8
P_10_10_45		2139.18	172.43	0	2.43	8	13	171	20	0	8
P_10_10_60		2009.92	132.43	0	2.43	7	13	171	20	0	8
P_100_100_15		2237.98	135.44	0	5.44	7	14	171	20	0	9
P_100_100_30		2009.92	132.43	0	2.43	7	13	171	20	0	8
P_100_100_45		1992.89	132.43	0	2.43	7	12	171	20	0	9
P_100_100_60		1992.89	132.43	0	2.43	7	12	171	20	0	9
P_0_0_15		34.16	682.25	0	0.00	21	0	169	17	2	7
P_0_0_30		0.00	565.89	0	0.00	18	0	169	0	0	7
P_0_0_45		0.00	605.48	0	0.00	19	0	170	0	0	7
P_0_0_60		0.00	565.89	0	0.00	18	0	169	0	0	7
P_1_1_15		54.96	524.42	0	0.00	17	0	167	20	0	9
P_1_1_30		16.73	410.00	0	0.00	14	0	150	3	0	9
P_1_1_45		16.73	410.00	0	0.00	14	0	150	3	0	10
P_1_1_60	0.05	16.73	410.00	0	0.00	14	0	150	3	0	9
P_10_10_15		2237.98	86.43	0	0.00	6	13	168	19	1	11
P_10_10_30		1767.95	49.99	0	0.00	5	7	118	15	5	13
P_10_10_45		1767.95	49.99	0	0.00	5	7	118	15	5	13
P_10_10_60		1767.95	49.99	0	0.00	5	7	118	15	5	13
P_100_100_15		3751.91	27.51	0	0.00	5	15	171	20	0	11
P_100_100_30		2535.38	8.57	0	0.00	4	7	125	20	0	14
P_100_100_45		2535.38	8.57	0	0.00	4	7	125	20	0	15
P_100_100_60		2535.38	8.57	0	0.00	4	7	125	20	0	14
P_0_0_15		40.99	650.70	0	0.00	21	0	169	17	2	7
P_0_0_30		0.00	444.32	0	0.00	15	0	152	0	0	7
P_0_0_45		0.00	444.32	0	0.00	15	0	152	0	0	7
P_0_0_60		0.00	404.32	0	0.00	14	0	152	0	0	7
P_1_1_15		65.95	493.30	0	0.00	17	0	167	20	0	9
P_1_1_30		2.23	210.00	0	0.00	9	0	111	1	0	10
P_1_1_45		2.23	210.00	0	0.00	9	0	111	1	0	10
P_1_1_60	0.06	2.23	210.00	0	0.00	9	0	111	1	0	10
P_10_10_15		1741.36	147.20	0	0.00	8	10	165	20	0	11
P_10_10_30		986.45	50.58	0	0.00	5	2	102	19	1	15
P_10_10_45		986.45	50.58	0	0.00	5	2	102	19	1	14
P_10_10_60		986.45	50.58	0	0.00	5	2	102	19	1	15
P_100_100_15		4240.39	11.23	0	0.00	5	15	171	20	0	11
P_100_100_30		2020.21	0.05	0	0.00	4	6	126	18	2	23
P_100_100_45		2020.21	0.05	0	0.00	4	6	126	18	2	19
P_100_100_60		2020.21	0.05	0	0.00	4	6	126	18	2	19

Table 2.5: Results of *BSC-MFLP-SB*$_{sum}$ and *BSC-MFLP-SB*$_{max}$ for Apulia region

Instance_id	ObjType	Obj	Φ_{tot}	Ψ_{tot}	δ	#BCs	#BSs	X_{tot}	Z_{tot}	T_{tot}	time (s)
P_0_0_15		190.00	676.14	0	1.81	21	0	169.67	18	1.33	34
P_0_0_30		186.00	635.30	0	0.81	20	0	170.33	7.33	0.67	33
P_0_0_45		172.00	635.73	0	0.81	20	0	171	6.67	0	33
P_0_0_60		172.00	635.73	0	0.81	20	0	171	6.67	0	34
P_1_1_15		262.35	596.13	0	1.81	19	2	169.67	18	1.33	43
P_1_1_30		209.45	595.72	0	0.81	19	1	171	6.67	0	42
P_1_1_45		209.45	595.72	0	0.81	19	1	171	6.67	0	43
P_1_1_60		209.45	595.72	0	0.81	19	1	171	6.67	0	41
P_10_10_15	SUM	1779.94	198.03	0	1.81	9	12	169	18	2	52
P_10_10_30		2231.86	119.53	0	0.81	7	13	157.67	6.67	13.33	51
P_10_10_45		2231.86	119.53	0	0.81	7	13	157.67	6.67	13.33	52
P_10_10_60		2231.86	119.53	0	0.81	7	13	157.67	6.67	13.33	48
P_100_100_15		3464.85	107.11	0	1.81	7	14	171	20	0	46
P_100_100_30		3164.63	95.60	0	0.81	7	13	171	20	0	50
P_100_100_45		3164.63	95.60	0	0.81	7	13	171	20	0	53
P_100_100_60		3164.63	95.60	0	0.81	7	13	171	20	0	48
P_0_0_15		636.57	695.45	0	5.44	21	0	171	20	0	37
P_0_0_30		689.64	652.44	0	2.43	20	0	171	20	0	32
P_0_0_45		689.64	612.44	0	2.43	19	0	171	20	0	32
P_0_0_60		689.64	652.44	0	2.43	20	0	171	20	0	31
P_1_1_15		698.07	615.44	0	5.44	19	2	171	20	0	49
P_1_1_30		719.69	612.43	0	2.43	19	1	171	20	0	45
P_1_1_45		691.19	612.43	0	2.43	19	0	171	20	0	44
P_1_1_60		691.19	612.43	0	2.43	19	0	171	20	0	43
P_10_10_15	MAX	1937.67	215.44	0	5.44	9	12	171	20	0	50
P_10_10_30		2161.01	172.43	0	2.43	8	13	171	20	0	48
P_10_10_45		2131.95	172.43	0	2.43	8	12	171	20	0	47
P_10_10_60		2161.01	172.43	0	2.43	8	13	171	20	0	50
P_100_100_15		2766.05	135.44	0	5.44	7	14	171	20	0	50
P_100_100_30		2454.30	132.43	0	2.43	7	13	171	20	0	57
P_100_100_45		2454.30	132.43	0	2.43	7	13	171	20	0	56
P_100_100_60		2454.30	132.43	0	2.43	7	13	171	20	0	54

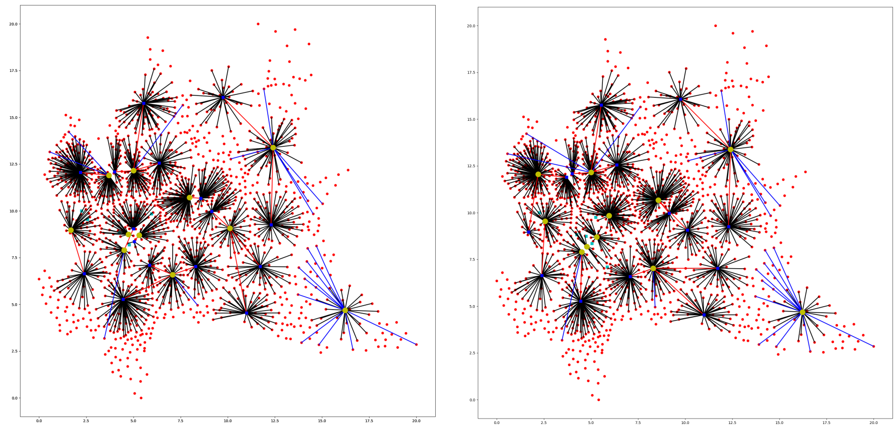
2.5.4 Experimentation on Lombardy region

The last step of our empirical analysis consists of additional experiments performed on Lombardy, the largest Italian region. Following the above adopted scheme, we split the results in two parts: Section (2.5.4), containing the results of the *BSC-MFLP-CB* with the centralized strategy; Section (2.5.4), concerning the results of the *BSC-MFLP-SB* variants.

Results of *BSC-MFLP-CB*

Table 2.6: Results of *BSC-MFLP-CB* for Lombardy region

Instance_id	α	Obj	Φ_{tot}	Ψ_{tot}	δ	#BCs	#BSs	X_{tot}	Z_{tot}	T_{tot}	time (s)
L_0_0_15	0.04	320.00	624.94	18.15	101.16	24	0	1230	20	0	583
L_0_0_30		320.00	607.85	6.68	101.16	24	0	1230	20	0	537
L_0_0_45		320.00	608.03	6.04	101.16	24	0	1230	20	0	538
L_0_0_60		320.00	608.03	6.04	101.16	24	0	1230	20	0	571
L_1_1_15		320.00	631.16	0.00	101.16	25	0	1230	20	0	1851
L_1_1_30		320.00	631.16	0.00	101.16	25	0	1230	20	0	1738
L_1_1_45		320.00	631.16	0.00	101.16	25	0	1230	20	0	1764
L_1_1_60		320.00	631.16	0.00	101.16	25	0	1230	20	0	1761
L_10_10_15		2168.86	315.11	0.00	101.16	17	14	1230	20	0	2021
L_10_10_30		2168.86	315.11	0.00	101.16	17	14	1230	20	0	2011
L_10_10_45		2177.66	314.23	0.00	101.16	17	12	1230	20	0	2049
L_10_10_60		1582.73	273.73	0.00	101.16	16	11	1230	20	0	2038
L_100_100_15		4204.30	111.16	0.00	101.16	12	17	1230	20	0	2056
L_100_100_30		7825.90	81.27	0.00	101.16	9	20	1230	20	0	2158
L_100_100_45		7193.40	64.94	0.00	101.16	9	15	1230	20	0	2143
L_100_100_60		7193.40	64.94	0.00	101.16	9	15	1230	20	0	2064
L_0_0_15	0.05	360.00	667.72	9.01	8.95	26	0	1230	20	0	609
L_0_0_30		360.00	601.97	4.40	8.95	26	0	1230	20	0	579
L_0_0_45		360.00	593.13	4.17	8.95	26	0	1230	20	0	593
L_0_0_60		360.00	517.29	38.34	8.95	26	0	1230	20	0	575
L_1_1_15		449.02	530.98	0.00	8.95	25	2	1230	20	0	1817
L_1_1_30		363.55	506.45	0.00	8.95	24	0	1230	20	0	1884
L_1_1_45		363.55	506.45	0.00	8.95	24	0	1230	20	0	1742
L_1_1_60		363.55	506.45	0.00	8.95	24	0	1230	20	0	1830
L_10_10_15		2110.60	128.67	40.27	8.95	16	10	1230	20	0	2171
L_10_10_30		1685.56	168.51	1.93	8.95	15	7	1230	20	0	2194
L_10_10_45		1689.20	169.08	0.00	8.95	15	7	1230	20	0	2355
L_10_10_60		1724.19	165.71	4.87	8.95	14	8	1230	20	0	2013
L_100_100_15		4517.80	24.42	0.00	8.95	11	15	1230	20	0	2247
L_100_100_30		4011.30	24.02	0.00	8.95	11	14	1230	20	0	2230
L_100_100_45		4011.30	24.02	0.00	8.95	11	14	1230	20	0	2243
L_100_100_60		4011.30	24.02	0.00	8.95	11	14	1230	20	0	2440
L_0_0_15	0.06	0.00	670.72	0.00	0.00	30	0	1213	0	0	622
L_0_0_30		0.00	472.23	40.53	0.00	23	0	1154	0	0	538
L_0_0_45		0.00	455.96	69.47	0.00	22	0	1175	0	0	525
L_0_0_60		0.00	383.31	78.97	0.00	19	0	1108	0	0	518
L_1_1_15		6.19	226.70	9.75	0.00	17	0	1135	20	0	2209
L_1_1_30		0.00	50.23	3.70	0.00	12	0	908	0	0	4229
L_1_1_45		0.00	50.23	3.70	0.00	12	0	908	0	0	4469
L_1_1_60		0.00	50.23	3.70	0.00	12	0	908	0	0	4850
L_10_10_15		618.12	122.04	1.69	0.00	15	3	1150	18	2	9104
L_10_10_30		374.43	10.32	1.43	0.00	11	0	820	19	1	7821
L_10_10_45		374.43	10.32	1.43	0.00	11	0	820	19	1	8231
L_10_10_60		374.43	10.32	1.43	0.00	11	0	820	19	1	6175
L_100_100_15		3122.59	0.04	0.00	0.00	12	13	1172	18	2	12659
L_100_100_30		594.23	0.02	0.00	0.00	11	1	836	12	8	13727
L_100_100_45		594.23	0.02	0.00	0.00	11	1	836	12	8	15512
L_100_100_60		594.23	0.02	0.00	0.00	11	1	836	12	8	13305



(a) L_100_100_15 solution (b) L_100_100_45 solution

Figure 2.5: Comparison of *BSC-MFLP-CB* solutions for Lombardy region ($\alpha = 0.05$)

The results obtained for these experiments are reported in Table 2.6. As for the Campania and Apulia regions, similar patterns are again observed concerning productivity and capacity restrictions, on the one hand, and with accessibility, on the other.

Therefore, here we only focus on an (novel) aspect of great relevance emerging from the results, which shed light on the proper functioning of the regional BSC. Lombardy, we recall, is the largest region in Italy in terms of population, and the region where most medical treatments are conducted as people migrate from all over Italy to seek care. However, Lombardy only manages to meet the self-sufficiency goal when the value of α equals 0.06 (note the non-zero values for δ in the Table). We want to emphasize that the assumed data of 5 donations per 1000 inhabitants is an aggregate figure for Italy, and the specific reference value for Lombardy is actually higher. Nevertheless, we highlight the importance of implementing policies aimed at maintaining this value to ensure the smooth operation of the system. Indeed, if the donation rate was to fall below the threshold of 6 donations per 1000 inhabitants, the region could face difficulties in sourcing the necessary blood for self-sufficiency, making it essential to purchase blood units from other Italian regions.

Figures 2.5a and 2.5b, representing the solutions obtained for the L_100_100_15 and L_100_100_45 instances (for $\alpha = 0.05$), give an insight into this situation. The solution obtained for L_100_100_15 solution utilizes a total of 26 BFs (comprising 11 BCs and 15 BSs) strategically scattered across the regional territory. Alternatively, by raising the accessibility threshold to $45km$, the L_100_100_45 solution achieves similar results with only 1 BSs less. This demonstrates that, despite increasing the accessibility threshold, the solution remains relatively unchanged due to the inability to meet the entire regional demand for blood.

Results of *BSC-MFLP-SB*

Table 2.7 present the results for two variations of the centralized strategy: *BSC-MFLP-SB*$_{sum}$ and *BSC-MFLP-MS*$_{max}$. When considering the two variants of *BSC-MFLP-SB*, we observe that regardless of the parameter settings, they produce very similar solutions in terms of both the number and type of blood facilities (BFs). This similarity is primarily attributed to the fact that, in scenarios with $\alpha = 0.04$ and $\alpha = 0.05$, the number of donations fails to meet the regional demand for blood (self-sufficiency). To maximize the amount of collected donations, the model tends to keep all existing BFs open due to the high value assigned to λ_3.

A similar observation holds when focusing on the comparison between *BSC-MFLP-SB* and *BSC-MFLP-CB*. The solutions generated by *BSC-MFLP-SB* closely resemble those of *BSC-MFLP-CB* when α is set equal to 0.05. This further demonstrates that the Lombardy region, like Apulia, is highly susceptible to fluctuations in blood donations, necessitating the maintece of a sufficiently high donation rate to address this issue. In contrast, Campania is less affected by these variations.

Table 2.7: Results of *BSC-MFLP-SB*$_{sum}$ and *BSC-MFLP-SB*$_{max}$ for Lombardy region

Instance_id	ObjType	Obj	Φ_{tot}	Ψ_{tot}	δ	#BCs	#BSs	X_{tot}	Z_{tot}	T_{tot}	time (s)
L_0_0_15	SUM	219.89	573.33	53.25	29.76	24	0	1230.0	13.33	0	3070
L_0_0_30		219.89	608.87	7.08	29.76	26	0	1230.0	13.33	0	3048
L_0_0_45		219.89	515.81	31.81	29.76	23	0	1230.0	13.33	0	3061
L_0_0_60		219.89	664.91	0.00	29.76	28	0	1230.0	13.33	0	3326
L_1_1_15		551.24	591.87	0.00	29.76	26	3	1229.67	13.33	0	11729
L_1_1_30		551.24	587.83	0.00	29.76	26	3	1230.0	13.33	0	14161
L_1_1_45		431.78	626.89	0.00	29.76	23	3	1230.0	13.33	0	13307
L_1_1_60		229.23	442.78	85.00	29.76	20	0	1230.0	13.33	0	11907
L_10_10_15		1883.75	352.03	0.00	29.76	20	11	1230.0	13.33	0	17088
L_10_10_30		1989.92	347.56	0.00	29.76	19	13	1229.67	13.33	0	13401
L_10_10_45		1989.92	318.95	0.00	29.76	19	13	1230.0	13.33	0	14124
L_10_10_60		1746.23	309.55	0.00	29.76	13	9	1230.0	13.33	0	15549
L_100_100_15		2574.01	50.64	4.36	29.76	12	13	1227.33	20	0	20122
L_100_100_30		2997.43	38.66	0.00	29.76	10	16	1227.33	20	0	18490
L_100_100_45		3012.99	35.06	0.00	29.76	10	16	1227.33	20	0	18918
L_100_100_60		3231.88	30.11	4.90	29.76	9	14	1227.33	20	0	17425
L_0_0_15	MAX	287.43	639.32	0.00	101.16	25	0	1230.0	20	0	3104
L_0_0_30		287.43	660.87	59.70	101.16	24	0	1230.0	20	0	2907
L_0_0_45		287.43	648.29	7.12	101.16	25	0	1230.0	20	0	3215
L_0_0_60		287.43	578.06	16.89	101.16	23	0	1230.0	20	0	2955
L_1_1_15		688.11	711.16	0.00	101.16	27	3	1230.0	20	0	12298
L_1_1_30		688.11	671.16	0.00	101.16	26	3	1230.0	20	0	12860
L_1_1_45		716.29	591.16	0.00	101.16	23	4	1230.0	20	0	13299
L_1_1_60		287.43	551.16	0.00	101.16	23	0	1230.0	20	0	12208
L_10_10_15		1975.67	432.19	0.00	101.16	20	13	1230.0	20	0	13950
L_10_10_30		2030.96	391.16	0.00	101.16	19	13	1230.0	20	0	12816
L_10_10_45		2030.96	391.16	0.00	101.16	19	13	1230.0	20	0	14916
L_10_10_60		1696.06	198.37	33.60	101.16	13	9	1230.0	20	0	15710
L_100_100_15		2966.70	111.80	0.00	101.16	12	12	1230.0	20	0	19130
L_100_100_30		3915.52	86.13	0.00	101.16	11	14	1230.0	20	0	20462
L_100_100_45		3892.48	74.97	0.00	101.16	11	14	1230.0	20	0	19091
L_100_100_60		5306.65	35.85	0.00	101.16	9	14	1230.0	20	0	22184

2.6 Conclusions

This Chapter investigates the reorganization of regional blood management systems from a strategic perspective. It is directed forward the closing or reconverting of blood facilities currently operating in a region, finding good trade-off solutions between the need of attracting donors (self-sufficiency objective) and the need of reducing the system management costs (efficiency objective). To this aim, we proposed two MILP formulations tackling a multi-echelon facility location problem in a case-based and scenario-based perspective.

The proposed formulations allow, on the one hand, to provide insights on the impact of several strategies and policies in the reorganization of the regional BSC and, on the other, to perform a sensitivity of the proposed solutions to the variation of several parameters and different scenarios. Thus, it can represent a useful decision support system for the competent authorities and stakeholders.

The model has been validated on three real cases related to the Campania, Apulia and Lombardy regions, for which several configurations have been simulated. The experimentation has allowed to derive several suggestions for an effective reorganization of the the "as-is" situations in the regions under investigation and has highlighted the importance of taking into account the specific demographic and topological features of a territory, as well as the structural deficiencies of the system.

Future research directions naturally include the integration of tactical and operational issues, such as workforce planning and mobile unit routing. As a consequence, the need of developing either heuristic and meta-heuristic solution methods or advanced multi-stage stochastic programming approaches could arise. Moreover, we recall that our methodology is specifically designed to address the requirements of a European Directive, which encompasses not only Italy but also other countries. As a result, we believe that our proposed modeling framework holds relevance for other countries that are currently undergoing or considering a reorganization of their BSC. Therefore, future work perspective will include the adaptation of our modeling framework to different national contexts, incorporating and tailoring different BSC peculiarities.

Chapter 3

Integrated Operating Room Planning and Scheduling (*IORPS*): an emergency responsive-based offline approach

As the global population ages, the demand for health services, particularly surgical procedures, has significantly increased. Efficient management of hospital resources, especially in the operating rooms (*ORs*), becomes fundamental due to their substantial contribution to hospital costs and revenues. In this context, this chapter investigate into the challenges faced in elective *OR* scheduling. By focusing on the Integrated Operating Room Planning and Scheduling (*IORPS*) for elective patients at a local hospital in Naples, this work aims to provide an automated solution for weekly surgical scheduling, ensuring maximum efficiency while addressing unpredictable events. We observe that in numerous contributions, *IORPS* not only encompass the surgery schedule but also necessitate the utilization of medical resources, personnel, and more. However, the challenge of determining the specific day and starting time for surgery remains within the scope of *IORPS*.

3.1 Introduction

The global aging of the population is a well-known demographic trend, with reports indicating a steady increase in the number of people aged 60 years and over [141]. According to a study by United Nations [289], the number of people in this age group reached 901 million in 2015, reflecting an increase of 48% since 2000. By 2030, this number is expected to reach 1.4 billion, and by 2050, it is projected to more than double to nearly 2.1 billion. This demographic shift is likely to increase the demand for health products and services, particularly in higher age groups where health needs are more frequent and significant [302]. Addressing this demand represents a multifaceted challenge, further complicated by the fact that public hospitals must contend with frequent regulatory changes from healthcare authorities, increasingly stringent quality requirements, and budget cuts, resulting in constraints on both human and material resources [249].

In this context, the vital role of hospitals, particularly the operating theatres (*OT*) and their associated operating rooms (*OR*s), in the healthcare supply chain and the effective management of patient flow, is unequivocally acknowledged [124]. This importance is further emphasized by the statistic revealing that around 60% of all patients necessitate the *OT* at some point during their hospitalization [187]. At the same time, numerous studies recognize the *OT* as the most financially intensive assets for both public and private hospitals. Surgical expenditures typically constitute around 40% of the overall resource budget, and surgical procedures contribute to approximately 67% of the total revenue [71]. In [252], the expenses associated with the utilization of valuable resources, such as surgeons, staff, and equipment, are estimated to range from 20 to 116 euros per minute.

In this context, as highlighted in [292], the need for effective and efficient management of *OR*s, encompassing planning and scheduling, arises. Operations Research can play a relevant role in supporting hospital executives to enhance hospital operations, reduce costs, and improve patient satisfaction, under limited resource constraints and diverse stakeholder priorities [105].

On this basis, in the following of this section, we establish the context for our research. Firstly, we provide an overview of the background related to *OT* planning and scheduling problems in order to position our work in Operations Research field. Next, we present the real problem that motivates our present work. Finally, we outline the specific methodological contributions of this study, and we provide an overview of its structure.

3.1.1 *OT* management background

The literature categorizes the main Operations Research contributions on *OT* management with respect to two main issues [63, 124]: decision levels and *OR* management strategies or procedures.

Concerning the first issue, existing literature identifies three hierarchical decision levels based on the decision-making time horizon (long-term, medium-term, short-term): strategic, tactical, and operational [321]. At a strategic level, the decision problem involves the forecasting of medium and long-term demand and the distribution of the *OR* time among the surgical specialties and/or surgeons (Capacity Planning Problem, Capacity Allocation Problem); at a tactical level, the decision problem deals with the definition of cyclic *OR* schedules, i.e., assigning *OR* time blocks to surgical specialties, usually with homogeneous lengths of half day or full day (Master Surgical Scheduling Problem); at the operational level, the decision problem consists in determining the date, the operating room and the starting time of each surgery, accounting also for related estimated completion time and medical resources allocation.

The operational level can be conceived as composed of two decision sub-problems [317]: the Advance Scheduling Problem (*ADP*), also referred to as *OR* planning in [136], consisting in assigning patients to specific surgical block (i.e., *OR* and day) within a planning horizon (usually a week); the Allocation Scheduling Problem (*ALP*), consisting in sequencing surgeries within each block. When tackled together, these the two sub-problems constitute the core of the so called Integrated Operating Room Planning and Scheduling (*IORPS*) problem [136]. However, we observe that in many contributions, *IORPS* generally refers to more complex variants including also decision about utilization of medical resources and personnel.

Concerning the second issue, the literature distinguishes three *OR* management strategies used in practice: open, block and modified-block [108, 249, 321]. In the open strategy, all the surgeries can be scheduled into any *OR* and in any day of the planning horizon, following a predetermined performance criteria. Under the block strategy, the *OR* scheduling is organized in surgical blocks, where a block corresponds to a set of consecutive time periods (hours, half-hours, etc.). Each block is pre-assigned to a day of the planning horizon and to a specific surgical specialties. Then, a surgery can be performed only in a surgical block dedicated to the corresponding specialty. Finally, the modified-block strategy combines the previous ones, i.e., if there is under-utilization of a pre-assigned *OR* block in a day, then this block can be exploited by a different surgical specialty in order to maximize the *OR* usage. The choice of the strategy is mainly motivated by the equipment of surgery department and by the personnel shifts policies of each hospital.

The two issues just discussed underlie the categorization of surgical patients, typically organized into emergency, urgent, and elective classes [124]. Emergency and urgent cases necessitate surgical intervention within two hours or few hours of the decision to operate, respectively. Elective cases, instead, are individuals who are admitted for a planned or scheduled medical procedure or a treatment that can be planned in advance. Elective patients are usually inserted into a waiting list and are scheduled through an off-line (ex-ante) planning process according to internal rules, performance criteria, and scheduling policies, typically one or two weeks before the surgery. Generally an initial schedule for elective patients represents a baseline for forthcoming elective surgeries.

Thus, in broad terms, *IORPS* decision are typically directed by elective patient cases, given their impact on system performance [26], whereas the management of non-elective patients represents a source of uncertainty generally addressed reactively at the time of occurrence.

However, over the past decade, several contributions have underscored the significance of proactively considering potential emergency arrivals, which are not-preemptive by nature, when defining (off-line) the weekly surgery schedule [294, 187]. Recognizing this importance is crucial because prompt delivery of emergency treatment is essential; any delays can result in increased morbidity and mortality [187].

Emergency surgeries are characterized by different levels of urgency, each associated with distinct Waiting Time Targets (*WTT*s), where the *WTT* is defined as the time within which the hospital must initiate an emergency surgery [26]. To optimize the system's responsiveness, one policy is to allocate a dedicated *OR* solely for emergency surgeries [94]. However, as noticed in [307], such option is not a good solution for hospitals with a limited number of *OR*s. Alternatively, emergency procedures can be integrated into the elective *OR* schedule, waiting a suitable moment between ongoing elective surgeries [292, 26]. Another option is to proactively incorporate buffer times into the baseline schedule, similarly to slack times introduced in [6], facilitating the seamless integration of unforeseen emergency arrivals on the day of surgery [147, 187]. The second and the the third policies are referred to as shared-room policies in literature. Their main distinctions lie in the distribution of the emergency insertion moments, referred to as 'break-in-moments' *BIM*, throughout the day and in the duration of the break-in interval (*BII*), which represents the maximum time between two consecutive *BIM*s for the available operating room *OR*s.

As hinted above, the provided discussion is preparatory to precisely contextualize the specific problem under investigation, presented in next section, in the *IORPS* literature.

3.1.2 Motivation

In this chapter, we address the *IORPS* problem faced by a local hospital in Naples, whose future broad-ranging goal is adopting an automatic tool, in place of the actual manual procedure, to provide the weekly surgery scheduling of its elective patients (both inpatients and outpatients). The objective is determining the schedule that maximizes the number of performed surgeries according to available resources, medical requirements and priority policies, while accounting for unforeseen emergencies.

In a nutshell, the specific *IORPS* problem under investigation, introduced in [46], can be schematized as follows. The planning horizon is fixed to one week (six working days) and each *OR* has a limited working time per day. The weekly schedule should assign a day, an *OR*, and a starting time to each elective surgery, while incorporating information about surgery duration and resulting room occupancy.

The elective patients, each of them associated to one of the six surgical specialties at the hospital, has to be selected from a known in advance waiting list for scheduling. Each elective patient surgery is characterized by three elements: clinical specialty, duration, and clinical severity. The clinical specialty indicates the equipment/instruments and skills/surgeons required for the surgery. The duration is an estimate of the required operating time and is generally defined based on the average duration of past surgeries. Finally, the clinical severity is a proxy for the complexity of the surgery and the patient's clinical status. The sequencing of elective patients each day should prioritize the scheduling of the most severe and time-consuming surgeries ahead of others.

Regarding non-elective patients, the hospital follows a shared operating room policy. Once an emergency arrives, the patient will be stabilized in the emergency department with initial treatments before being transferred to an *OR*. Thus, according to this setting, emergency patients who arrive during operating hours compete with elective surgeries for *OR* capacity.

The hospital seeks a schedule guaranteeing strict responsiveness to potential emergency arrivals, i.e., the timely response to the most urgent potential emergency has to be guaranteed whenever it arrives during the open hours [187]. This requirement is common for emergencies in cardiology, trauma, and gynecology, which are the specific cases dealt by the hospital. For all these emergencies, the hospital sets a *WTT* of 20 minutes and, consistently, wants to keep the estimated *BIM* on the baseline elective schedule within 20 minutes, as well.

Currently, the hospital uses a block strategy scheme for the four available *OR*s. However, the hospital is interested in having a tool that can manage other scheduling strategies, such as block-modified and open, in order to evaluate possible temporary re-configurations of the *OT* and to simulate the performance of potential upgrades and expansions.

Summarizing, the *IORPS* problem under investigation deals with the operational level, all the *OR* management strategies and include strict *BII, BIM* and *WTT* requirements for emergency management in a shared operating room policy.

More details will be given in Section 3.3 devoted to problem definition and formulation and in Section 3.4.1 devoted to hospital instance description.

3.1.3 Contribution of the work

From a methodological point of view, the *IORPS* has been generally conceived as a combinatorial optimization problem with multiple objectives, such as: maximizing the utilization of *OR*s, minimizing the number of the postponed surgeries, minimizing the overtime work of the *OR*s and of the surgeons, maximizing the patient satisfaction, etc.. This problem, as widely known, is NP-Hard and it can lead to some severe inefficiencies in the *OR* management if it is not solved appropriately [226, 71].

As widely discussed in the following, the literature on this topic is extensive and diverse, with contributions varying in both the solution approaches employed and the specific real problems investigated. However, there remains a limited body of literature addressing the integration of emergency management during the off-line planning of elective patients within a shared room policy. To the best of the author's knowledge, works explicitly addressing this aspect, albeit in a less stringent version than ours, are discussed in [292, 147, 187].

On this basis, this work, building upon the one presented in [46], aims to contribute to the literature in the field by the following advancements:

1. presenting a compact and efficient Integer Linear Programming (*ILP*) formulation of the *IORPS* problem under investigation, integrating the 20-minutes strict responsiveness requirement into off-line planning for emergency management;

2. introducing a novel objective function based on the definition of a surgery priority index, envisaging surgery severity and duration, to meet the daily scheduling priority criterion used by the hospital;

3. proposing a unified *ILP* modeling framework allowing the hospital to easily evaluate the usage of the different *OR* management strategies.

The proposed objective function differs from typical cost-based ones proposed in the literature, reflecting the obligation of accommodating patients in Italian public hospitals, despite budget constraints.

It is noteworthy that hospitals do not often modify their scheduling policies, given that such changes entail long-term strategic decisions. Specifically, due to specialized equipment needed, in most cases hospitals use block or modified-block strategies. Instead, open strategies are infrequently used as they require operating rooms equipped with setting for all specialties. However, as occurred during the COVID-19 pandemic, the surgery demand was characterized by a great variability, and many hospitals needed to frequently re-think and adapt their scheduling policy and *OR* usage on the basis of a contingency plan. Thus, the motivation behind the latest contribution consists in supporting hospital executives in these decisions. Hence, the motivation behind the latest contribution lies in providing support to hospital executives in making these decisions.

In light of the preceding discussion and following the classification proposed in [256], our work falls within the category of practice-oriented research. In this context and in alignment with the hospital's comprehensive objective of implementing an automated solution for their *IORPS* problem, the proposed approach has been integrated into a prototype tool utilizing off-the-shelf optimization solvers. This tool, currently undergoing the validation phase at the hospital, has been tested on real instances provided by the local hospital in Naples. The results, evaluated in terms of both objective function and computation time, demonstrate its applicability as an effective decision support tool for the hospital.

3.1.4 Structure of the work

The paper is organized as follows. In Section 3.2 we provide the literature review on *IORPS* problem. Section 3.3 details the specific problem tackled in this work and is devoted to related mathematical formulation. Section 3.4 reports and discusses the experimental results. Finally, in Section 3.5 some conclusions and future research directions are outlined.

3.2 Literature Review

Optimization problems in healthcare have received considerable attention for over three decades [231]. Moreover, the literature addressing *IORPS* is notably extensive, as evidenced by numerous survey works on the subject published in the last 15 years [63, 124, 256, 321, 226].

In this section, our focus is on the *IORPS* at the operational level, which, as mentioned above, encompasses two decision sub-problems, i.e., the *ADP* and the *ALP*. Considerable research has explored *IORPS* in its more comprehensive form, addressing both sub-problems simultaneously, or tackling each sub-problem individually. For the sake of the brevity, we concentrate on contributions from the former group in the subsequent discussion, while we do not revise the literature related to each sub-problem. Interested readers are addressed to the following state-of-the-art contributions in the field: [125, 108, 60, 61] for the *ADP* and [292, 95, 179, 26] for the *ALP*.

Two main approaches have been proposed in the literature for the *IORPS*: sequential and integrated. The sequential approach assumes a hierarchy between the two sub-problems and solves them in sequence. Significant contributions on the topic are presented in [144, 109, 23, 147, 168, 320]. In particular, the work presented in [5] envisages also a rolling horizon for re-scheduling activities.

However, it is widely recognized that sequential approaches can have significant drawbacks in defining efficient and effective solutions due to the mutual dependency between the two sub-problems [63]. Integrated approach simultaneously solves the two *IORPS* sub-problems. Literature on the topic is rich and, as discussed in [272, 250], from the methodological point of view, contributions mainly fall in the area of mathematical programming, heuristics and metaheuristics, constraint programming and decomposition techniques.

Regardless of the methodology employed, contributions mainly differ in the way they consider two types of uncertainties, specifically, those pertaining the surgery duration surgery and non-elective patients management (emergency arrivals).

Concerning the first issue, as noted in the cited surveys, a majority of contributions adopt a deterministic approach to surgery duration. This choice is often justified by the inherent complexity of the *IORPS* problem, even in its deterministic form. Moreover, as discussed in [294, 245] and recently in [27], a pragmatic approach is commonly adopted by the majority of hospitals in formulating off-line the baseline schedule for elective patients. Thus, deterministic surgery duration values are regularly considered. These values typically derive from the time series average duration of specific surgeries or are directly provided by surgeons who are aware of severity of the surgery and the clinical status of the patient. This choice is coherent with the understanding that the baseline schedule serves as a provisional plan for surgeries, subject to adjustments based on unforeseen events such as prolonged durations, cancellations, or the arrival of emergency cases.

Numerous contributions have addressed the off-line deterministic version of the *IORPS*, with notable contributions from [249, 244, 176, 298, 177, 136, 250, 27], which, in their chronological sequence, stand as essential references in the literature. It is worth noting that the majority of these contributions address a richer and more complex variant of the *IORPS*, often driven by practical applications. However, we highlight that none of them explicitly incorporate the management of non-elective patients.

The challenge posed by non-elective patients has been treated in two different ways. One approach involves proactively establishing an off-line baseline schedule for elective patients that is responsive to unforeseen emergency arrivals. Alternatively, there is the development of on-line re-scheduling algorithms, as made in [26], aimed at efficiently handling such situations. Obviously, the two approaches can be used in tandem.

In the following discussion, our focus will be on the first group of approaches. While managing uncertainty naturally implies the use of stochastic approaches, the literature on off-line stochastic *IORPS* remains limited, as discussed in the cited surveys. Noteworthy contributions in this context include in [159, 283]. Remaining within the deterministic framework, the management of emergency arrivals has led to the conceptualization of the *BIM* and *BII* concepts in [292]. Starting from this work, several contributions have explored *BII* and *BIM*, e.g., [112, 147]. Additionally, we reference the work presented in [187], where the introduction of buffer times addresses some limitations of the *BII* and *BIM* approaches.

Building upon the preceding discussion, we can affirm that our work is related to the proactive off-line deterministic *IORPS*. It can be considered as a specialized case of [187], differing in the way emergency insertion is ensured. More precisely, the hospital under consideration wants to manage all emergencies within a narrow time-window of 20 minutes. As a result, the proactive off-line schedule has to be capable of accommodating the initiation of emergency treatments within this crucial window. Furthermore, our work distinguishes itself from the existing literature through the choice of the objective function. While existing works in the literature primarily concentrate on cost minimization, our approach involves an objective function that is patient-oriented, in line with the obligations of public hospitals.

On this basis, we propose an original *ILP* formulation that extends the one presented in [46], derived from [244], to take into account all these aspects.

In conclusion of this section it is also important to observe that most of the *IORPS* contributions are derived from real-world applications, which are highly specific to each hospital's structure, resources, objectives, and internal regulations. Although there are common features and practices shared among hospitals in different countries, each contribution is designed to address a particular variant of the *IORPS*, making it difficult to apply the proposed methodologies without significant modification. This is also mentioned in [185, 272] where a classification of the works on a geographical basis is provided.

3.3 Problem Definition and Formulation

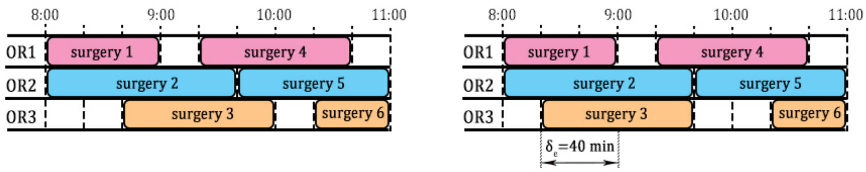
This section is devoted to the presentation of the *ILP* formulation for the *IORPS* problem under investigation. Before going into mathematical details, we summarize the key guidelines and requirements employed by the hospital in the off-line manual definition of the elective patient schedule.

The elective weekly schedule must adhere to:

1. Maximizing the number of surgeries performed.

2. Sequencing elective surgeries prioritizing the most severe and long-lasting ones. In other words, most severe and longest surgeries should be scheduled as early as possible within the planning horizon of a given day.

3. Preserving a specified percentage of each *OR*'s capacity during the planning stage. Thus, a slack time is allocated for each *OR*.



(a) Feasible schedule (b) Unfeasible schedule

Figure 3.1: Toy example of a surgery scheduling, with 6 surgeries, 3 *ORs*, 3 surgical specialties and 9 periods

4. Managing all the potential emergencies within a predefined *WTT*. In other words, ensuring that either an *OR* becomes available within a time equal to or less than the *WTT*, or there is an unoccupied *OR* ready for use at any time.

5. Performing surgeries within predefined work shifts without any overtime.

The initial two points represent guidelines that the hospital adheres to in order to optimize its performance. The other ones pertain to the requirements that must be met in the settings of the *OT*.

For the sake of clarity, we provide additional explanation for the third and fourth requirements. Concerning the third one, we highlight that the baseline schedule relies on estimated duration of surgeries, yet unforeseen complications or events may extend the actual duration beyond initial expectations. Additionally, emergency arrivals may increase the number of surgeries to be performed. To address these issues, the hospital strategically allocates only a percentage of each *OR*'s capacity during the planning stage to minimize the chance of postponements.

Regarding the fourth guideline, we provide an illustrative toy example. Let us consider the schedule of six surgeries over a planning horizon, as shown in Figure 3.1. The schedule in Figure 3.1a adheres to a *WTT* of 20 minutes. Specifically, $surgery$ 1 and $surgery$ 2 begin at $8:00$ in $OR1$ and $OR2$, respectively, while $surgery$ 3 is scheduled at $8:40$. It cannot be scheduled earlier without violating the 20-minute *WTT* if an emergency arises between $8:00$ and $8:40$, as shown in Figure 3.1b.

It is evident that these requirements impacting the elective schedule result in underutilization of the *OR*s. While the practice of keeping *OR*s unoccupied in anticipation of unforeseen elective surgeries may seem unconventional [26], it is an acceptable solution for the hospital. As mentioned above, the emergencies predominantly handled by the hospital are of a highly critical nature, leaving no room for acceptable delays.

Finally, we observe that resources other than the *OR*s (e.g., beds) do not limit the surgical activity of the *OT*.

On this basis, the rest of the section is organized as follows: Section 3.3.1 outlines the problem settings, while Section 3.3.2 provides the mathematical formulation for the block strategy; Section 3.3.3 discusses its adaptation to the modified-block and open strategies; finally, Section 3.3.4 presents two variants of the model. For the sake of generality, the mathematical formulation will be presented without explicitly referring to the specific hospital under investigation.

3.3.1 Problem Setting

Let us consider a generic surgery department, where the set of available operating rooms and surgical specialties are denoted by R and S, respectively. The department operates on a time horizon of one week, divided into days and periods, represented by sets D and P, respectively. Finally, the set of surgeries composing the weekly waiting list is denoted by I.

Based on this setting, we define several parameters. In particular, each surgery $i \in I$ is characterized by five parameters: clinical weight w_i, duration in time periods Δ_i, number of days of hospitalization required before and after the surgery ϵ_i^- and ϵ_i^+, respectively, and the penalty for an unscheduled surgery Ψ_i.

Furthermore, we define a parameter, denoted as B_s, for each specialty $s \in S$ that represents the number of available beds. In addition, two more parameters are defined: δ_e, which specifies the *WTT* value, and $\alpha \in [0, 1]$, which indicates the allowed percentage of *OR* utilization in terms of time slots.

Finally, we use three matrices to express more complex relationships among the defined sets:

1. Ω: this matrix links a surgery $i \in I$ to a surgical specialty $s \in S$. Its generic element σ_{is} is equal to 1 if surgery i belongs to the specialty s, and 0 otherwise.

2. $\mathcal{M}$: this is a four-index matrix linking sets R, S, P, and D, representing the used master surgery schedule. Its generic element M_{rpds} is equal to 1 if the period $p \in P$ on day $d \in D$ of operating room $r \in R$ is assigned to specialty $s \in S$, and 0 otherwise.

3. $\mathcal{C}$: this is a four-index matrix linking all the defined sets and expressing the clinical priority of a surgery as a function of its clinical weight, duration, and starting time. Its generic element C_{ipdr} can be computed as follows:

$$C_{ipdr} = \sum_{s \in S} \sigma_{is} \sum_{k=p}^{p+\Delta_i-1} \frac{w_i^2}{k} M_{rkds} \quad \forall i \in I, p \in P, d \in D, r \in R, s \in S \tag{3.1}$$

By construction, the C_{ipdr} value associated with a generic surgery i (where $i \in I$) decreases throughout the day.

For the sake of readability, the problem setting is summarized in Table 3.1.

Table 3.1: Notation used for the *IORPS* formulation

Problem setting notation	
Sets:	
I	set of surgeries
R	set of *ORs*
S	set of surgical specialties
D	set of days of time horizon
P	set of periods
Parameters:	
w_i	severity of surgery i
Ψ_i	penalty for not scheduling surgery i
Δ_i	total time required by surgery i
δ_e	emergency time threshold
σ_{is}	equal to 1 if surgery i belongs to specialty s, 0 otherwise
ε_i^+	number of days of hospitalization after surgery i
ε_i^-	number of days of hospitalization before surgery i
M_{rpds}	equal to 1 if *OR* r is reserved to specialty s on period p of day d, 0 otherwise
C_{ipdr}	weight for scheduling the surgery i of the specialty s, in the period p, of the day d, of the room r
B_s	beds available for the specialty s
α	$\in [0, 1]$ safety coefficient that represent the wanted percentage utilization of the operation room

3.3.2 *IORPS-ILP* formulation for the block strategy

In accordance with the notation presented above, we introduce a binary decision variable, denoted by x_{ipdr}, which takes a value of 1 if the surgery $i, i \in I$ is scheduled to start in room $r, r \in R$, during period $p, p \in P$ of day $d, d \in D$, and 0 otherwise.

Based on this, the IORPS problem can be formulated using the following *ILP* model, referred to as the *IORPS-ILP*:

$$\mathit{max} \sum_{i \in I} \sum_{p \in P} \sum_{d \in D} \sum_{r \in R} C_{ipdr}\, x_{ipdr} - \sum_{i} (1 - \sum_{p \in P} \sum_{d \in D} \sum_{r \in R} x_{ipdr}) \Psi_i \tag{3.2}$$

$s.t.$

$$\sum_{p \in P} \sum_{d \in D} \sum_{r \in R} x_{ipdr} \leq 1 \qquad \forall i \in I \tag{3.3}$$

$$\sum_{i \in I} \sum_{k=p-\Delta_i+1}^{p} x_{ikdr} \leq 1 \qquad \forall p \in P, d \in D, r \in R \tag{3.4}$$

$$\sum_{i \in I} \sum_{p \in P} \Delta_i x_{ipdr} \leq \lceil \alpha\, |P| \rceil \qquad \forall d \in D, r \in R \tag{3.5}$$

$$\sum_{i \in I | \Delta_i > \delta_e} \sum_{r \in R} \sum_{k=p-\Delta_i+\delta_e | k+\Delta_i-1 \neq p}^{p} x_{ikdr} \leq |R| - 1 \qquad \forall p \in P, d \in D \tag{3.6}$$

$$\sum_{i \in I} \sum_{p \in P} \sum_{r \in R} \sum_{t=\varepsilon_i^-}^{\varepsilon_i^+} x_{ip(d+t)r}\, \sigma_{is} \leq B_s \qquad \forall d \in D, s \in S \tag{3.7}$$

$$x_{ipdr} \sum_{k=p}^{p+\Delta_i-1} M_{rkds} = \Delta_i x_{ipdr} \qquad \forall i \in I, p \in P, d \in D, r \in R, s \in S | \sigma_{is} = 1 \tag{3.8}$$

$$x_{ipdr} \in \{0, 1\} \qquad \forall i \in I, p \in P, d \in D, r \in R \tag{3.9}$$

The objective function (3.2) consists of two terms. The first one prioritizes the scheduling of surgeries during the day based on clinical priority. Indeed, by computing C_{ipdr} in a specific manner, the model attempts to schedule high-priority surgeries as early as possible within the day. The second term aims to maximize the total number of surgeries performed.

Constraints (3.3) ensure that each surgery is scheduled at most once. Constraints (3.4) prohibit overlapping surgeries. Constraints (3.5) require that each operating room has at least $\lceil(1-\alpha)*p\rceil$ empty time slots left. Constraints (3.6) ensure that any occurring emergency can be handled within the predetermined time limit. Constraints (3.7) ensure that a surgery is scheduled only if there are sufficient beds available for hospitalization in the respective surgical department. Finally, constraints (3.8) enforce compliance with the block strategy, ensuring that a surgery is scheduled only in an *OR* period designated for its specialty. Finally, constraints (3.9) specify the nature of the decision variables.

3.3.3 *IORPS-ILP* formulations for Open and Modified-Block strategies

This section describes two adaptations of the *IORPS-ILP* model to the open and modified-block strategies. For the first adaptation, the *IORPS-ILP* can be easily modified to handle the open strategy by setting the matrix $\mathcal{M}$ to a matrix of ones. Then, constraints (3.8) ensure that any scheduled surgery is completed within the work shift.

The second adaptation, however, requires the definition of a new matrix $\overline{\mathcal{M}}$, a reshaping of the matrix $\mathcal{C}$, and the modification of constraints (3.8).

Specifically, let $\overline{\mathcal{M}}$ be defined as the complement of the matrix $\mathcal{M}$. The generic element $\overline{M}_{rpds}$ equals 1 if the room $r \in R$ on day $d \in D$ and period $p \in P$ is not assigned to specialty $s \in S$ in the master surgical schedule, and 0 otherwise.

To reshape the matrix $\mathcal{C}$, we need to introduce an additional parameter: β. Accordingly, the generic element C_{ipdr} is computed as follows:

$$C_{ipdr} = \sum_{s \in S} \sigma_{is} \sum_{k=p}^{p+\Delta_i-1} \left(\frac{w_i^2}{k} M_{rkds} + \frac{\beta}{k} \overline{M}_{rkds}\right) \qquad \forall i \in I, p \in P, \; d \in D, r \in R, s \in S \tag{3.10}$$

where β is a constant parameter representing the surgical weight of a surgery performed in a block not dedicated to its specialty. Its value is strictly lower than $min\, \mathcal{C}$. This choice ensures that we first try to schedule surgeries according to the master surgical schedule and then schedule those that do not comply with it.

Finally, the introduction of $\overline{\mathcal{M}}$ requires modifying constraints 3.8 as follows:

$$x_{ipdr} \sum_{k=p}^{p+\Delta_i-1} (M_{rkds} + \overline{M}_{rkds}) = \Delta_i x_{ipdr} \qquad \begin{matrix} \forall i \in I, p \in P, d \in D, \\ r \in R, s \in S | \sigma_{is} = 1 \end{matrix} \quad (3.11)$$

In this way, surgeries that comply with the master surgical schedule, as well as those that do not, cannot be performed outside the work shift. This means that using overtime is not allowed during planning stage.

3.3.4 Variants of the *IORPS-ILP* formulation

In the proposed model, each room is used only for a percentage of its allowed capacity. However, the free time slots can be spread on convenience overall a day of the planning horizon. In this section, we describe a variant of the *IORPS-ILP* where we impose that the free time slots must be consecutive.

To achieve this, we introduce a dummy surgery j^* whose duration is equal to $\lceil (1 - \alpha) * |P| \rceil$, representing the complement of the room's allowed percentage utilization. Furthermore, we define the set $I^* = I \cup j^*$.

Then, the *IORPS-ILP* can be easily modified by removing constraints (3.5) and introducing the following ones:

$$\sum_{p} x_{j^*pdr} = 1 \qquad \forall d \in D, r \in R \quad (3.12)$$

Constraints (3.12) ensure that a dummy surgery is performed in each room and each day, thereby ensuring that the free time slots are consecutive.

Furthermore, we can enahance this modification to ensure that the free time slots in each available room do not overlap during a day. This can be achieved by including the following constraints:

$$\sum_{r} \sum_{k=p-\Delta_{j^*}+1}^{p} x_{j^*kdr} \leq 1 \qquad \forall d \in D \quad (3.13)$$

This condition enables a turnover of the *ORs*, which can be advantageous in terms of flexibility in responding to emergencies. From a managerial perspective, it is similar to having a dedicated room for managing non-elective surgeries, which provides an additional layer of contingency planning.

3.4 Experimental Results

To validate the mathematical model proposed in Section 3.3, we tested it on a set of instances derived from a real case study of a small hospital in Naples. We are unable to provide detailed information due to a non-disclosure agreement. This section is structured as follows: Section 3.4.1 provides details of the real case study, Section 3.4.2 describes the settings of the instances used, and finally, Section 3.4.3 reports the experimental results.

3.4.1 The Real Case Study of the Hospital in Naples

The hospital being investigated performs surgeries for six surgical specialties: cardiology, general surgery, gynecology, ophthalmology, orthopedics, and senology. The hospital treats both elective and non-elective patients, and the surgeries are performed in a surgery department consisting of four operating rooms managed with a block strategy. The master surgical schedule used by the hospital is shown in Figure 3.2.

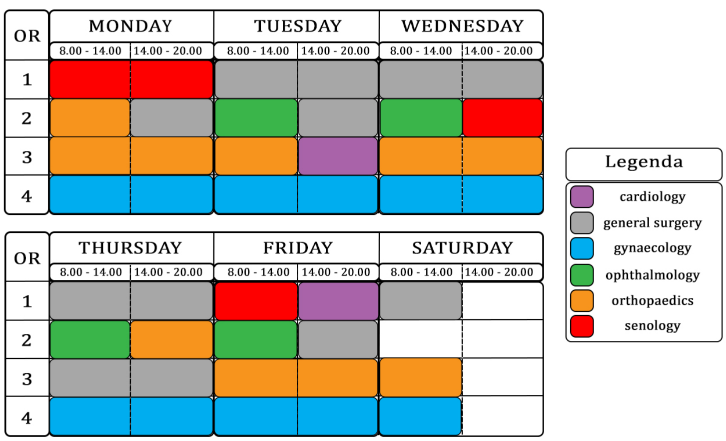


Figure 3.2: Master surgical schedule used for the case study

All the rooms are used from Monday to Friday, from 8:00 to 20:00. On Saturday, Rooms 1, 3, and 4 are also used from 8:00 to 14:00, while room 2 is used only for emergencies.

During the planning stage, no elective surgery can be scheduled in overtime in any *OR*. However, in practice, overtime can be used to manage unexpected events such as emergency surgeries and surgical complications.

The clinical weight assigned to each surgery is defined in terms of two medical parameters: ASA (American Society of Anesthesiologists) and NHS (National Institute for Health and Care Excellence).

The ASA index provides information about the health status of a patient undergoing surgery or standing for it. It uses a classification into 6 levels [12]:

- 1: no organic, biochemical, or psychiatric alteration;
- 2: mild systemic disease;
- 3: severe but not disabling systemic pathology;
- 4: severe systemic disease with a severe prognosis that affects survival regardless of intervention;
- 5: a dying patient who will not survive the next 24 hours, who undergoes surgery as a last chance;
- 6: a patient declared brain dead lying bodies aimed at explant.

Patients with an ASA index between 1 and 3 are classified as elective patients, while values 4 and 5 are used for non-elective cases, requiring immediate decisions.

The NHS index provides information about the complexity of a surgical procedure and uses a classification into 4 degrees [203]:

- 1: minor surgery (such as removal of a skin lesion or drainage of a breast abscess, etc.);
- 2: medium surgery (such as primary repair of an inguinal hernia, removal of varicose veins in the leg, etc.);
- 3: medium-high surgery (such as full hysterectomy, partial removal of the prostate, removal of the thyroid, etc.);
- 4: high-very high surgery (such as total joint replacement, lung operations, removal of part of the liver, etc.).

The hospital defines the priority of each elective patient based on their ASA and NHS values. The surgeons decide the priority based on their experience. According to the hospital's directives, surgeries should be scheduled following the decreasing clinical priority order during the day. Surgeries with the highest clinical priority are scheduled in the first hours of the morning, while surgeries with the lowest clinical priority are scheduled in the latest hours of the evening.

The motivation for this choice is twofold. On the one hand, surgeries performed in the morning are least likely to be postponed during the day and in the following days. On the other hand, high-priority surgeries generally require longer monitoring activities, which can be more easily managed during the regular work shift than during the night one.

3.4.2 Real Instances Setting

The hospital has provided information about surgeries performed over a period of 30 weeks. Specifically, for each surgery, the following information was provided:

1. ASA and NHS indicators;
2. Duration and cleaning time required in minutes;
3. Reference surgical specialty.

Based on this information and interaction with the surgery department manager, we have been able to customize the problem setting defined in Section 3.3:

- The planning horizon has been divided into 36 time slots of 20 minutes each from 8:00 to 20:00. This decision is consistent with the surgery duration, including cleaning activities;
- The value of α has been set to 0.82;
- The weight w_i is computed as the product of the ASA and NHS indexes, which is consistent with the surgeon practice;
- The value of β has been set to 0.25 since the minimum value of an element of $\mathcal{C}$ can be equal to $0.28(1/36)$;
- The value of Ψ_i has been determined by performing a tuning phase to balance the two objective function components;
- The value of δ_e has been set to 20 minutes;
- The matrices $\mathcal{M}$ and $\overline{\mathcal{M}}$ have been set according to the master surgical schedule shown in 3.2;
- The matrix $\mathcal{C}$ has been computed based on the defined setting according to the expressions described in Section 3.3.

- Upon analyzing the hospital data, it was found that the number of available beds for each specialty does not represent a bottleneck for the surgeries to be scheduled. This has been validated by the hospital personnel, and for this reason, these values have been set large enough not to affect the surgery schedule.

Regarding the setting of the value δ_e, i.e., the value of the *WTT*, further discussion is needed. As discussed in [243], medical emergencies can have five levels ranging from level 1 (critical) to level 5 (not urgent). Each level corresponds to a maximum waiting time before treatment: level 1 requires immediate treatment, level 2 requires treatment within 15-20 minutes, and so on. Therefore, the chosen value is consistent with this classification and, based on hospital surgeon experience, ensures an appropriate response time to emergency.

3.4.3 Computational Results of *IORPS-ILP* Formulations

In this section, we present and discuss the computational results of the experimentation performed to validate the proposed *IORPS-ILP*. The experimentation has a twofold objective. Firstly, it aims to demonstrate the usability of the proposed approach with the three previously defined scheduling strategies and the effect of the *WTT*. On this basis, we provide managerial insights that can be used either in the definition of a new master surgical schedule or in the validation of the current one. Secondly, the experiment compares the three scheduling strategies in terms of emergency management, highlighting their benefits and drawbacks.

Concerning the effect of the *WTT* constraints (constraints (3.6)), its evaluation has been performed experiencing the *IORPS-ILP* in three additional variants:

- *IORPS-ILP*$_r$, obtained removing constraints (3.6) from *IORPS-ILP*;
- *IORPS-ILP*^{ms}, modifying the objective function of *IORPS-ILP* in order to maximize the number of scheduled surgeries, i.e., using the follows expression:

$$max \sum_{i \in I} \sum_{p \in P} \sum_{d \in D} \sum_{r \in R} x_{ipdr} \tag{3.14}$$

- *IORPS-ILP*$_r^{ms}$, obtained removing constraints (3.6), as done in *IORPS-ILP*$_r$, and modifying the objective function, as done in *IORPS-ILP*^{ms};

Each of the proposed formulations was solved using Gurobi 9.0.1 as *ILP* solver in its default settings. All instances were run on an Intel i7, 2.60GHz processor with 16.00GB of RAM and a time limit of 3600 seconds. Additionally, these formulations are implemented in the Python language to enable integration into the optimization tool, which is also developed in Python 3.

For completeness, in Figure 3.3, we present a sample solution generated by the proposed approach for a real working week.

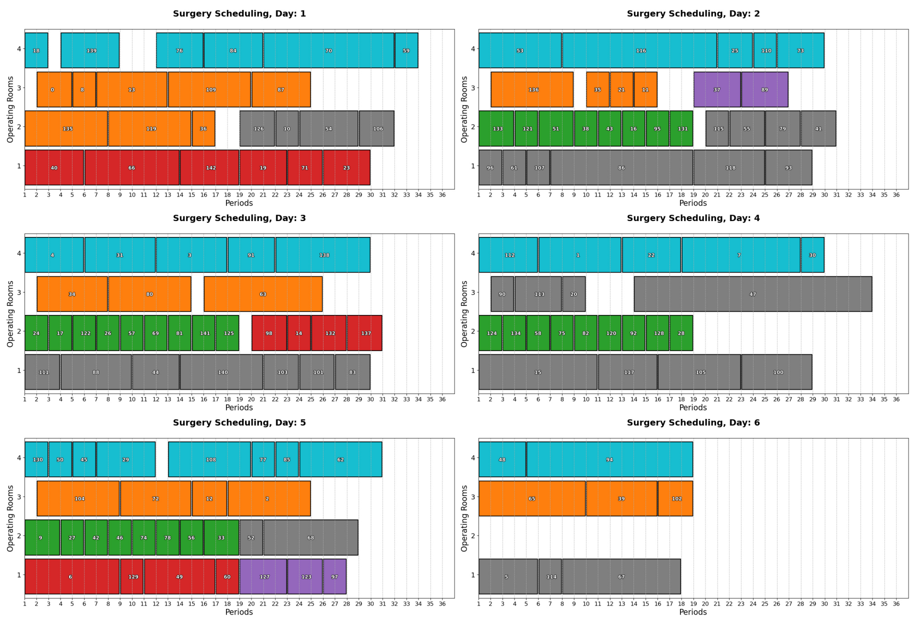


Figure 3.3: Example of the proposed approach output

Comparison among *OR* Management Strategies

Tables 3.2, 3.3, and 3.4 present the results of the experimentation for the block, modified-block, and open scheduling strategies, respectively. Each table displays the following information:

- Instance ID;
- Number of elective patients in the waiting list ($|I|$);
- %SgS: Percentage of scheduled surgeries from the waiting list;

- %OUt: Percentage of *OR* department utilization in terms of periods used for surgeries
- Time: Running time (in seconds)
- %Gap: Optimality gap (%Gap = 1 - lower bound / upper bound)

As a general comment to the results, we highlight that all proposed formulations, including the variants, have been solved to near optimality or optimality within the designated time limit. This confirms the effectiveness of the proposed formulations and their usability in the weekly planning of the hospital.

Table 3.2, which refers to the block strategy, shows that all four variants of the *IORPS-ILP* provide good results in terms of scheduled surgeries, with %SgS values around 94.5%. The effect of surgery priority on the overall number of scheduled surgeries can be observed by comparing the results of *IORPS-ILP* and *IORPS-ILP*^{ms} to those of *IORPS-ILP*$_r$ and *IORPS-ILP*$_r^{ms}$. We observe that neglecting surgical priority leads to a very slight increase in the total number of scheduled surgeries in only a few instances. This confirms that the tuning of the two objective function components has been performed effectively.

A similar discussion can be made for the modified-block strategy, considering the results of Table 3.3. For all four models, the percentage of scheduled surgeries is around 98.5%, on average, and the effect of the two objective functions is negligible.

Finally, we can derive from Table 3.4 that the behavior of the open strategy is similar to the other two strategies. We experience again a high %SgS value, which is around 98.5%, and the effect of surgical priority is still negligible.

To conclude this section, it is important to compare the effect of the strategies on the weekly surgical schedule. As expected, the average percentage of scheduled surgeries and the average operating room utilization of the block strategy are lower than the values obtained by the modified-block and open strategies, confirming the higher flexibility of the latter two. However, from the experimentation performed, we observe that the modified-block strategy performs surgeries related to at least three different specialties in the same room and in each block of the day (morning or afternoon). This situation is even more pronounced in the open strategy, where we observe an average of four different specialties in each room and in each block of the day. As a consequence of such spreading of the specialties in each room and day, higher coordination is required among all involved personnel and surgeons and a more complex management of the workload and shifts from one specialty to another, due to room setup.

On this basis, the results of our approach confirm the effectiveness of the master surgical schedule already used by the hospital. It shows good compromise solutions between different hospital needs and provides solutions that are very close to the modified-block and open strategies in terms of scheduled surgeries and room utilization.

WTT and Emergency Management Evaluation

This section aims to demonstrate the importance of integrating the *WTT* in terms of responsiveness to incoming emergencies.

To achieve this aim, we computed two indicators, denoted as M_p and Σ_p, for the *IORPS-ILP*r and the *IORPS-ILP*r^{ms} variants. M_p corresponds to the maximum number of time periods an emergency has to wait before it can be scheduled, while Σ_p counts the number of time periods in which it is not possible to handle an emergency within the *WTT*. These values are computed by simulating the arrival of an emergency in each period of the time horizon. It should be noted that these values are not evaluated for the *IORPS-ILP* and *IORPS-ILP*^{ms} as constraints (3.6) ensure that these M_p and Σ_p values are always equal to 1 and 0, respectively.

Table 3.4 reports the results of *IORPS-ILP*r and the *IORPS-ILP*r^{ms} in terms of Σ_p and M_p for each instance and strategy.

Based on the observations from Tables 3.2, 3.3, 3.4, and 3.5, we note that, on the one hand, regardless of the strategies used, comparing the results of *IORPS-ILP*r and the *IORPS-ILP*r^{ms} with those of the *IORPS-ILP* and *IORPS-ILP*^{ms}, the first two models return slightly higher %SgS and %OUt. On the other hand, emergencies may not always be handled within the *WTT*. Specifically, for the block strategy, the *IORPS-ILP*r has an average value of 4 and 6 for M_p and Σ_p and a maximum of 7 and 10, respectively. We report a similar situation for the *IORPS-ILP*$_r^{ms}$, with average values of 4 and 5 for M_p and Σ_p, and maximum values of 7 and 9, respectively. For the modified-block and open strategies, we observe a similar but even more pronounced trend. Indeed, for the modified-block strategy, the *IORPS-ILP*r has an average value of 5 and 10 and a maximum value of 7 and 18 of M_p and Σ_p, respectively. The *IORPS-ILP*r^{ms} has similar average values of M_p but with a maximum Σ_p of 20. For the open strategy, the observed values of M_p and Σ_p are on average of 6 and 19 for the *IORPS-ILP*r and 5 and 11 for the *IORPS-ILP*r^{ms}, with maximum values of 9 and 26 for the *IORPS-ILP*$_r$ and 7 and 20, respectively.

We would like to point out that a maximum wait of 9 periods corresponds to almost 2 hours and 30 minutes, and this could lead to severe and unacceptable consequences on the health of patients. In conclusion, we can affirm that the increment in the percentage of surgeries scheduled obtained without considering the *WTT* is negligible compared to the ensured responsiveness to emergencies.

Table 3.2: Results for block strategy

INSTANCE		*IORPS*				$IORPS_r$				$IORPS^{ms}$				$IORPS_r^{ms}$			
ID	**|I|**	**%SgS**	**%OUt**	**%Time**	**Gap**	**%SgS**	**%OUt**	**Time**	**Gap**	**%SgS**	**%OUt**	**%Time**	**Gap**	**%SgS**	**%OUt**	**Time**	**Gap**
INST1	143	0.94	0.71	126	0.00	0.94	0.71	54	0.00	0.94	0.71	28	0.00	0.94	0.71	24	0.00
INST2	139	0.96	0.76	3627	0.73	0.96	0.76	3624	0.05	0.96	0.75	28	0.00	0.96	0.75	24	0.00
INST3	164	0.90	0.76	127	0.00	0.90	0.76	53	0.00	0.90	0.76	28	0.00	0.90	0.76	26	0.00
INST4	125	0.98	0.69	251	0.00	0.98	0.70	41	0.00	0.98	0.68	23	0.00	0.98	0.68	22	0.00
INST5	160	0.94	0.78	1992	0.00	0.94	0.78	70	0.00	0.94	0.76	69	0.00	0.94	0.76	26	0.00
INST6	146	0.94	0.78	3627	0.95	0.94	0.78	50	0.00	0.95	0.76	27	0.00	0.95	0.76	25	0.00
INST7	162	0.91	0.76	3630	0.60	0.91	0.76	77	0.00	0.91	0.75	65	0.00	0.91	0.75	38	0.00
INST8	136	0.92	0.66	910	0.00	0.92	0.66	3625	0.19	0.93	0.65	24	0.00	0.93	0.65	23	0.00
INST9	164	0.85	0.73	3630	1.40	0.85	0.73	3628	0.28	0.85	0.72	27	0.00	0.85	0.72	26	0.00
INST10	155	0.91	0.71	192	0.00	0.92	0.71	47	0.00	0.92	0.71	22	0.00	0.92	0.71	21	0.00
INST11	136	0.97	0.66	877	0.00	0.97	0.66	29	0.00	0.97	0.66	21	0.00	0.97	0.66	20	0.00
INST12	149	0.97	0.78	1798	0.00	0.97	0.78	52	0.00	0.97	0.76	26	0.00	0.97	0.76	23	0.00
INST13	129	1.00	0.68	40	0.00	1.00	0.68	25	0.00	1.00	0.68	21	0.00	1.00	0.68	20	0.00
INST14	133	0.95	0.71	3628	0.66	0.95	0.71	47	0.00	0.95	0.71	23	0.00	0.95	0.71	22	0.00
INST15	137	0.99	0.73	187	0.00	0.99	0.73	39	0.00	0.99	0.73	22	0.00	0.99	0.73	21	0.00
INST16	141	0.94	0.72	406	0.00	0.94	0.72	46	0.00	0.94	0.71	23	0.00	0.94	0.71	22	0.00
INST17	157	0.88	0.75	3629	0.47	0.88	0.75	118	0.00	0.88	0.74	229	0.00	0.88	0.75	344	0.00
INST18	162	0.89	0.78	3630	1.78	0.89	0.78	120	0.00	0.90	0.77	63	0.00	0.90	0.77	41	0.00
INST19	159	0.92	0.77	485	0.00	0.92	0.77	99	0.00	0.92	0.76	32	0.00	0.92	0.76	25	0.00
INST20	137	0.98	0.73	1161	0.00	0.98	0.73	41	0.00	0.98	0.73	24	0.00	0.98	0.73	23	0.00
INST21	110	0.97	0.61	47	0.00	0.97	0.61	39	0.00	0.97	0.60	21	0.00	0.97	0.60	20	0.00
INST22	149	0.93	0.77	3629	0.38	0.93	0.78	3625	0.17	0.93	0.75	43	0.00	0.93	0.76	25	0.00
INST23	128	0.99	0.67	3627	0.83	0.99	0.68	37	0.00	0.99	0.68	22	0.00	0.99	0.67	23	0.00
INST24	130	0.94	0.66	110	0.00	0.94	0.66	48	0.00	0.94	0.65	22	0.00	0.94	0.65	22	0.00
INST25	122	0.99	0.67	295	0.00	0.99	0.67	34	0.00	0.99	0.67	23	0.00	0.99	0.67	22	0.00
INST26	139	0.97	0.72	3628	1.05	0.97	0.72	69	0.00	0.97	0.71	25	0.00	0.97	0.71	23	0.00
INST27	142	0.96	0.72	3627	0.05	0.96	0.72	109	0.00	0.96	0.71	32	0.00	0.96	0.71	28	0.00
INST28	132	0.99	0.69	84	0.00	0.99	0.69	27	0.00	0.99	0.69	24	0.00	0.99	0.69	22	0.00
INST29	143	0.89	0.70	872	0.00	0.89	0.70	61	0.00	0.89	0.69	39	0.00	0.89	0.69	33	0.00
INST30	137	0.97	0.72	207	0.00	0.97	0.72	59	0.00	0.97	0.71	26	0.00	0.97	0.71	24	0.00

Table 3.3: Results for modified-block strategy

INSTANCE		*IORPS*				$IORPS_r$				$IORPS^{ms}$				$IORPS_r^{ms}$			
ID	**|I|**	**%SgS**	**%OUt**	**%Time**	**Gap**	**%SgS**	**%OUt**	**Time**	**Gap**	**%SgS**	**%OUt**	**%Time**	**Gap**	**%SgS**	**%OUt**	**Time**	**Gap**
INST1	143	1.00	0.82	468	0.00	1.00	0.82	107	0.00	1.00	0.82	44	0.00	1.00	0.82	32	0.00
INST2	139	0.99	0.81	3624	1.03	0.99	0.81	579	0.00	0.99	0.81	51	0.00	0.99	0.81	38	0.00
INST3	164	0.95	0.82	3624	0.64	0.96	0.82	594	0.00	0.96	0.81	52	0.00	0.96	0.81	39	0.00
INST4	125	1.00	0.74	417	0.00	1.00	0.74	51	0.00	1.00	0.74	29	0.00	1.00	0.74	33	0.00
INST5	160	0.98	0.82	3625	0.18	0.98	0.82	677	0.00	0.98	0.81	51	0.00	0.98	0.81	44	0.00
INST6	146	0.98	0.81	2476	0.00	0.98	0.81	61	0.00	0.98	0.81	49	0.00	0.98	0.81	38	0.00
INST7	162	0.95	0.82	3641	1.35	0.95	0.82	3636	1.24	0.95	0.81	57	0.00	0.95	0.81	39	0.00
INST8	136	1.00	0.79	4013	0.04	1.00	0.79	270	0.00	1.00	0.79	43	0.00	1.00	0.79	30	0.00
INST9	164	0.93	0.82	3626	0.72	0.93	0.82	615	0.00	0.93	0.81	60	0.00	0.93	0.81	38	0.00
INST10	155	0.97	0.82	3626	0.82	0.97	0.82	706	0.00	0.97	0.82	58	0.00	0.97	0.82	36	0.00
INST11	136	1.00	0.70	279	0.00	1.00	0.70	37	0.00	1.00	0.70	33	0.00	1.00	0.70	28	0.00
INST12	149	0.99	0.82	2345	0.00	0.99	0.82	84	0.00	0.99	0.82	59	0.00	0.99	0.82	40	0.00
INST13	129	1.00	0.68	78	0.00	1.00	0.68	29	0.00	1.00	0.68	31	0.00	1.00	0.68	41	0.00
INST14	133	0.99	0.81	3622	0.36	0.99	0.81	649	0.00	0.99	0.81	58	0.00	0.99	0.81	34	0.00
INST15	137	1.00	0.77	365	0.00	1.00	0.77	37	0.00	1.00	0.77	37	0.00	1.00	0.77	28	0.00
INST16	141	0.99	0.82	984	0.00	0.99	0.82	199	0.00	0.99	0.82	52	0.00	0.99	0.82	35	0.00
INST17	157	0.94	0.82	3624	0.12	0.94	0.82	523	0.00	0.94	0.82	51	0.00	0.94	0.82	38	0.00
INST18	162	0.94	0.82	3626	0.09	0.94	0.82	1480	0.00	0.95	0.82	51	0.00	0.95	0.82	41	0.00
INST19	159	0.97	0.82	3624	0.08	0.97	0.82	485	0.00	0.97	0.82	51	0.00	0.97	0.82	35	0.00
INST20	137	1.00	0.78	797	0.00	1.00	0.78	58	0.00	1.00	0.78	33	0.00	1.00	0.78	29	0.00
INST21	110	1.00	0.66	46	0.00	1.00	0.66	32	0.00	1.00	0.66	25	0.00	1.00	0.66	27	0.00
INST22	149	0.97	0.81	3625	1.45	0.97	0.82	3627	0.49	0.97	0.81	55	0.00	0.97	0.81	34	0.00
INST23	128	1.00	0.68	341	0.00	1.00	0.68	34	0.00	1.00	0.68	29	0.00	1.00	0.68	32	0.00
INST24	130	1.00	0.74	123	0.00	1.00	0.74	52	0.00	1.00	0.74	29	0.00	1.00	0.74	29	0.00
INST25	122	1.00	0.67	277	0.00	1.00	0.67	38	0.00	1.00	0.67	36	0.00	1.00	0.67	26	0.00
INST26	139	1.00	0.78	1306	0.00	1.00	0.78	35	0.00	1.00	0.78	33	0.00	1.00	0.78	30	0.00
INST27	142	1.00	0.77	357	0.00	1.00	0.77	88	0.00	1.00	0.77	37	0.00	1.00	0.77	29	0.00
INST28	132	1.00	0.72	339	0.00	1.00	0.72	31	0.00	1.00	0.72	34	0.00	1.00	0.72	28	0.00
INST29	143	0.99	0.81	3624	0.9	0.99	0.82	620	0.00	0.99	0.81	47	0.00	0.99	0.81	35	0.00
INST30	137	1.00	0.76	52	0.00	1.00	0.76	42	0.00	1.00	0.76	33	0.00	1.00	0.76	29	0.00

Table 3.4: Results for open strategy

INSTANCE		*IORPS*				$IORPS_r$				$IORPS^{ms}$				$IORPS_r^{ms}$			
ID	**|I|**	**%SgS**	**%OUt**	**%Time**	**Gap**	**%SgS**	**%OUt**	**Time**	**Gap**	**%SgS**	**%OUt**	**%Time**	**Gap**	**%SgS**	**%OUt**	**Time**	**Gap**
INST1	143	1.00	0.82	1630	0.00	1.00	0.82	62	0.00	1.00	0.82	43	0.00	1.00	0.82	35	0.00
INST2	139	0.99	0.81	2947	0.00	0.99	0.81	3624	1.74	0.99	0.81	46	0.00	0.99	0.81	41	0.00
INST3	164	0.96	0.82	2304	0.00	0.96	0.82	1041	0.00	0.96	0.81	49	0.00	0.96	0.81	37	0.00
INST4	125	1.00	0.74	112	0.00	1.00	0.74	36	0.00	1.00	0.74	28	0.00	1.00	0.74	31	0.00
INST5	160	0.98	0.82	3628	1.71	0.98	0.82	1347	0.00	0.98	0.81	53	0.00	0.98	0.81	38	0.00
INST6	146	0.98	0.81	3159	0.00	0.98	0.81	1639	0.00	0.98	0.81	49	0.00	0.98	0.81	36	0.00
INST7	162	0.95	0.82	3629	2.63	0.95	0.82	3629	2.35	0.95	0.81	57	0.00	0.95	0.81	37	0.00
INST8	136	1.00	0.79	892	0.00	1.00	0.79	60	0.00	1.00	0.79	45	0.00	1.00	0.79	33	0.00
INST9	164	0.93	0.82	3626	0.13	0.93	0.82	3306	0.00	0.93	0.81	59	0.00	0.93	0.81	38	0.00
INST10	155	0.97	0.82	3628	1.39	0.97	0.82	2181	0.00	0.97	0.82	48	0.00	0.97	0.82	31	0.00
INST11	136	1.00	0.70	91	0.00	1.00	0.70	39	0.00	1.00	0.70	28	0.00	1.00	0.70	32	0.00
INST12	149	0.99	0.82	2750	0.00	0.99	0.82	301	0.00	0.99	0.82	49	0.00	0.99	0.82	35	0.00
INST13	129	1.00	0.68	85	0.00	1.00	0.68	42	0.00	1.00	0.68	29	0.00	1.00	0.68	29	0.00
INST14	133	0.99	0.81	3627	0.15	0.99	0.82	3406	0.00	0.99	0.81	54	0.00	0.99	0.81	32	0.00
INST15	137	1.00	0.77	152	0.00	1.00	0.77	462	0.00	1.00	0.77	31	0.00	1.00	0.77	34	0.00
INST16	141	0.99	0.82	3042	0.00	0.99	0.82	173	0.00	0.99	0.82	49	0.00	0.99	0.82	34	0.00
INST17	157	0.94	0.82	3632	1.65	0.94	0.82	1209	0.00	0.94	0.82	49	0.00	0.94	0.82	39	0.00
INST18	162	0.94	0.82	3635	1.61	0.94	0.82	1238	0.00	0.95	0.82	49	0.00	0.95	0.82	36	0.00
INST19	159	0.97	0.82	3628	0.78	0.97	0.82	1153	0.00	0.97	0.82	49	0.00	0.97	0.82	38	0.00
INST20	137	1.00	0.78	462	0.00	1.00	0.78	62	0.00	1.00	0.78	30	0.00	1.00	0.78	31	0.00
INST21	110	1.00	0.66	331	0.00	1.00	0.66	41	0.00	1.00	0.66	25	0.00	1.00	0.66	30	0.00
INST22	149	0.97	0.82	3636	1.55	0.97	0.82	3627	1.12	0.97	0.81	56	0.00	0.97	0.81	37	0.00
INST23	128	1.00	0.68	118	0.00	1.00	0.68	35	0.00	1.00	0.68	28	0.00	1.00	0.68	29	0.00
INST24	130	1.00	0.74	293	0.00	1.00	0.74	40	0.00	1.00	0.74	28	0.00	1.00	0.74	28	0.00
INST25	122	1.00	0.67	487	0.00	1.00	0.67	41	0.00	1.00	0.67	27	0.00	1.00	0.67	28	0.00
INST26	139	1.00	0.78	279	0.00	1.00	0.78	64	0.00	1.00	0.78	32	0.00	1.00	0.78	33	0.00
INST27	142	1.00	0.77	203	0.00	1.00	0.77	37	0.00	1.00	0.77	30	0.00	1.00	0.77	35	0.00
INST28	132	1.00	0.72	137	0.00	1.00	0.72	38	0.00	1.00	0.72	29	0.00	1.00	0.72	31	0.00
INST29	143	0.99	0.82	3625	0.06	0.99	0.82	1995	0.00	0.99	0.81	43	0.00	0.99	0.81	35	0.00
INST30	137	1.00	0.76	122	0.00	1.00	0.76	53	0.00	1.00	0.76	29	0.00	1.00	0.76	33	0.00

Table 3.5: Emergency management threshold comparison

INSTANCE		*BLOCK*				*MOD-BLOCK*				*OPEN*			
		IORPS-ILP$_r$		**IORPS-ILP**$_r^{ms}$		***IORPS-ILP***$_r$		**IORPS-ILP**$_r^{ms}$		***IORPS-ILP***$_r$		**IORPS-ILP**$_r^{ms}$	
ID	**| I |**	M_p	Σ_p	M_p	Σ_p	M_p	Σ_p	M_p	Σ_p	M_p	Σ_p	M_p	Σ_p
INST1	143	5	4	3	4	3	6	3	2	6	25	3	2
INST2	139	4	7	5	9	6	13	5	11	6	25	5	11
INST3	164	7	8	4	4	5	9	4	7	5	18	4	7
INST4	125	5	10	5	8	4	11	6	18	6	19	6	18
INST5	160	7	7	3	1	7	8	4	6	5	21	4	6
INST6	146	5	10	5	9	6	12	4	9	6	22	4	9
INST7	162	5	10	5	6	6	8	4	4	7	18	4	4
INST8	136	4	2	4	3	4	9	7	19	5	17	7	19
INST9	164	3	2	5	4	4	13	4	7	7	20	4	7
INST10	155	3	3	4	4	5	10	4	5	6	20	4	5
INST11	136	4	4	3	1	4	5	5	9	4	12	5	9
INST12	149	4	5	4	8	5	11	4	8	5	20	4	8
INST13	129	4	5	3	1	3	2	5	14	5	9	5	14
INST14	133	5	8	4	6	6	13	6	8	6	26	6	8
INST15	137	5	6	4	6	4	9	4	12	9	23	4	12
INST16	141	4	7	4	2	7	18	5	12	5	22	5	12
INST17	157	5	7	4	4	5	11	5	8	4	21	5	8
INST18	162	5	10	4	4	4	10	6	9	6	23	6	9
INST19	159	3	5	5	4	5	13	4	9	4	11	4	9
INST20	137	3	7	6	6	4	11	5	20	9	19	5	20
INST21	110	5	3	6	5	5	8	4	13	5	16	4	13
INST22	149	4	7	3	3	4	10	4	7	7	21	4	7
INST23	128	6	8	4	4	6	7	4	10	4	14	4	10
INST24	130	3	2	3	2	4	5	5	13	5	21	5	13
INST25	122	5	8	7	8	6	11	5	13	4	15	5	13
INST26	139	3	2	4	4	5	12	5	15	6	16	5	15
INST27	142	3	3	4	2	4	7	4	13	5	18	4	13
INST28	132	5	8	4	3	4	6	4	10	6	17	4	10
INST29	143	4	4	4	3	4	15	4	8	5	22	4	8
INST30	137	5	5	5	8	5	6	6	18	5	18	6	18

3.5 Conclusions

In this study, we presented an *IORPS* problem arising in a local hospital in Naples. The problem includes specific strict requirements for emergency management and potential surgery delays. We proposed a novel *ILP* modeling framework, which allows the hospital to effectively perform the weekly elective schedule of its patients and simulate potential modifications of the master surgical schedule. The performed experimentation demonstrates the usability of the proposed approach, indicating its potential to replace the manual procedure currently made by hospital personnel. The proposed models constitute also the core of an optimization-based prototypal tool, actually under evaluation by hospital management. Future work perspectives encompass exploring the scalability of the approach to larger hospitals, integrating an on-line rescheduling procedure to handle emergency arrivals, and considering additional elements of the *IORPS* problem, such as staffing resources.

Chapter 4

Re-optimizing *IORPS* in case of unexpected events: an online approach

Addressing the complexities of non-elective surgeries in *ORs* scheduling remains a challenge in healthcare optimization. Unlike elective surgeries, non-elective cases introduce significant unpredictability. This chapter explores an *ILP* approach to real-time *OR* re-scheduling, prioritizing efficiency and responsiveness. Using data from a Naples hospital, we demonstrate the efficacy and practicality of this methodology in improving *OR* scheduling outcomes. We underline that, throughout this chapter, readers may encounter elements and concepts previously introduced in Chapter 3. However, such repetition serves to ensure the self-contained comprehensibility of this chapter.

4.1 Introduction

In recent times, hospitals are facing the challenge of providing better and more efficient services while containing costs within limits. This is especially true in operating rooms (*ORs*), which are considered the financial backbone of modern hospitals as they bring together executives, doctors, nurses, and patients, and also consume a significant portion of the hospital resources, including infrastructure and procurement [252]. In most of the literature, the decisions involved in *OR* planning and scheduling can be broadly classified into three levels: strategic, tactical, and operational. At the strategic level, the problems encompass medium- and long-term demand forecasting, as well as *OR* time allocation to medical specialties/surgeons. This includes addressing capacity planning, capacity allocation, and case mix problems. The decision-making process at this level covers a relatively extended period and involves higher-level decisions. At the tactical level, the primary concern is cyclic *OR* schedules, also known as the master surgical scheduling problem. At the operational level, we recall that, the challenge typically involves the Integrated Operating Room Planning and Scheduling (*IORPS*, which requires determining the date, starting time, and specific resources allocated to each surgical case. The *IORPS* primarily handles elective patients, meaning surgical procedures that can be scheduled in advance. These patients are typically placed on a waiting list and scheduled based on internal rules, policies, and performance criteria, usually one or two weeks before the surgery. In the definition of *IORPS*, great relevance is given to the utilization policies of *ORs*. The literature identifies three primary strategies: open, block, and modified-block. An open strategy allows all surgeries to be scheduled in any *OR* on any day. A block strategy structures the *OR* scheduling into surgical blocks, with each block corresponding to a set of consecutive time periods. Each block is pre-assigned to a specific surgical specialty and day in the planning horizon. The modified-block strategy combines elements of both open and block strategies, allowing under-utilized *OR* blocks to be used by different surgical specialties to maximize *OR* usage [321]. Generally, *IORPS* can encompass non-elective patients in addition to elective ones. The non-elective patients are the ones needing emergency surgical procedures that can not be scheduled in advance. However, it is worth noting that the number of non-elective surgeries is typically lower than the elective. Non-elective surgeries are typically performed in *ORs* that are either dedicated or shared. Dedicated rooms are exclusively set aside for emergency cases, while shared rooms are used for both elective and non-elective surgeries. This approach, as shown in [94], allows for more efficient utilization of operating resources but requires careful room management. Given the unpredictable nature of non-elective patients, the number, arrival times, and duration of emergency surgeries remain uncertain. However, delays in addressing those surgeries can significantly increase patients risk of morbidity and mortality. Furthermore, to further complicate the situation, many emergency surgeries have response time requirements based on their severity. In practice, operating block managers typically design an initial schedule for elective patients at the end of each week for the following one. They select a group of surgeries from a waiting list, taking into consideration various factors such as the length of time a patient has been on the waiting list, the clinical priority of each surgery, and the hospital available resources. Then, the initial schedule must then be quickly and iteratively adjusted in the event of unexpected developments. A poor handling of this task can result in ineffective surgical programs that operate below capacity or that are unable to accommodate emergencies. Nevertheless, a study presented in [181] has shown that a significant portion of patients requiring emergency surgery do not receive timely access to operating rooms, with over 86% of delays being attributed to systemic inefficiencies that could have been prevented. Ineed, despite the best efforts in planning and scheduling, the findings from the study highlight a significant gap in effectively handling emergencies, underscoring the need for improved strategies in the *IORPS* to address the complexities and uncertainties inherent in surgical scheduling, particularly in responding to unforeseen emergencies and managing variable procedure durations.

The *IORPS* has been extensively studied in the literature. The majority of research has focused on the offline problem of creating a weekly surgical schedule, with any unforeseen contingencies being managed by hospital staff on a daily basis. Recently, a new area of interest has emerged i.e., the need to re-plan surgical schedules when emergencies arise that modify the original one. This poses numerous challenges from both modeling and computational perspectives. As mentioned earlier, the primary sources of unpredictability within this field are the unexpected arrival of emergencies and the variability in the duration of surgical procedures. However, it is worth noting that most literature contributions focus on only one of these sources of unpredictability, and typically plan for a single day.

To address the challenges of surgical scheduling, we propose an extension of the approach presented in Chapter 3. This extension consists of a two-phase Integer Linear Program (*ILP*) approach that balances efficiency and responsiveness. In Phase I, we design an initial weekly surgical schedule for elective patients that takes into account unforeseeable events. In Phase II, we re-optimize the schedule in response to exogenous and endogenous variations, such as emergency arrivals or changes in surgery duration. The objective of Phase II is to minimize the deviation of the new schedule from the initial one while maintaining scheduling efficiency and responsiveness. To achieve this objective, we integrate the two phases into an optimization tool. The tool aims to re-schedule contingencies in a way that minimizes the impact on the overall schedule. Specifically, it attempts to re-optimize surgeries on the same day without affecting the schedule of other days. When this is not possible, the tool iteratively adds the following day to the rescheduling planning horizon until the total horizon is reached. In summary, our proposed approach offers a comprehensive solution to the complex problem of surgical scheduling, providing an efficient and responsive framework for managing unforeseen events while minimizing disruption to the overall schedule. The proposed tool has been tested and validated on real instances from a local hospital in Naples as part of a research agreement between the University of Naples Federico II and the hospital. The results show significant improvements in scheduling efficiency by using this methodology as an off-the-shelf tool.

In a nutshell, we can summarize the contribution of this in developed an *ILP*-based optimization tool for surgery scheduling, that aims at:

- Creating efficient weekly schedules for elective patients.
- Able to respond to emergency surgeries promptly.
- Adapting the schedule dynamically to changes and unforeseen events.

The tool operates in two phases:

- Phase I: Utilizes an *ILP* to formulate an initial schedule, emphasizing surgeries with higher priorities and integrating emergency response contingencies into the planning.
- Phase II: Re-optimizes the schedule from Phase I for unexpected events, with the goal of minimal disruptions while maintaining emergency responsiveness.

To our knowledge, this is the first attempt to address the *IORPS* in an iterative and real-time manner, utilizing an exact method that extends the planning horizon beyond a single day and considering the unpredictable nature of both emergency arrivals and surgery duration.

The reminder of the chapter is organized as follows. In Section 4.2 we provide the literature review on deterministic Surgical Scheduling Problem (*IORPS*). Section 4.3 defines the problem under investigation. Section 4.4 is dedicated to presenting the mathematical formulations of the problem as well as the optimization tool that is based on these formulations. Section 4.5 reports and discusses the experimental results. Finally, in Section 4.6 some conclusions and future research directions are outlined.

4.2 Literature review

Healthcare optimization problems have received significant attention in the past few decades [231]. Among the topics covered, the literature concerning the *IORPS* is remarkably broad and investigated by many surveys ([63], [227], [321]).

The problem of *IORPS* is typically approached from two perspectives. The first perspective involves offline planning of surgeries assuming that all necessary information is available at the initial stage, and the literature on this topic is well-established.

On the other hand, the second perspective deals with the online problem of *IORPS*, which complements the first one. In the following, we address offline and online *IORPS* issues in a comprehensive way and examine the literature on both subjects. It should be noted that our primary focus is on the literature concerning *IORPS-on*, which is the main topic of interest in this part. This work builds upon the research presented in Chapter 3, where the literature on *IORPS-off* was thoroughly examined.

The *IORPS-on* belongs to the broader topic of real time management (*RTM*) of the *ORs*. The *RTM* was introduced in [93] where the authors proposed a model, along with an online optimization algorithm to deal with the *IORPS* in a reactive fashion, maximizing *OR* utilization and ensuring that each patient is operated within the time limit set by their urgency level.

Despite the importance of the *IORPS-on*, it is a relatively recent and scarcely investigated topic in the scientific community. It takes into account two types of unpredictable issues: the arrival time of non-elective surgeries the potential variation in the duration of both elective and non-elective surgeries compared to their planned schedules. The majority of existing studies only address the unpredictability of emergency arrival times, often overlooking the potential fluctuations in the actual duration of surgeries.

In particular, [101], [25] and[186] only address the emergencies arrival time unpredictability, while [271] takes into account both the issues of arrival time and surgery duration.

In [101] is proposed a mixed-integer linear programming (*MILP*) model for rescheduling elective surgeries after an emergency arrival. Their formulation considers on the one hand the overtime and the *ORs* costs while, on the other hand, the cost of postponing or anticipating elective surgeries and the cost of turning down the emergency patients. In other words, they can decide to accept or not an emergency (deviate the emergency to another hospital). The objective of their *MILP* model is to minimize the costs associated with postponing and anticipating the elective surgery patients, declining the emergency patients, and overtime of *ORs* and post-anesthesia care units.

In [25] the authors propose a stochastic programming model to determine the optimal policy for inserting emergencies into a predetermined schedule. They assume that the elective surgeries can be delayed but can not be moved to other *ORs*, and that the arrival pattern of emergencies follows a non-homogeneous Poisson process. The objective is to minimize the expected cost of responsiveness requirements for emergencies, delays in elective surgeries, and surgery team overtime. The results show that the responsiveness of the elective schedule is worse when considering the "shortest expected processing time first" rule compared to the "best insertion method" rule.

In [186] the authors addressed the *IORPS* by proposing a mixed-integer programming (*MIP*) model to design an initial schedule that takes into account possible emergency arrivals. The goal of this stage is to minimizing number of non-scheduled elective surgeries. They also develop two heuristics based on the proposed model. In the second stage, they present a reactive method to reschedule surgeries on the surgery day, with the goal of minimizing its impact on the initial schedule for elective surgeries. However, they do not consider the possibility of an elective surgery taking longer than expected as a condition to adjust the schedule. Indeed, they assumed that elective surgeries duration is deterministic.

In [271] the authors address the *IORPS* of elective and non-elective surgeries in a shared operating room policy. They consider uncertainty in both the arrival time and duration of surgeries and use a stochastic process to determine the probability of arrival, expected duration, and response requirements for each surgery. They solve the problem by dividing the workday into time periods and designing an approximate dynamic programming algorithm to dynamically reschedule elective cases on a single day planning horizon.

Building upon the existing literature and to the best of our knowledge, this is the first work to tackle the *IORPS*: in a iterative and real-time manner; employing an exact method that extends the planning horizon beyond a single day; takes into account the unpredictable nature of both emergency arrivals and surgery duration.

4.3 Problem definition

In this section, we discuss the specific challenges faced by a local hospital in Naples. The hospital addresses this problem by splitting it into two correlated subproblems. These two subproblems fall within the aforementioned classifications of *IORPS-off* and *IORPS-on*. For the sake of clarity, we report the definition of these two problems in 4.3.1 and 4.3.2, respectively.

4.3.1 Offline *IORPS*

In the *IORPS-off* approach, hospital staff generates a weekly surgical schedule for the upcoming week, assigning a specific day, *OR*, and start time to each elective surgery using a block strategy. In more detail, a subset of patients is selected from a waiting list based on their clinical priority. Each patient is then assigned to an *OR* and a specific surgical session (e.g., day, morning/afternoon). Concurrently, surgeries are sequenced within their designated sessions. According to literature, the hospital simultaneously addresses the *ADP* and *ALP* during this process. Each elective surgery is defined by four elements: clinical specialty, duration, clinical priority, and required days of hospitalization. The clinical specialty determines the necessary medical equipment and surgeon skills, while the surgery duration is estimated based on historical data and surgeon expertise. Clinical priority serves as a proxy for surgery complexity and patient health status. While, the required days of hospitalization define the length of stay that each surgery needs before and after being performed. It is important to note that bed availability may be a bottleneck, as it may not be possible to perform surgery if there are no beds available in the respective department.

The initial schedule is made using the planned duration of the surgeries. However, surgeries could last longer than expected due to operative complications and unforeseeable events, and emergency arrivals may increase the number of surgeries that need to be performed. To address these issues, only a percentage of the capacity of each *OR* is used during the planning stage, which avoids surgery postponements that can be perceived by patients and competent authorities as poor quality of service.

Additionally, the schedule is prepared in a way that allows for the management of incoming emergencies within their responsiveness requirement. This may result in underutilization of the *ORs*, but it ensures the ability to manage emergencies within an adequate time, thereby increasing the probability of successful emergency treatments. To such aim, in each *OR*, a surgery can only start if another *OR* is free or if a surgery ends within the minimum responsiveness requirement. The minimum responsiveness requirement has been widely discussed in Chapter 3, and interested readers can refer to that Chapter for more information.

The principal objective of the hospital is to find a good trade-off solution that considers the maximization of surgeries performed and the ability of the schedule to respond to emergencies.

4.3.2 Online *IORPS*

In addition to the initial schedule defined in 4.3.1, the hospital must take into account two types of contingencies. The first is the arrival of non-elective patients, whose number and arrival time, as well as the time required for their surgeries, cannot be predicted. To accommodate these cases, the hospital has implemented a shared operating room policy where both elective and non-elective surgeries are performed in the same *ORs* [94]. This means that non-elective patients may compete with elective surgeries for *OR* capacity during operating hours.

The second contingency is associated with variations in surgery duration due to complications or unforeseeable events during the procedure. In this context, surgery duration is a dynamic variable. Prior to the start of the surgery, an average duration is assumed based on the criteria used in the initial schedule. However, as soon as the surgery begins, this value is updated with information from the *ORs* and converges to the actual duration of the surgery over time.

Therefore, the initial schedule must be flexible enough to accommodate the arrival of emergency patients within their responsiveness requirements, absorb the variation in surgery duration, and be adjusted throughout the planning period to ensure adequate responsiveness to emergencies.

The hospital objective is to maintain the schedule as close as possible to the initial plan within the planning horizon. This is because changing the operating day for a patient can lead to organizational problems for the hospital staff and occupy a bed for an extended period, which can prevent the hospital from performing other surgeries. Furthermore, a session of a certain specialty may not be available two consecutive days, which could lead to the postponement of a surgery for several days and result in a low quality of service perceived by patients.

4.4 Mathematical models for *IORPS*

This section presents the mathematical model and problem settings of the *IORPS* under investigation. Specifically, Sections 4.4.1 and 4.4.2 introduce the problem settings and mathematical formulations for the *IORPS-off* and the

IORPS-on, respectively. The formulation for the *IORPS-off* is referred to as *ILP-IS*, and the formulation for the *IORPS-on* is referred to as *ILP-EM*. Finally, in Section 4.4.3, we introduce and discuss the optimization tool developed for the problem being investigated.

4.4.1 Formulation for the offline *IORPS*

Let us consider a generic surgery department, where the set of available operating rooms is denoted as R and the set of surgical specialties is denoted as S. The department operates over a one-week time horizon, divided into days and periods, which are denoted by sets D and P, respectively. The set of surgeries on the weekly waiting list is denoted as I. Based on this setting, are defined several parameters. In particular, each surgery $i \in I$ is characterized by five parameters: the clinical weight w_i, the duration in time periods Δ_i, the number of days of hospitalization required before and after the surgery, ϵ_i^- and ϵ_i^+, respectively, and the penalty for not scheduling it Ψ_i. Additionally, is introduced a parameter B_s representing the number of available beds for each specialty $s \in S$.

Concerning the emergency management issue, two additional parameters needs to be defined: δ_e, which sets the EMT value, and $\alpha \in [0, 1]$, which represents the allowed utilization percentage of *ORs* in terms of time slots.

Moreover, we need to introduce three matrices to express more complex relationships among the defined sets. Specifically:

1. Ω: links a surgery $i \in I$ to a surgical specialty $s \in S$. Its generic element σ_{is} is equal to 1 if surgery i belongs to specialty s, and 0 otherwise.

2. $\mathcal{M}$: is a four-index matrix linking sets R, S, P and D, and represents the master surgery schedule. Its generic element (M_{rpds}) is equal to 1 if the period $p \in P$ on day $d \in D$ in *OR* $r \in R$ is assigned to specialty $s \in S$, and 0 otherwise.

3. $\mathcal{C}$: is a four-index matrix linking all defined sets, and expresses the clinical priority of a surgery as a function of its clinical weight, duration, and starting time. Its generic element C_{ipdr} can be calculated as:

$$C_{ipdr} = \sum_{s \in S} \sigma_{is} \sum_{k=p}^{p+\Delta_i - 1} \frac{w_i^2}{k} M_{rkds} \qquad \forall i \in I, p \in P, d \in D, \; r \in R, s \in S \tag{4.1}$$

By construction, the C_{ipdr} value for a generic surgery $i \in I$ decreases over a day.
On this basis, we define a binary decision variable x_{ipdr} which assumes value 1 if the starting of surgery $i, i \in I$, is scheduled in room $r, r \in R$, in period $p, p \in P$, of day $d, d \in D$, 0 otherwise.

Consistently with above introduced notation, the *ILP-IS* can be formulated as follows:

$$max \sum_{i \in I} \sum_{p \in P} \sum_{d \in D} \sum_{r \in R} C_{ipdr}\, x_{ipdr} - \sum_{i} (1 - \sum_{p \in P} \sum_{d \in D} \sum_{r \in R} x_{ipdr}) \Psi_i \tag{4.2}$$

$s.t.$

$$\sum_{p \in P} \sum_{d \in D} \sum_{r \in R} x_{ipdr} \leq 1 \qquad \forall i \in I \tag{4.3}$$

$$\sum_{i \in I} \sum_{k=p-\Delta_i+1}^{p} x_{ikdr} \leq 1 \qquad \forall p \in P, d \in D, r \in R \tag{4.4}$$

$$\sum_{i \in I} \sum_{p \in P} \Delta_i x_{ipdr} \leq \lceil \alpha\, |P| \rceil \qquad \forall d \in D, r \in R \tag{4.5}$$

$$\sum_{i \in I | \Delta_i > \delta_e} \sum_{r \in R} \sum_{k=p-\Delta_i+\delta_e | k+\Delta_i-1 \neq p}^{p} x_{ikdr} \leq |R| - 1 \qquad \forall p \in P, d \in D \tag{4.6}$$

$$\sum_{i \in I} \sum_{p \in P} \sum_{r \in R} \sum_{t=\varepsilon_i^-}^{\varepsilon_i^+} x_{ip(d+t)r}\, \sigma_{is} \leq B_s \qquad \forall d \in D, s \in S \tag{4.7}$$

$$x_{ipdr} \sum_{k=p}^{p+\Delta_i-1} M_{rkds} = \Delta_i x_{ipdr} \qquad \begin{array}{r} \forall i \in I, p \in P, d \in D, \\ r \in R, s \in S | \sigma_{is} = 1 \end{array} \tag{4.8}$$

$$x_{ipdr} \in \{0, 1\} \qquad \begin{array}{r} \forall i \in I, p \in P, \\ d \in D, r \in R \end{array} \tag{4.9}$$

The objective function (4.2) consist of two components. The first part focuses on scheduling surgeries based on their clinical priority. This is achieved by trying to schedule high priority surgeries as early as possible in the day, as reflected in the computation of C_{ipdr}. The second part aims to maximize the overall number of surgeries performed.

Shifting focus to the logical and operative conditions that need to be observed. Constraints (4.3) ensure that each surgery is scheduled only once. Constraints (4.4) prevent surgery overlapping. Constraints (4.5) require that at least $\lceil (1 - \alpha) * p \rceil$ time slots remain empty in each *OR*. It is important to note that, given the definition of the $\mathcal{C}$ matrix and the objective function are defined, the empty slots are concentrated towards the end of the day. This arrangement enables the schedule to accommodate emergencies and variations in surgical procedures, aiming to avoid overtime work. Constraints (4.6) guarantee that emergency situations can be handled within the specified time threshold. Constraints (4.7) ensure that a surgery can only be scheduled if there is sufficient bed availability in the relevant surgical department. Constraints (4.8) enforce the block strategy, meaning that a surgery can only be scheduled during an *OR* period dedicated to its specialty.

Finally, constraints (4.9) define the nature of the decision variables.

4.4.2 Formulation for the online *IORPS*

Building on the configurations introduced in section 4.4.2, we must adjust certain sets and parameters while introducing new ones to represent the *ILP-EM* model. Specifically, we need to divide the set of surgeries on the weekly waiting list (I) into four distinct sets: I^{kl}, F^{kl}, O^{kl}, and E^{kl}. These sets represent the surgeries to be scheduled, completed, ongoing, and emergencies, for period $k \in P$ of day $l \in D$, respectively. Moreover, it is necessary to update the parameter Δ_i into Δ_i^{kl}, which represent the estimated duration of surgery i at period $k \in P$ of day $l \in D$, based on the information available online. We also need to introduce three new parameters related to the initial schedule: p_i^*, d_i^*, and r_i^*, which represent the starting period, day, and *OR* for each surgery i in the initial schedule, respectively. Additionally, are required three parameters related to the last available re-schedule: p_i^{-1}, d_i^{-1}, and r_i^{-1}, representing the starting period, day, and *OR* for each surgery i in the last available re-schedule, respectively. Furthermore, two parameters related to emergencies, p_i^e and d_i^e, are needed, representing the arrival period and the arrival day for each emergency i.

Finally we need to define a new four index matrix $\mathcal{P}$, defined as follows:

$$\overline{\mathcal{C}} = \begin{cases} \overline{C}_{ipdr} = 0 & if\ p = p^* \wedge d = d^* \\ \overline{C}_{ipdr} = p - p^* & if\ p > p^* \wedge d = d^* \\ \overline{C}_{ipdr} = |P|(d - d^*) & if\ d > d^* \end{cases}$$

The generic element of this matrix is denoted as $\overline{C}_{ipdr}$. This element is used to compute the deviation of the revised schedule from the original one. If the two schedules are identical, the weight is assigned a value of zero. Moreover, moving a surgery from one day to another incurs a higher weight compared to shifting it within the same day. This is due to the fact that extending a patient stay in the hospital by one or more additional days can lead to logistic challenges for both the patient and the hospital.

Consistently with the additional notation introduced, as well as the one presented in the preceding section, the *IORPS-EM* can be formulated as follows:

$$max \sum_{i\in I} \sum_{p=k}^{P} \sum_{d=l}^{l+z} \sum_{r\in R} |\overline{C}_{ipdr}|\ x_{ipdr} - \sum_{i} (1 - \sum_{p\in P} \sum_{d\in D} \sum_{r\in R} x_{ipdr}) \Psi_i \quad (4.10)$$

$s.t.$

$$x_{ip_i^{-1} d_i^{-1} r_i^{-1}} = 1 \qquad \forall i \in F^{kl} \cup O^{kl},\ i \in I^{kl} : d_i^{-1} \notin l, \dots, l+z \quad (4.11)$$

$$\sum_{p\in P} \sum_{d\in D} \sum_{r\in R} x_{ipdr} \leq 1 \qquad \forall i \in F^{kl} \cup O^{kl} \cup I^{kl} \quad (4.12)$$

$$\sum_{i\in F^{kl}\cup O^{kl}\cup I^{kl}} \sum_{j=p-\Delta_i^{kl}+1}^{p} x_{ijdr} \leq 1 \qquad \forall p \in P, d \in D, r \in R \quad (4.13)$$

$$\sum_{i\in O^{kl}\cup I^{kl} | \Delta_i^{kl} > \delta_e} \sum_{r\in R} \sum_{j=p-\Delta_i^{kl}+\delta_e | j+\Delta_i^{kl}-1\neq p}^{p} x_{ijdr} \leq |R| - 1 \qquad \forall p \in P,\ d \in D \quad (4.14)$$

$$\sum_{i\in F^{kl}\cup O^{kl}\cup I^{kl}} \sum_{p\in P} \sum_{r\in R} \sum_{t=\varepsilon_i^-}^{\varepsilon_i^+} x_{ip(d+t)r}\ \sigma_{is} \leq B_s \qquad \forall d \in D, s \in S \quad (4.15)$$

$$x_{ipdr} \sum_{j=p}^{p+\Delta_i^{1l}-1} M_{rjds} = \Delta_i^{kl} x_{ipdr} \qquad \begin{array}{r} \forall i \in I, p \in P, d \in D, \\ r \in R, s \in S | \sigma_{is} = 1 \\ -E^{11} \cup \cdots \cup E^{kl} \end{array} \tag{4.16}$$

$$\sum_{p=p_i^e}^{k+d_i^e} \sum_{r \in R} x_{ipd_i^e r} = 1 \qquad \forall i \in E^{(k-d_i^e)l} \cup \cdots \cup E^{kl} \tag{4.17}$$

$$\sum_{p \in P} \sum_{r \in R} \sum_{j=1}^{d_i^*-1} x_{ipjr} = 0 \qquad \forall i \in I^{kl} \tag{4.18}$$

$$x_{ipdr} \in \{0,1\} \qquad \begin{array}{r} \forall i \in I, p \in P, \\ d \in D, r \in R \end{array} \tag{4.19}$$

It is important to note that this new model is a variation of the *IORPS-IS* that takes into account the new dynamics and contingencies of the system. For the sake of simplicity, we will concentrate on the constraints that either introduce new aspects or modify their existing meaning.

The objective function (4.10) consists of two components. The first part aims to minimize the deviation from the initial schedule, consistently with the matrix $\overline{\mathcal{C}}$. We underline that the generic element of the matrix, namely $\overline{C}_{ipdr}$, is taken in its absolute value to prevent any deviation from the initial schedule, regardless of it being positive or negative.

Constraints (4.11) ensure that surgeries, once started or completed, maintain their positions in the updated schedule. Constraints (4.12, 4.13, 4.14, 4.15) keep their original purpose from the *IORPS-IS* model, but have been modified to align with the additional notation introduced in section 4.4.2. Constraints (4.16) represent a modified version of constraints (4.8), accounting for the possibility that an initiated surgery, whose estimated duration aligns with the block strategy, can be completed even if it extends beyond its allocated time block. Emergencies are excluded from these constraints, as they can be accommodated in any operating room. Constraints (4.17) ensure that every emergency is scheduled within its maximum allowable waiting time from the arrival moment. Constraints (4.18) allow surgeries to be delayed but prevent them from being performed earlier than their initially scheduled time. Meanwhile, constraints (4.19) specify the characteristics of the decision variables.

4.4.3 The optimization tool

In this section, we describe an *ILP*-based optimization tool developed for a case study of a hospital in Naples. The tool is designed to achieve the following goals:

1. Assist medical staff in creating an effective and efficient weekly schedule for elective patients.

2. Ensure responsiveness to emergency surgeries.

3. Incorporate emergency surgeries into the schedule within the required response time.

4. Automatically adjust the schedule when emergencies or variability in surgery duration make it ineffective.

To this aim the optimization tool consists of two phases. In Phase I, the tool inputs the elective surgeries planned for the upcoming week and creates an initial schedule using the mathematical model *ILP-IS*. The average surgery duration is used as input, based on historical data and individual surgeons' experiences, since real-time information is not yet available. This phase has two objectives: to maximize the number of surgeries performed while considering their medical priority and to create a resilient schedule that can accommodate fluctuations in external conditions.

In Phase II, the tool inputs the initial schedule from Phase I and assesses if there are any changes, such as emergency arrivals or updated surgery duration from the operating room. If any changes occur, the schedule is re-optimized using the mathematical model *ILP-EM*. The objective of Phase II is twofold: to minimize deviations from the initial schedule without cancelling surgeries and to maintain responsiveness throughout the planning horizon. Maintaining the initial schedule minimizes last-minute changes and surgery postponements, while ensuring responsiveness to emergencies is crucial in saving patients' lives.

It is important to note that while the responsiveness requirement can be guaranteed during the initial planning (Phase I), this may not be possible in Phase II due to external factors. For example, if an emergency arrives and no operating room is available, the emergency can be scheduled when another surgery ends. However, if an emergency arrives at the same time as a surgery experiencing complications, it may not be possible to guarantee responsiveness. In such cases, the tool schedules the emergency in the first available slot, sacrificing responsiveness in that period to ensure feasibility of the solution.

We highlight that, it is possible to account for these edge cases in the initial schedule, but this would result in a conservative plan with many unused surgical slots and longer waiting times, as events with low probability are considered.

In summary, the optimization tool provides support in creating an operating plan as follows: Phase I creates an initial preliminary schedule to maximize the number of surgeries and consider potential emergencies. In Phase II, the schedule is re-optimized if there are changes that affect responsiveness, with the goal of minimizing deviations from the initial schedule. The re-optimization takes place incrementally, starting with modifying the schedule for one day, then two days, and so on, until the last day. Any surgeries that cannot be performed at the end of the last day be carried over to the following week.

4.5 Computational results

To validate the mathematical model presented in Section 4.4, we conducted experimentation on a set of instances based on a real-world case study of a small hospital in Naples. Due to a non-disclosure agreement, the specific details of the case study are not exposed. This section is organized into three parts: Section 4.5.1 provides an overview of the real-world case; Section 4.5.2 outlines the configuration of the instances used; finally, Sections 4.5.3 and 4.5.4 presents the experimental results. All the proposed formulations were solved using the developed prototype tool, which employs Gurobi 9.0.1 with its default settings as an *ILP* solver. The instances were executed on an Intel i7 processor with a speed of 2.60GHz and 16.00GB of RAM. The time limit for the *IORPS-off* was set to 3600 seconds, while there was no time limit for the *IORPS-on*.

4.5.1 The case study

The hospital under investigation provides surgical services for six specialties: cardiology, general surgery, gynecology, ophthalmology, orthopedics, and senology. Their surgical department includes four *ORs*, managed using a block strategy to accommodate both elective and non-elective patients. It is crucial to note that the hospital operates under a shared rooms policy. We recall that this is indicating that emergency patients may compete with elective surgeries for *OR* capacity during regular hours.

The *ORs* are open from 8:00 am to 8:00 pm, Monday through Friday. On Saturdays, rooms *1*, *3*, and *4* are open from 8:00 am to 2:00 pm, while room *2* is reserved for emergency cases only.

During the planning stage, elective surgeries cannot be scheduled outside regular operating hours. However, in exceptional situations like emergency surgeries or complications during surgery, overtime may be utilized.

Non-elective surgeries, due to their unpredictable nature, cannot be scheduled in advance. However, if a non-elective surgery arises, it must be performed without delay due to the urgent medical condition of the patient.

To each surgery is assigned a clinical weight based on two medical parameters: *ASA* (American Society of Anesthesiologists) and *NHS* (National Institute for Health and Care Excellence). The *ASA* index reflects the patient health status, with a classification into six levels ranging from no organic, biochemical, or psychiatric alterations to patients declared brain dead. Elective patients are those with an *ASA* index between 1 and 3, while non-elective patients have an index of 4 or 5, requiring immediate attention. According to the defined levels, surgeries with *ASA* indicator ranging from 1 to 3 are considered in the weekend planning for the following week. Meanwhile, patients with *ASA* indicator of level 4 or 5 are considered as emergencies and are promptly addressed as soon as they arise, with maximum attention given to their timely and efficient management.

The *NHS* index, on the other hand, classifies the complexity of surgical procedures into four degrees, ranging from minor to high-very high. This classification system is used to categorize surgeries based on their complexity. Surgeons determine the priority of each elective patient based on their *ASA* and *NHS* values, as well as their experience. Based on this classification, surgeries are scheduled based on decreasing clinical priority during the day, with surgeries of the highest priority scheduled in the morning and those of lower priority in the evening.

This approach has two key benefits. Firstly, surgeries performed in the morning are less likely to be postponed or delayed in the following days. Secondly, high-priority surgeries require longer monitoring activities that are more efficiently managed during regular work shifts.

4.5.2 Test bed

The hospital provided information on surgeries performed over a period of 30 weeks. For each surgery, they provided details on the *ASA* and *NHS* indicators, duration, cleaning time required (expressed in minutes), and reference surgical specialty.

Based on this information and discussions with the surgery department managers, we have established the problem setting for the *IORPS-off* as outlined in 4.5.2, and the corresponding setting for the *IORPS-on* as described in 4.5.2.

Instance setting for the offline *IORPS*

The planning horizon has been divided into 36 time slots of 20 minutes each, from 8:00 to 20:00, which is consistent with surgery duration, cleaning activities and possible emergency arrivals. To assess the practical effectiveness of the proposed schedule in *IORPS-off*, we have set the value of α to different values, specifically $0.8, 0.9$, and 1.0. The objective is to determine the optimal percentage of room utilization that achieves the best performance, taking into account emergencies and complications during surgeries. By evaluating the trade-off between the utilization of operating rooms and the number of postponed surgeries at the end of a working week, we can determine the ideal room utilization rate. This approach allows us to achieve a balance between optimizing the use of *ORs* and reducing the number of surgeries that get postponed due to unexpected events. Overall, these tests enable us to make informed decisions about how to plan (*IORPS-off*) and then adjust the proposed schedule (*IORPS-on*), leading to better outcomes for both patients and staff.

Concerning the parameter definition in more details, the weight w_i is computed as the product of the *ASA* and *NHS* indexes, which is in line with surgeon practice. We have set the value of β to 0.25 since the minimum value of an element of $\mathcal{C}$ can be equal to $0.28(1/36)$. The value of Ψ_i has been determined by performing a tuning phase aimed at balancing the two objective function components. We have set the value of δ_e to 20 minutes. The matrix $\mathcal{M}$ have been defined consistently with the master surgical schedule used in practice. The matrix $\mathcal{C}$ has been computed based on the settings defined above, according to the expressions described in 4.4.

Upon analyzing the data provided by the hospital, it was determined that the number of available beds for each specialty is not a bottleneck for scheduling surgeries. This assertion has been validated by hospital personnel, and therefore, we have set these values large enough to not affect the surgery schedule. However, we have decided to incorporate this aspect into the mathematical model, as it is a part of an optimization tool intended for clinical use. By incorporating this aspect into the model, we ensure that, should conditions change and it becomes a bottleneck in the future, we are well-prepared to intervene and minimize its effect on the scheduling process.

Regarding the setting of the value δ_e, which is the responsiveness requirement for each emergency, further discussion is necessary. As discussed in [243], medical emergencies can be classified into five levels, from level 1 (critical) to level 5 (not urgent), with each level corresponding to a maximum waiting time before treatment. For instance, level 1 requires immediate treatment, and level 2 requires treatment within 15-20 minutes. Therefore, the chosen value of δ_e is consistent with this classification and, based on hospital surgeon experience, it ensures an appropriate waiting time to manage an emergency. This is particularly true for the hospital under investigation, where most emergencies are gynecological in nature (mainly of level 2).

Instance setting for the online *IORPS*

To test and evaluate the effectiveness of the optimization tool, we needed to create models for both the arrival of emergencies and the potential variation in surgeries duration, which often necessitates modifying the initial schedule. To accomplish this, we utilized historical data provided by the hospital, which indicated that an average of 10 emergencies are managed each week and roughly 10% of surgeries experience a positive duration variation.

On this basis, we conducted a simulation by generating emergencies and varying the duration of surgeries. Our aim was to ensure that the simulation closely resembled the operational reality of the hospital. Thus, we iteratively generated emergencies and surgery duration variations according to the data provided by the hospital. Specifically, for each period and day within the time horizon ($\forall p \in P, d \in D$), there was a possibility of an emergency occurring or a change in the duration of a surgical procedure, both of which needed to be accommodated in the schedule.

In particular, to simulate emergencies, we generated them with a probability of $p = \frac{10}{198}$ for each subsequent period after the initial one. The duration of the emergency was determined by sampling from a Gaussian distribution with mean 5 periods ($\mu = 5$) and variance equal to 2 periods ($\sigma^2 = 2$). The resulting value was then rounded to the nearest integer. This duration is typically estimated by the surgical team based on their experience and expertise with the type of emergency at hand.

Concerning the variation in the duration of surgeries, we observed that negative variations (where surgeries take less time than expected) are significantly less frequent than positive ones (where surgeries take more time than expected). Additionally, negative variations typically do not necessitate changes to the initial schedule and can be managed by simply waiting, and allowing the schedule to progress normally. As a result, negative variations were not taken into account in this study.

For each period following the first, each ongoing surgeries has a probability of varying their duration of $P_r(i) = \frac{0.1*\delta_i}{|R||P|}$, where the potential increase in estimated duration could be $1, 2$, or 3, with respective probabilities of $0.5, 0.3$, and 0.2. This method was devised in collaboration with the hospital medical staff to simulate various levels of complications that might occur during surgeries. The probabilities were set to decrease with the increasing severity of the complications. More in details, if a surgery is selected to have a variation in its duration, then based on the proposed rule, an increment, which we can refer to as var, is also selected. Consequently, the new duration of surgery $i, i \in I$, in period $p, p \in P$ of day $d, d \in D$ is calculated as $\Delta_i^{(p-1)d}$=$\Delta_i^{pd} + var$.

We want to highlight two key points. Firstly, the estimated duration of a surgeries can change at different points in time, as new complications may arise during the procedure. Secondly, emergencies are not excluded to potential duration variations.

4.5.3 Results and discussion

In the Table 4.1, 4.2, 4.3 are reported the experimentation made for 30 weeks provided by the Surgical Department of an hospital in Naples. In particular, the table reports the following information. Column 1, labeled "ID", displays the instance ID associated with both the *IORPS-off* and *IORPS-on* models. In column 2, "|I|" represents the total number of surgeries for each given instance.

In addition, addressing the specifics of the *IORPS-off* model, column 3, referred as "%SgS", reports the percentage of surgeries scheduled for every instance. Column 4, named "%OUt", represents the average utilization of the *ORs* over the week. Column 5, labeled "%Time", showcases the computational time (in seconds). Column 6, denotes as "Gap", reports the percentage optimality gap.

Meanwhile, the *IORPS-on* is addressed in columns ranging from 7 to 11. In particular, column 7, titled "#emg", reports the number of emergency simulated for each instance. In column 8, labeled "#s-var", the focus shifts to surgeries, reporting the number of those which deviated from their initially estimated durations. Column 9, referred as "%OUt$_{last}$", reports the concluding utilization rates of the *ORs* (i.e. at the last period of the last day of the week). Column 10, named "#Sg$_{post}$", highlights the number on surgeries that underwent rescheduling, reporting the number of those shifted to alternate days. Finally, column 11, denoted as "#Sg$_{n}s$", emphasizes surgeries that were not accommodated within the current schedule and were subsequently postponed to the following week.

Table 4.1: Results comparison for *IORPS-off* and *IORPS-on*, $\alpha = 0.8$

		IORPS-off				IORPS-on				
ID	|I|	%SgS	%OUt	%Time	Gap	#emg	#s-var	%OUt$_{last}$	#Sg$_{post}$	#Sg$_{n}s$
INST1	130	0.94	0.66	110	0	9	16	0.75	0	0
INST2	133	0.95	0.71	3628	0.66	17	16	0.81	3	1
INST3	135	0.97	0.61	347	0	11	8	0.68	1	0
INST4	136	0.92	0.66	910	0	10	18	0.75	0	0
INST5	136	0.97	0.66	877	0	10	9	0.72	4	0
INST6	137	0.97	0.72	207	0	6	18	0.77	3	0
INST7	141	0.94	0.72	406	0	8	17	0.78	2	0
INST8	142	0.96	0.72	3627	0.05	14	13	0.82	1	0
INST9	143	0.94	0.71	126	0	7	18	0.79	4	0
INST10	143	0.89	0.7	872	0	19	12	0.81	2	1
INST11	146	0.94	0.78	3627	0.95	9	17	0.85	1	0
INST12	149	0.93	0.77	3629	0.38	12	16	0.86	3	1
INST13	155	0.91	0.71	192	0	15	17	0.84	0	1
INST14	157	0.88	0.75	3629	0.47	15	19	0.88	1	1
INST15	159	0.92	0.77	485	0	16	17	0.87	1	2
INST16	160	0.94	0.78	1992	0	13	20	0.9	0	2
INST17	159	0.92	0.77	485	0.01	9	24	0.91	2	2
INST18	157	0.88	0.75	3629	0.47	11	13	0.85	2	0
INST19	155	0.91	0.71	192	0	6	10	0.8	0	0
INST20	158	0.92	0.77	485	0.01	14	13	0.86	3	0
INST21	162	0.91	0.76	3630	0.6	12	16	0.86	3	1
INST22	162	0.89	0.78	3630	1.78	13	20	0.87	3	1
INST23	164	0.9	0.76	127	0	11	14	0.86	4	1
INST24	164	0.85	0.73	3630	1.4	11	19	0.81	2	0
INST25	164	0.85	0.73	3630	1.4	9	17	0.8	1	0
INST26	161	0.89	0.78	3630	1.78	14	21	0.92	4	3
INST27	162	0.91	0.76	3630	0.6	17	9	0.88	3	1
INST28	169	0.9	0.76	127	0.01	12	15	0.86	3	0
INST29	165	0.9	0.76	127	0.01	13	13	0.88	4	1
INST30	161	0.89	0.78	3630	1.78	11	18	0.89	4	0

In table 4.1 is reported a overview of the results provided from *IORPS-off* and the *IORPS-on*, for an allowed initial utilization of Operating Rooms of 80% $\alpha = 0.8$.

In particular, we note a stable percentage of surgeries scheduled across all instances, with a majority clustering near the 90% threshold. This consistency underscores the effectiveness and precision of the initial surgical planning and scheduling. The operating room utilization, as depicted by the "%OUt" column, consistently lies between 66% and 78%. These figures clearly highlight a balanced utilization of the *ORs*, ensuring they are neither underutilized nor reaching saturation. While the "%Time" column showcases fluctuations in computation time—with instances like INST2, INST8, and INST21 being notably prolonged—a closer look in tandem with the "Gap" column reveals that even for these lengthier computations, the solutions achieved are proximate to the optimal.

Moving to the *IORPS-on* model results, the table presents several metrics to understand the model adaptability. The number of emergencies simulated, indicated by "#emg", varies across instances but does not exceed 19. This gives us an idea about the unpredictability factor in the data. The "#s-var" column, marking surgeries with duration variations, suggests that many surgeries do not adhere strictly to their initial estimated duration. Yet, even with these emergencies and duration variations, the final utilization of *ORs*, "%OUt$_{last}$", remains quite high, always surpassing, as expected, the "%OUt" of the *IORPS-off* model. Additionally, the columns "#Sg$_{post}$" and "#Sg$_{n}s$" draw attention to surgeries that were either rescheduled or not scheduled at all. Although these numbers are typically low, certain instances, such as INST26, are notable—with 4 surgeries being rescheduled and 3 left unscheduled, suggesting challenges the model could faces under particular circumstances of variability. Nevertheless, the results denotes, for every all instances, the *IORPS-off* capacity to absorb unforeseen challenges without heavily compromising the initial schedule.

Table 4.2: Results comparison for *IORPS-off* and *IORPS-on*, $\alpha = 0.9$

		IORPS-off				IORPS-on				
ID	|I|	%SgS	%OUt	%Time	Gap	#emg	#s-var	%OUt$_{last}$	#Sg$_{post}$	#Sg$_{n}s$
INST1	130	0.95	0.69	243	0	9	16	0.76	1	1
INST2	133	1	0.78	1112	0	17	16	0.91	4	3
INST3	135	1	0.72	254	0	11	8	0.8	3	1
INST4	136	0.96	0.73	1921	0	10	18	0.79	1	0
INST5	136	1	0.68	1024	0	10	9	0.74	2	0
INST6	137	1	0.75	243	0	6	18	0.8	3	1
INST7	141	0.95	0.75	872	0	8	17	0.81	3	1
INST8	142	0.97	0.74	2231	0	14	13	0.82	3	1
INST9	143	0.97	0.77	413	0	7	18	0.84	5	1
INST10	143	1	0.74	413	0	19	12	0.89	3	3
INST11	146	0.95	0.88	3613	0.13	9	17	0.95	5	6
INST12	149	1	0.79	3627	0.32	12	16	0.87	2	2
INST13	155	0.91	0.71	192	0	15	17	0.84	0	1
INST14	157	0.9	0.78	3627	0.23	15	19	0.87	4	2
INST15	159	0.94	0.81	391	0	16	17	0.94	2	5
INST16	160	1	0.83	2472	0	13	20	0.94	2	9
INST17	159	0.95	0.83	1231	0	9	24	0.92	3	6
INST18	157	0.94	0.84	3612	0.31	11	13	0.91	6	4
INST19	155	1	0.81	1129	0	6	10	0.9	1	5
INST20	158	1	0.84	1121	0	14	13	0.92	5	4
INST21	162	1	0.78	3614	0.05	12	16	0.87	4	4
INST22	162	0.95	0.83	3632	0.81	13	20	0.92	6	4
INST23	164	0.95	0.85	284	0	11	14	0.92	8	3
INST24	164	0.91	0.79	3632	0.72	11	19	0.87	4	3
INST25	164	0.92	0.82	3612	0.42	9	17	0.89	2	4
INST26	161	0.93	0.83	3616	0.97	14	21	0.96	13	10
INST27	162	0.95	0.85	3631	0.78	17	9	0.94	6	7
INST28	169	0.94	0.82	145	0	12	15	0.9	4	8
INST29	165	0.93	0.81	267	0	13	13	0.91	5	8
INST30	161	0.98	0.88	3630	0.12	11	18	0.94	9	6

In Table 4.2 is presented a detailed comparison of results provided by *IORPS-off* and the *IORPS-on*, but this time for a value of α = 0.9. This implies that the average utilization of the *ORs* permitted for the *IORPS-off* model is set at a higher threshold of 90%, compared to the previous (80%).

Firstly, by focusing on the *IORPS-off* model columns, we notice that the "%SgS" values, are generally higher than those presented in Table 4.1. The increase in the number of scheduled surgeries is a direct consequence of the higher allowed utilization rate of 90%. This implies that more surgeries are being fit into the schedule. Similarly, the "%OUt" column shows values that are higher than those in Table 4.1. These highlights underscore that *ORs* are being utilized more intensively.

However, while the increased allowed value of α seems advantageous, it is crucial to evaluate its repercussions. Directing our attention to the *IORPS-on* model columns provides some insight. The columns "#Sg$_{post}$" and "#Sg$_{n}s$", representing rescheduled surgeries and those not scheduled at all, respectively, have values that are generally higher than those observed in Table 4.1. This suggests that while more surgeries are initially scheduled, the higher threshold also leads to a more significant number of surgeries being postponed or even canceled. Such disruptions can induce logistic and reorganization challenges, likely leading to dissatisfaction among patients and medical staff. These challenges can ultimately lead to a perceived decrease in the quality of service among all stakeholders involved.

Comparing the two tables, 4.1 and 4.2, it is evident that while the latter adjusted threshold of 90% does ensure more surgeries are scheduled and a higher average utilization of *ORs*, it comes with its set of challenges. The trade-off between maximizing *ORs* utilization and maintaining a smooth, predictable surgical schedule becomes evident. On the one hand, the value of α equal to 0.9 allows for a denser schedule, results in a higher number of surgeries performed. On the other hand, the increased postponements and cancellations can have disadvantageous effects on the overall service quality, affecting both patients and medical staff.

In summary, while the 90% α threshold might seem more efficient in terms of numbers, it is essential to weigh the qualitative aspects. The potential organizational challenges and dissatisfaction arising from frequent rescheduling and cancellations can overshadow the numerical benefits. Therefore, adopting a balanced approach that takes into account both quantitative and qualitative factors is essential for achieving optimal surgeries management.

Lastly, Table 4.3 presents results for the *IORPS-off* and the *IORPS-on* with

Table 4.3: Results comparison for *IORPS-off* and *IORPS-on*, $\alpha = 1.0$

		IORPS-off				IORPS-on				
ID	|I|	%SgS	%OUt	%Time	Gap	#emg	#s-var	%OUt$_{last}$	#Sg$_{post}$	#Sg$_{n}s$
INST1	130	0.98	0.78	341	0	9	16	0.83	7	3
INST2	133	1	0.78	1043	0	17	16	0.91	4	3
INST3	135	1	0.72	413	0	11	8	0.8	3	1
INST4	136	0.96	0.73	2024	0	10	18	0.79	1	0
INST5	136	1	0.68	1045	0	10	9	0.74	2	0
INST6	137	1	0.75	199	0	6	18	0.8	3	1
INST7	141	1	0.8	912	0	8	17	0.85	8	6
INST8	142	0.97	0.74	1993	0	14	13	0.82	3	1
INST9	143	1	0.79	512	0	7	18	0.85	7	7
INST10	143	1	0.74	537	0	19	12	0.89	3	3
INST11	146	0.98	0.91	3625	0.07	9	17	0.97	4	10
INST12	149	1	0.79	3629	0.32	12	16	0.87	2	2
INST13	155	0.91	0.71	421	0	15	17	0.83	0	1
INST14	157	0.97	0.88	3636	0.18	15	19	0.97	4	17
INST15	159	1	0.92	391	0	16	17	0.97	3	17
INST16	160	1	0.83	1987	0	13	20	0.94	2	9
INST17	159	0.95	0.83	1231	0	9	24	0.92	3	6
INST18	157	0.96	0.85	1789	0	11	13	0.92	7	4
INST19	155	1	0.81	1129	0	6	10	0.9	1	5
INST20	158	1	0.84	1121	0	14	13	0.92	5	4
INST21	162	1	0.78	3613	0.05	12	16	0.87	4	4
INST22	162	1	0.89	3632	0.17	13	20	0.96	6	14
INST23	164	1	0.91	245	0	11	14	0.96	3	10
INST24	164	0.98	0.87	3654	0.63	11	19	0.93	3	8
INST25	164	0.96	0.87	3614	0.21	9	17	0.93	3	6
INST26	161	0.93	0.83	3613	0.97	14	21	0.96	13	10
INST27	162	0.95	0.85	3631	0.78	17	9	0.94	6	7
INST28	169	1	0.91	142	0	12	15	0.93	6	9
INST29	165	0.95	0.87	245	0	13	13	0.94	7	10
INST30	161	0.98	0.88	3631	0.12	11	18	0.94	9	6

a value of α of 1.0. This, again, indicates that the average utilization of the *ORs* permitted for the *IORPS-off* is at the maximum threshold of 100%.

By examining the "%SgS" column. we observe that the values are higher or comparable to those from the previous table 4.2. This is consistent with the notion that a 100% utilization threshold allows for maximum surgeries to be scheduled. The "%OUt" column, indicative of the average utilization of *ORs*, also showcases values that are generally higher than those in Table 4.2, emphasizing that *ORs* are utilized to their full capacity.

However, the real implications of this 100% threshold become evident when examining the repercussions on the *IORPS-on*. Indeed, the columns "#Sg$_{post}$" and "#Sg$_{n}s$" display values that are even higher than those in Table 4.2 and considerably higher than Table 4.1. This trend suggests that while an increased number of surgeries are scheduled initially, this aggressive scheduling leads to a significant rise in the number of surgeries postponed or not scheduled at all.

Comparing tables 4.1, 4.2 and 4.3, we observe a clear trend:

1. As the allowed utilization percentage increases, straightforwardly, the number of scheduled surgeries also rises, and *ORs* are utilized more in-

tensively.

2. However, the higher the utilization threshold, the more pronounced the disruptions become in the form of rescheduled surgeries or those not scheduled at all. The differences between the 100% and 90% thresholds are evident, and the deviations between the 100% and 80% thresholds are even more rigid.

In conclusion, while allowing a 100% α threshold maximizes the number of scheduled surgeries and *ORs* utilization, it does so at a considerable cost. The disruptions in the form of postponed and canceled surgeries can lead to significant organizational challenges, as the implications observed with the 90% threshold in Table 4.2. This further amplifies the trade-off between maximizing utilization and ensuring a smooth, reliable surgical schedule. Prioritizing one aspect might compromise the other, underscoring the need for a balanced approach in surgical scheduling and management.

To provide a comprehensive perspective on the findings, we additionally reported a graphical representation in figures 4.1, 4.2, and 4.3. These histograms present three significant metrics: the average utilization percentage of *ORs*, the number of surgeries rescheduled to the subsequent days versus their initially planned date, and the surgeries that were canceled or deferred to the weeks ahead, respectively.

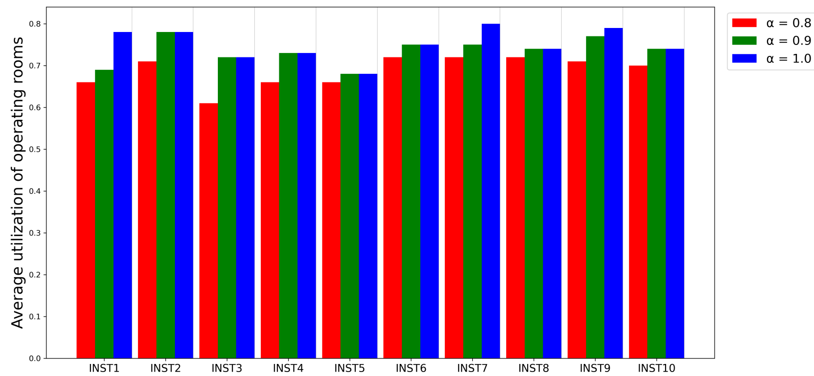
In these figures, the x-axis enumerates different instances of surgeries, consistently with the data presented in tables 4.1, 4.2, and 4.3. To enhance clarity, we have categorized these instances across three rows. The first row focuses on instances where the total number of surgeries (or cardinality) ranges from 130 to 145. The second row address the instances with cardinalities between 146 and 160 surgeries. Finally, the third row is dedicated to the instances where the number of surgeries surpass the 160 mark.

Within each instance, three bars are depicted for each figure, which relate to the value of the parameter α. Specifically, red bars indicate results when α is set at 0.8, green bars represent α at 0.9, and blue bars delineate the outcomes for α equal to 1.0.

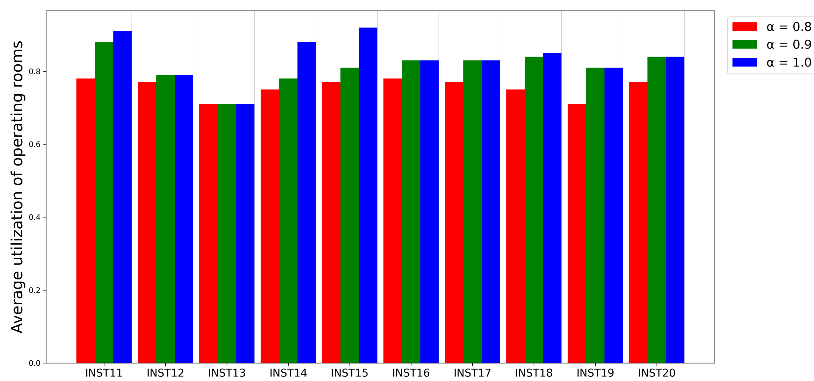
An analysis of figure 4.1 aligns with our expectations: as α increases, there is a tiny rise in the average utilization of *ORs*. This trend is also validated by data in the accompanying tables. Intuitively, this enforces that when there is a propensity to utilize *ORs* more aggressively during the planning phase, it results in higher room utilization.

Upon a more detailed analysis of Figure 4.1, when compared with figures

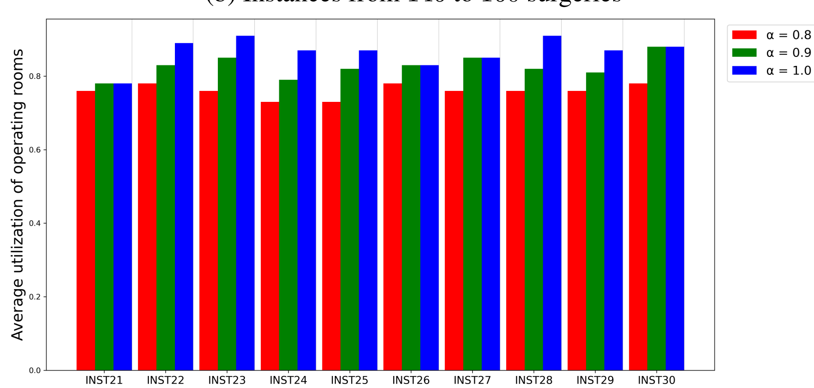
4.2 and 4.3, is depicted a more intricate scenario. A higher value of α is correlated with an increase in the number of surgeries being either postponed or canceled. This keeps enforcing a crucial consideration: while aggressive initial scheduling might project an ostensible increase in operational efficiency, but t could inadvertently compromise the overall service quality. The adverse effects of such an approach extend beyond simple rescheduling; they manifest in the form of logistical challenges pertaining to re-coordination and potential dissatisfaction among the hospital staff. Thus, while on the surface, a packed schedule might appear advantageous, the underlying repercussions, including service quality degradation and operational challenges, cannot be overlooked.



(a) Instances from 130 to 145 surgeries

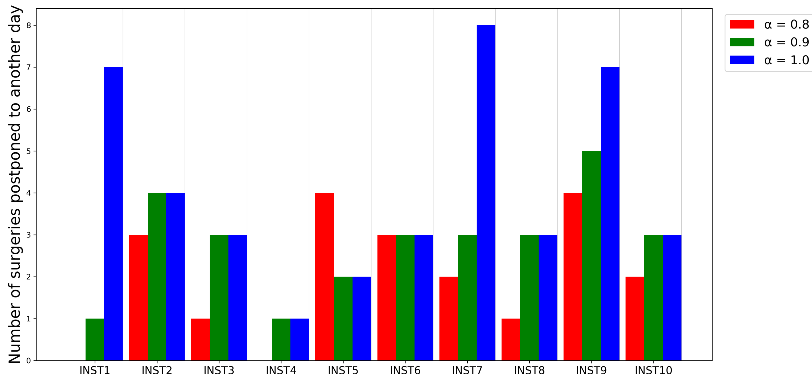


(b) Instances from 146 to 160 surgeries

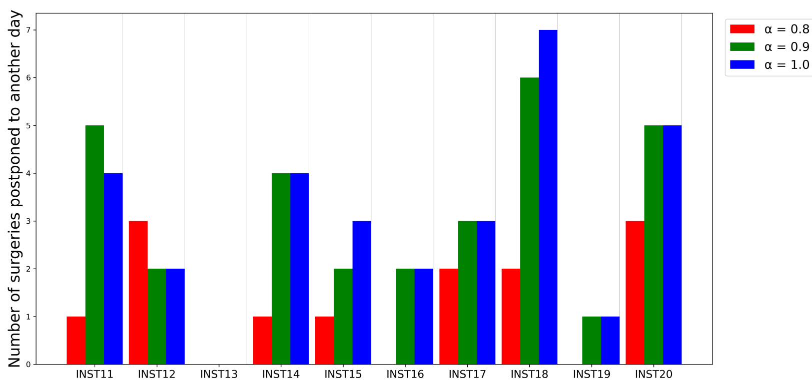


(c) Instances with more than 160 surgeries

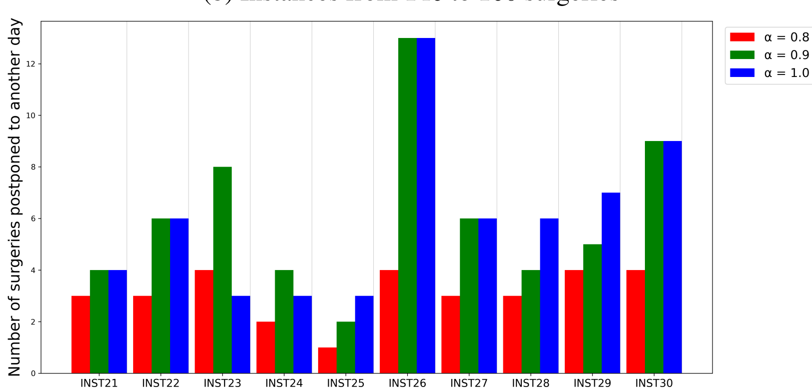
Figure 4.1: Room utilization comparison increasing the value of α



(a) Instances from 130 to 145 surgeries

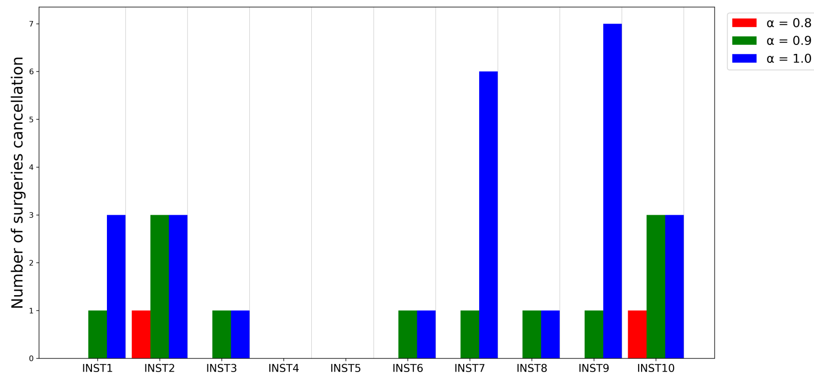


(b) Instances from 146 to 160 surgeries

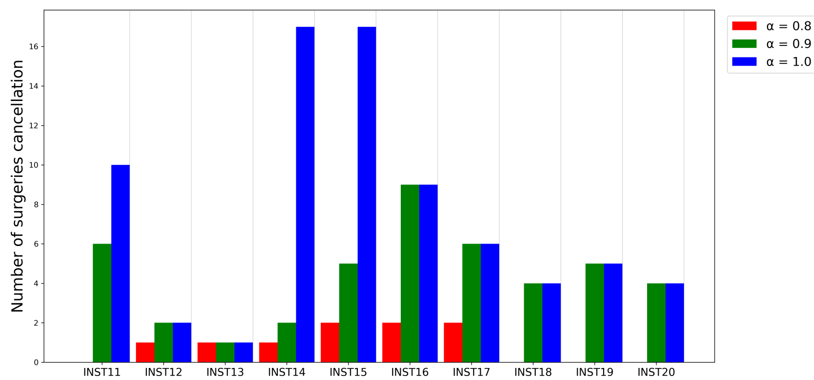


(c) Instances with more than 160 surgeries

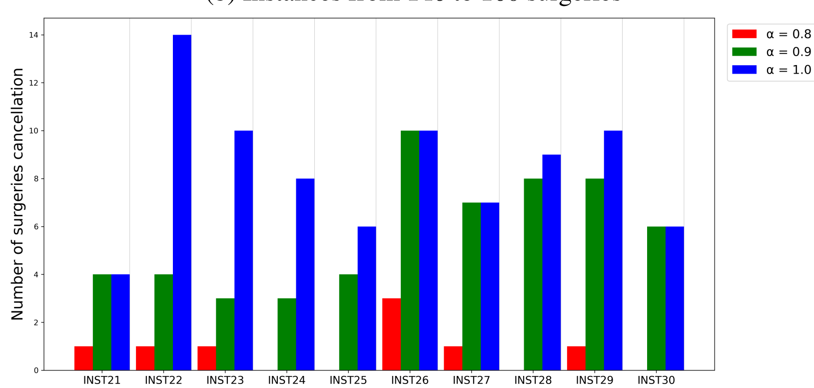
Figure 4.2: Surgeries' day shift comparison increasing the value of α



(a) Instances from 130 to 145 surgeries



(b) Instances from 146 to 160 surgeries



(c) Instances with more than 160 surgeries

Figure 4.3: Surgeries cancelled comparison increasing the value of α

4.5.4 Full knowledge comparison

In this section, we present a comparison between the results provided by *IORPS-on* and those based on hypothetical full knowledge. The results for this assumed knowledge essentially comprise those obtained using the *IORPS-off* model, which takes into account all information regarding variations in surgery durations, as well as the quantity and duration of emergencies, with the assumption that this information is available and does not emerge progressively over the course of the day. In Table 4.4, we showcase this comparison for the same 30-week period described in Section 4.5.2.

Specifically, the table is organized as follows: Column 1, labeled "ID", displays the instance ID. Column 2, denoted as "|I|", represents the total number of surgeries for each instance. Column 3, titled "#emg", details the number of emergency cases simulated for each instance. In Column 4, labeled "#s-var", the focus is on surgeries, specifically reporting the number that deviated from their initially estimated durations. Columns 5 and 7, referred to as "%OUt$_{last}$", evaluate the efficiency of OR utilization. They report the OR utilization rates at the end of each week, calculated based on the "*IORPS-on*" model and "*IORPS-off*" model assuming full knowledge with an α value of 0.8. Likewise, Columns 6 and 8, named "#Sg$_{post}$", enumerate the surgeries rescheduled to different days for the respective models.

The table also provides a comparative perspective by replicating this information under different scenarios in the subsequent columns. Columns 9 through 12 replicate the data from Columns 5 to 8 but for an α value of 0.9, while Columns 13 to 16 present similar information for an α value of 1.0.

The data presented in Table 4.4 provides an insightful comparison between the *IORPS-on* model and the hypothetical scenario of full knowledge across various instances. A notable observation is that, even when assuming full knowledge from the onset, the number of postponed surgeries (#Sg$_{n}s$) remains the same for many instances. Moreover, for the instances where there is a variation, the change in the number of postponed surgeries is marginal. This highlights the effectiveness of the *IORPS-on* model; in many cases, its solutions would be optimal or near-optimal even if full knowledge were available from the beginning.

For example, in instances INST1, INST3, INST4, INST5, and INST6, there is no difference in the number of postponed surgeries between the *IORPS-on* and full knowledge scenarios across various α values. While, in instances where the "*IORPS-on*" model does not achieve complete optimality, such as INST14, INST15, and INST22, the deviation from the optimal number

Table 4.4: Results comparison for *IORPS-on* and the full knowledge

INSTANCE				IORPS-on		Full Knowledge		IORPS-on		Full Knowledge		IORPS-on		Full Knowledge	
				α=0.8				α=0.9				α=1.0			
ID	|I|	#emg	#s-var	%OUt$_{last}$	#Sg$_{ns}$	%OUt$_{last}$	#Sg$_{ns}$	%OUt$_{last}$	#Sg$_{ns}$	%OUt$_{last}$	#Sg$_{ns}$	%OUt$_{last}$	#Sg$_{ns}$	%OUt$_{last}$	#Sg$_{ns}$
INST1	130	9	16	0.75	0	0.75	0	0.76	1	0.77	0	0.83	3	0.85	1
INST2	133	17	16	0.81	1	0.82	0	0.91	3	0.91	3	0.91	3	0.91	3
INST3	135	11	8	0.68	0	0.68	0	0.8	1	0.81	0	0.8	1	0.81	0
INST4	136	10	18	0.75	0	0.75	0	0.79	0	0.79	0	0.79	0	0.79	0
INST5	136	10	9	0.72	0	0.72	0	0.74	0	0.74	0	0.74	0	0.74	0
INST6	137	6	18	0.77	0	0.77	0	0.8	1	0.8	0	0.8	1	0.8	0
INST7	141	8	17	0.78	0	0.78	0	0.81	1	0.82	0	0.85	6	0.86	5
INST8	142	14	13	0.82	0	0.82	0	0.82	1	0.82	0	0.82	1	0.82	1
INST9	143	7	18	0.79	0	0.79	0	0.84	1	0.85	0	0.85	7	0.88	5
INST10	143	19	12	0.81	1	0.82	0	0.89	3	0.9	2	0.89	3	0.9	2
INST11	146	9	17	0.85	0	0.85	0	0.95	6	0.96	5	0.97	10	0.97	10
INST12	149	12	16	0.86	1	0.86	0	0.87	2	0.89	0	0.87	2	0.89	0
INST13	155	15	17	0.84	1	0.85	0	0.84	1	0.84	0	0.83	1	0.84	0
INST14	157	15	19	0.88	1	0.88	1	0.87	2	0.88	1	0.97	17	0.97	17
INST15	159	16	17	0.87	2	0.88	1	0.94	5	0.95	4	0.97	17	0.97	17
INST16	160	13	20	0.9	2	0.9	2	0.94	9	0.94	9	0.94	9	0.94	9
INST17	159	9	24	0.91	2	0.91	2	0.92	6	0.93	5	0.92	6	0.93	5
INST18	157	11	13	0.85	0	0.85	0	0.91	4	0.92	3	0.92	4	0.92	3
INST19	155	6	10	0.8	0	0.8	0	0.9	5	0.92	3	0.9	5	0.92	3
INST20	158	14	13	0.86	0	0.86	0	0.92	4	0.93	3	0.92	4	0.93	3
INST21	162	12	16	0.86	1	0.87	0	0.87	4	0.89	2	0.87	4	0.89	2
INST22	162	13	20	0.87	1	0.87	1	0.92	4	0.93	3	0.96	14	0.96	14
INST23	164	11	14	0.86	1	0.86	0	0.92	3	0.94	1	0.96	10	0.96	10
INST24	164	11	19	0.81	0	0.81	0	0.87	3	0.9	0	0.93	8	0.94	7
INST25	164	9	17	0.8	0	0.8	0	0.89	4	0.91	2	0.93	6	0.94	5
INST26	161	14	21	0.92	3	0.92	3	0.96	10	0.96	10	0.96	10	0.96	10
INST27	162	17	9	0.88	1	0.88	0	0.94	7	0.94	7	0.94	7	0.94	7
INST28	169	12	15	0.86	0	0.86	0	0.9	8	0.92	6	0.93	9	0.93	9
INST29	165	13	13	0.88	1	0.89	0	0.91	8	0.93	6	0.94	10	0.95	8
INST30	161	11	18	0.89	0	0.89	0	0.94	6	0.94	6	0.94	6	0.94	6

of postponed surgeries is still relatively small. This suggests that, the *IORPS-on* model provides solutions that are close to what would be achieved with complete prior knowledge.

Furthermore, the table illustrates a consistent trend in OR utilization rates ($\%$OUt$_{last}$). The similarity in these rates between the *IORPS-on* and full knowledge scenarios across most instances reinforces the model's reliability in efficient resource management.

In summary, Table 4.4 effectively demonstrates the robust performance of the *IORPS-on* model in surgical scheduling and emergency management. Its ability to closely match or even replicate the outcomes of a full knowledge scenario underscores its practical applicability in real-world hospital settings.

4.6 Conclusions

In this Chapter, we addressed a surgical scheduling problem that arises in a local hospital in Naples. The problem includes specific requirements for managing emergencies and potential variations in surgery duration due to complications experienced in practice. To address this, we developed two *ILP* formulations, referred to as *IORPS-off* and *IORPS-on*. These models form the core of an optimization tool that can be used by the hospital to plan an initial weekly surgery schedule and quickly modify it when deviations from the plan occur. The tool can also provide valuable insights on potential surgery department re-configurations, upgrades, and expansions.

Our experimentation confirmed the usefulness of the proposed tool in surgical practice. Furthermore, the mathematical models are flexible and can be adapted in the future to include decisions about surgeons, anesthesiologists, post-operative resources, and mobile equipment, making it a valuable for other hospitals as well. Indeed, it shown that, by accurately tuning the planning paratemers to be utilized in the *IORPS-off*, the *IORPS-on* model demonstrates a good capability to adapt to surgical duration variations and emergencies without drastically modifying the original surgery schedule. Such resilience ensures that surgeries are conducted efficiently, even considering uncertainties, safeguarding both patient welfare and optimizing the resources, guaranteeing in that way a good trade off between costs and quality of the service provided.

From a methodological perspective, future work may integrate the tool with a heuristic framework to overcome computational issues that could arise for larger hospitals.

Finally, we emphasize that the hospital plans to integrate the optimization tool into a broader digitalization project to enhance surgical outcomes. Information technologies can simplify procedures and reduce administrative burdens, ultimately improving healthcare processes and patient outcomes [152, 58]. However, securely sharing information in the context of hospital logistics and supply chain remains a concern [57].

Chapter 5

Optimizing truck and drone delivery system for last-mile logistic in an healthcare perspective: a branch-and-cut approach

Innovations in last-mile logistics are critical for improving delivery efficiency. One such promising approach combines trucks and drones, creating a unique hybrid delivery system. While trucks serve as mobile depots, drones tackle limited delivery ranges, relying on the trucks for battery swaps and parcel transfers. Despite its apparent divergence from healthcare, its application for drone routing in the medical sector, as hinted at by [235], is direct and paves the way for future exploration. Indeed, drones have been widely used in the context of blood products delivery and pharmaceutical industry logistics, as highlighted in 1. The use of drones in these contexts could bring several benefits, such as expedited delivery times, reduced transportation costs, enhanced access to remote or hard-to-reach areas, improved supply chain efficiency, and increased safety by minimizing human exposure to unsafe environments. Furthermore, its implementation can be tailored to actual scenarios, given that drug packages are typically lightweight and compact, thus not posing issues related to size. Additionally, the designated delivery points, represented by pharmacies where staff is consistently available to receive the drugs, eliminate the need for coordination between the drone and the recipient.

5.1 Introduction and healthcare motivation

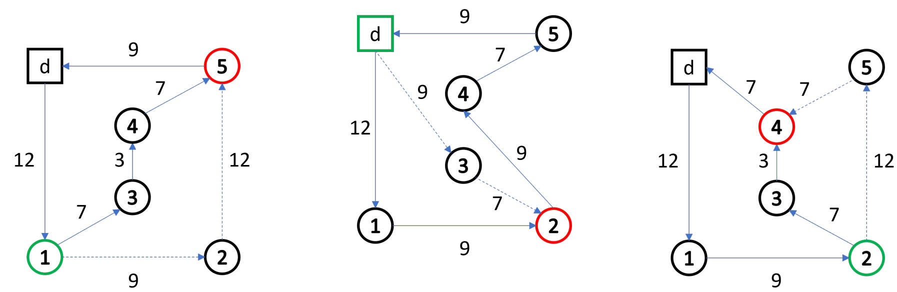
A truck-and-drone system is a hybrid delivery system composed of trucks and drones. The two kinds of vehicles can serve customers in tandem or independently. However, the drone is not completely autonomous because of its limited payload and battery capacity. Thus, the truck acts also as a mobile depot for the drone, swapping its battery and providing the parcels to be delivered. The system is aimed at serving all the customers minimizing the overall delivery time. The usage of such delivery systems and the deriving routing problems represent cutting edge research topic for the Operations Research community and related literature has grown extremely fast, as witnessed by the survey works reported in [170, 248]. In particular, a bunch of variants of truck-and-drone routing problems has been proposed differing for the side constraints envisaging either the operating conditions or different hybrid truck-and-drone delivery system structures. For the first case, we mention, among the others, the variants involving multiple customers within each drone sortie [221], the dependence of the drone energy consumption on the carried weight [178], and the presence of no-fly zones [311]. For the second case, instead, we cite the variants involving multi-truck and multi-drone homogeneous and heterogeneous fleets [200, 300].

Generally speaking, a truck-and-drone system can be conceived as a two-echelon distribution system, whose main complexity issue is represented by the synchronization between the two kinds of involved vehicles. Such an issue characterizes all the two-echelon distribution systems and it is declined in multiple ways depending on the specific features of the considered vehicles. On this basis, with reference to a truck-and-drone system, we can define two different operational schemes: a drone is launched from and recovered at the same location where the truck stops for the whole duration of the sortie; the truck launches a drone at a location and recovers it at a different location. It is easy to understand that the drone battery capacity, or in other words its endurance, differently affects these two operational schemes. Indeed, in the first case, the truck waits for the drone at a location and it is always available for the recovering operation. Thus, such a scheme requires no actual synchronization between the vehicles, but, on the other hand, it does not fully exploit the advantages of such a hybrid delivery system. Instead, the second scheme presents opposite features. Indeed, it has the advantage of simultaneously exploiting the truck and the drone for the delivery operations, but, on the other side, it requires a significant coordination between the two vehicles, which have to necessarily meet at the recovery location within a time period imposed by the drone endurance.

From the routing problem perspective, the first operational scheme configures a truck-and-drone problem which partially recalls the ideas underlying the p-median path problem [18] and the truck-and-trailer routing problem [68]. On the other side, the second operational scheme matches the definition of the first truck-and-drone routing problem defined in the literature [199], referred to as the Flying Sidekick Traveling Salesman Problem (*FS-TSP*).

In short, the *FS-TSP* operations and main assumptions can be summarized as follows. A set of customers must be served by a hybrid delivery system combining a truck and a drone, both placed at a starting depot. The truck plays a twofold role: it serves the customers met along its route and acts as a mobile depot for the drone. More in detail, the truck carries the drone to a location where it is loaded with a package, and it takes off to serve a customer. While the drone is performing the delivery, the truck either serves other customers or directly moves to the landing location where the drone is picked-up after completing the delivery. Without loss of generality, the take-off and landing locations correspond to the customer locations. Drone delivery operations are subject to two main restrictions: limited flight range, due to its battery endurance, and limited load capacity, since it can transport only one package at a time. For safety reasons, the drone can land only at the depot or on the top of the truck. Therefore, the drone must hover at a landing location until the truck arrives, so requiring the synchronization of their movements. After landing on the truck, the drone battery is replaced with a fully charged one, thus it can start a new delivery without idle time. The set of all the operations performed by the drone (take-off, delivery, and landing), is referred to as sortie in the literature. The objective of the system is to minimize the time required for the completion of the delivery operations.

For the sake of comprehension, Figure 5.1 sketches three different solutions of the *FS-TSP* on a small instance with one depot and 5 customers, represented by square and circles, respectively. Take-off and landing nodes are highlighted in green and red, respectively. The solid arrow indicates the movement of the truck (both with and without the drone on board), while the dashed arrow indicates the movement of the drone in flight. The values on the arcs represent the travel times. For the sake of simplicity, it is assumed that the two vehicles move at the same speed. The time required to complete the deliv-



(a) Delivery time = 42

(b) Delivery time = 46

(c) Delivery time = 47

Figure 5.1: Three different *FS-TSP* solutions

ery operations corresponds to the time when both vehicles return to the depot, after having served all the customers. It can be calculated as the sum of two components: the routing time of the truck moving with the drone on board and the time spent by the drone to complete each sortie. For example, in the first solution, the truck moves with the drone on board to serve customer 1 with a travel time of 12. At node 1, the drone is launched to serve customer 2 while the truck serves customers 3 and 4. Then, the truck serves customer 5 where it also picks-up the drone. The duration of the drone path is 21 while that of the truck is 17. Thus, because of synchronization constraints, the duration of the drone sortie is equal to the maximum between these two values. Finally, the two vehicles return together to the depot with a travel time of 9. Consequently, the completion time is 12+21+9 = 42. Coherently with these considerations, it is possible to compute the travel time of the two vehicles, the duration of the sorties, and the delivery time for the other two solutions. It is also important to underline that the validity of such solutions depends on the endurance of the drone. If the drone endurance is lower than the duration of a sortie, such a sortie is unfeasible, thus making the corresponding solution unfeasible as well. For example, if the drone endurance was equal to 20, the only feasible solution would be the third one.

It is straight to understand that the *FS-TSP* represents a more complex variant of the *TSP* which includes synchronization requirements. Thus, it is a NP-hard problem as well, and it is extremely difficult to optimally and/or effectively solve even on instances with few customers. As a result of this complexity, most contributions in literature tackle the *FS-TSP* by heuristic approaches [45, 82, 84], while just a few mathematical programming exact methods have been developed. Moreover, as is highlighted in the next section, available exact methods either suffer of dimensionality drawbacks due to the usage of an exponential number of variables, or suffer of tightness deficiency due to the weakness of formulation constraints. In particular, formulations characterized by path-based variables generally belong to the first class of methods while formulations using three-index and/or time variables in combination with *big-M* constraints usually fall in the second one. This work is aimed at overcoming these inconveniences by providing the following threefold methodological contribution:

- We introduce a new mixed-integer linear programming (*MILP*) arc-based formulation for the *FS-TSP* with a polynomial number of three-index variables.
- We present a Branch-and-Cut (*B&C*) algorithm for solving the *FS-TSP* based on new families of valid inequalities and different variable fixing strategies. In particular, the valid inequalities are grounded on the vehicle coordination while the variable fixings are based on drone endurance.
- We provide an extensive experimentation on the benchmark instances used by the state-of-the-art algorithms for the *FS-TSP* [47, 83, 247]. The computational results show that our approach is competitive or outperforms the state-of-the-art approaches. Indeed, it either provides the optimal solution for some instances that have never been solved before or returns improved upper bounds to those that are still unsolved.

The remainder of the chapter is organized as follows: in Section 5.2, we revise exact solution methods for the *FS-TSP*; in Section 5.3, we recall the *FS-TSP* and we present the *MILP* formulation; Section 5.4 describes the Branch-and-Cut algorithm, the valid inequalities, and the variable fixings; Section 5.5 is devoted to the computational results; finally, conclusions are given, and perspectives on future works on this topic are discussed in Section 5.6.

5.2 Literature review

The research activity on truck-and-drone delivery systems and, in particular, related routing problems have been significant in the last few years. For a complete review on variants of the *FS-TSP* the interested reader is referred to [178, 202], on single-truck multi-drone to [65, 200], and on multi-truck single/multi-drone to [262, 300]. In this work, we focus our revision on exact solution methods for the *FS-TSP*. We describe each method from a twofold perspective to evaluate its effectiveness. From a mathematical point of view, the evaluation is based on the kinds of variables and constraints used to formulate the problem. In particular, arc-based formulations generally present a polynomial number of two/three-index variables, while path-based formulations use an exponential number of variables. The first kind of formulations is generally easy to implement within currently available solver frameworks, while the second one involves the development of ad-hoc solution methods. Moreover, some formulations address the time dimension and/or the synchronization between the two vehicles by using *big-M* constraints that generally lead to weak linear relaxations. From a computational point of view, we highlight the instances used for the experimentation and briefly describe the resulting performance since the instance size and the computational burden are aspects generally used to evaluate a solution method. On this basis, we point out that the discussed methods consider one or more of the following three test beds for the *FS-TSP*:

- The set of 72 instances proposed in [199] with 10 customers and different endurance values (20 and 40 minutes) (*M10-instances*).
- The set of 240 instances proposed in [199] for a different truck-and-drone problem and adapted in [83] for the *FS-TSP*. These instances have 20 customers, and the endurance is set to either 20 or 40 minutes (*M20-instances*).
- The set of 300 instances proposed in [222] for a variant of the *FS-TSP* with up to 39 customers and an endurance equal to 20 minutes (*POI-instances*).

The first work addressing the *FS-TSP* is reported in [199]. In this work, the authors present for the first time the *FS-TSP* and propose a *MILP* model to solve it which uses three-index and time variables and *big-M* constraints. None of the *M10-instances* is solved by a commercial *MILP* solver within a time limit of 1800 seconds.

Starting from the *MILP* proposed in [199], a two- and a three- index formulation with time variables and *big-M* constraints are presented in [85] and tested on the *M10-instances*. The best results are obtained using the two-index formulation which solves 33 and 26 instances with a drone endurance of 20 and 40 minutes, respectively. The same authors successively propose an improved formulation with only two-index variables and *big-M* constraints in [83], able to optimally solve for the first time all *M10-instances*. In addition, they test the new formulation on the *M20-instances*, providing the optimal solution of 58 and 2 instances with a drone endurance equal to 20 and 40 minutes, respectively. Recently, the same authors present a solution method based on a branch-and-bound scheme coupled with an *assignment problem* in [84]. They are able to solve only the *M10-instances* and a few instances with up to 19 customers of the *POI-instances*.

Almost contemporarily to [83], two new formulations for the *FS-TSP* are proposed in [263]. Two-index and time variables characterize both formulations, but the second one uses a joint representation of the drone sortie to reduce the number of variables. They solve to optimality: all the *M10-instances*; 32 *M20-instances* with an endurance equal to 20 minutes; 3 *M20-instances* with an endurance equal to 40 minutes .

A new formulation based on an extended graph representation of the *FS-TSP* with path-based variables but without *big-M* constraints is proposed in [47]. The authors develop a column-and-row generation approach to solve the proposed formulation. They are able to solve all the *M10-instances*, and 80 and 5 *M20-instances* of the considering a drone endurance equal to 20 and 40 minutes, respectively. The same authors then extend their method in [48] and [45], including cyclic sortie and trying to use the proposed method in combination with data science techniques, respectively.

The method currently able to optimally solve the largest *FS-TSP* instances is presented in [247]. The authors propose a compact formulation considering the truck and drone as a single entity leading to an exponential number of variables. Therefore, the authors develop a branch-and-price method to solve the formulation where the pricing problem is solved using a dynamic programming approach. The method is tested on a modified version of the *POI-instances*. In particular, the authors round up all the travel time to make them integer. The results show that the proposed method solved almost all the instances with up to 39 customers.

From this review, it is clear that our method is the only one considering an arc-based formulation with a polynomial number of three-index variables and without time variables and *big-M* constraints.

5.3 Flying sidekick traveling salesman problem (*FS-TSP*)

In this section, we first briefly recall the *FS-TSP* description as defined in [47, 83, 199, 247] and simultaneously introduce the problem notation, then we present the proposed original formulation.

5.3.1 Description and notation

The *FS-TSP* considers a set of customers, C, that have to be served exactly once, either by the truck or the drone. The two vehicles depart from and return to a single depot once. The proposed formulation splits the depot into a source node, s, and a destination node, t. Moreover, let V be the set of nodes, $V = C \cup \{s, t\}$, and A the set of arcs, $A = \{(i, j) : i \in C \cup \{s\}, j \in C \cup \{t\}\}$. The travel time of the truck (drone) on each arc $(i, j), (i, j) \in A$, is indicated with $t_{ij}(d_{ij})$.

The truck has an infinite capacity and serves all the customers met along its route. Moreover, it acts as a mobile depot for the drone providing the packages and dispatching the drone to serve the customers.

The drone can serve one customer per sortie. Therefore, a sortie is characterized by a launch node, a rendezvous node, and a customer node. The launch and the rendezvous nodes can be either the depot or a customer node. In addition, they must be different except if there is only one sortie starting from and ending at the depot. The duration of a sortie is limited by the drone endurance (Dtl). The truck must collect the drone before its flight time exceeds its endurance because the drone cannot land due to safety reasons. On this basis, it is clear that the two vehicles must be synchronized at the beginning and at the end of a sortie. Moreover, setup times arise when the drone is launched (launch time, *SL*) or retrieved (recovery time, *SR*). We point out that there is no launch time in the depot and that the recovery time is included in the endurance computation. For the sake of readability, the notation is also reported in Table 5.1.

The objective is to minimize the duration of the delivery process, i.e., the time needed by the two vehicles to serve all the customers and return to the depot.

On the basis of this description, the *FS-TSP* involves three kinds of decisions: identification of the customers served by the drone; selection of launch and rendezvous nodes for each drone sortie; definition of the truck route.

Table 5.1: Notation used for the *FS-TSP* formulation.

Problem notation	
Sets:	
C	set of customers
$\{s\}$	tandem starting node
$\{t\}$	tandem ending node
V	set of nodes, $C \cup \{s, t\}$
A	set of arcs
Parameters:	
t_{ij}	truck travel time on the arc (i,j)
d_{ij}	drone travel time on the arc (i,j)
Dtl	drone endurance
SL	launch service time
SR	recovery service time

5.3.2 Formulation

To model the *FS-TSP* by a *MILP* formulation, we introduce the following sets of variables based on the previous problem description and notation:

- y_{ij}, $(i, j) \in A$, is equal to 1 if the truck travels along the arc (i, j) (with or without the drone on board), 0 otherwise.

- x_{ij}, $(i, j) \in A$, is equal to 1 if the drone flies along the arc (i, j), 0 otherwise.

- γ_{ij}^{h}, $(i, j) \in A$, $h \in C$, is equal to 1 if the arc (i, j) is traveled by the truck while the drone is performing the drone sortie serving customer h, 0 otherwise.

- θ^{h}, $h \in C$, is equal to 1 if customer h is served by the drone, 0 otherwise.

- ω_{i}^{h}, $i \in C \cup \{s\}, h \in C$, is equal to 1 if node i is the launch node of the drone sortie serving customer h, 0 otherwise.

- δ_{j}^{h}, $j \in C \cup \{t\}, h \in C$, is equal to 1 if node j is the rendezvous node of the drone sortie serving customer h, 0 otherwise.

- σ^h, $h \in C \cup \{t\}$, is a continuous variable indicating the truck waiting time at the destination node of the sortie serving customer h. If customer h is served by the drone, then $\sigma^h = 0$.

Consistently with this notation and variables, the makespan of the delivery process, i.e., the objective function, can be written as:

$$\min \sum_{(i,j)\in A} t_{ij} y_{ij} + \sum_{h\in C} (SL + SR)\, \theta^h - \sum_{h\in C} SL\, \omega_s^h + \sum_{h\in C\cup\{t\}} \sigma^h$$

where the first term measures the duration of the truck path, the second and the third term evaluate the total launch and recovery service time (the launch service time at the depot is equal to 0), and the fourth term assesses the total truck waiting time.

Then, the set of constraints can be divided into the following six families.

Truck routing constraints

$$\sum_{j:(s,j)\in A} y_{sj} = \sum_{i:(i,t)\in A} y_{it} = 1 \qquad (5.1)$$

$$\sum_{j:(i,j)\in A} y_{ij} = \sum_{j:(j,i)\in A} y_{ji} \leq 1 \qquad i \in C \quad (5.2)$$

$$\sum_{i,j\in S|(i,j)\in A} y_{ij} \leq \sum_{h\in S\setminus\{q\}} (1-\theta^h) \qquad 2 \subseteq S \subseteq V,\ \ q \in S \setminus \{s,t\} \quad (5.3)$$

They impose that there is a truck path from the origin node s to the destination node t, in any feasible solution. Constraints (5.3) are the subtour elimination constraints. As in the formulation of the *orienteering problem* [126], given a subset of nodes S, constraints (5.3) impose that the number of arcs with origin ad destination in S traveled by the truck must be less than the number of nodes served by the truck in the subset S.

Truck path and launch/rendezvous linking constraints

$$\sum_{j:(s,j)\in A} \gamma_{sj}^h = \omega_s^h \qquad h \in C \quad (5.4)$$

$$\sum_{j:(i,t)\in A} \gamma_{it}^h = \delta_t^h \qquad h \in C \quad (5.5)$$

$$\sum_{j:(i,j)\in A} \gamma_{ij}^h - \sum_{j:(i,j)\in A} \gamma_{ji}^h = \omega_i^h - \delta_i^h \qquad i, h \in C \quad (5.6)$$

They impose that there is a truck path without the drone onboard from the launch node to the rendezvous node of each drone sortie serving customer h. The constraints (5.4) ensure that if a sortie includes the tandem starting node, then the launch node of the sortie must be the starting node itself. Similarly, the constraints (5.5) ensure that a sortie must finish at the tandem ending node if it is part of the sortie. Lastly, the constraints (5.6) guarantee that a sortie must start at a launch node and end at a rendezvous node.

Single assignment constraints

$$y_{sj} + x_{sj} \leq 1 \qquad (s,j) \in A \quad (5.7)$$

$$y_{it} + x_{it} \leq 1 \qquad (i,t) \in A \quad (5.8)$$

$$y_{ij} + x_{ij} + x_{ji} \leq 1 \qquad i,j \in C|(i,j) \in A \quad (5.9)$$

$$\sum_{j:(h,j)\in A} y_{hj} + \theta^h = 1 \qquad h \in C \quad (5.10)$$

Constraints (5.7) and (5.8) impose that each arc cannot be traveled by the truck and the drone simultaneously for the arcs exiting from the starting node and for the arcs entering into the ending node, respectively. Constraints (5.9) impose that each arc between two customers can not be traveled by the truck and the drone simultaneously and avoid loops in which the launch and recovery nodes of a sortie are the same node. Constraints (5.10) impose that each customer must be served either by the truck or by the drone.

Consistency constraints

$$\sum_{h\in C} \gamma_{ij}^h \leq y_{ij} \qquad (i,j) \in A \quad (5.11)$$

$$\sum_{i\in V\setminus\{t,h\}} \omega_i^h = \sum_{j\in V\setminus\{s,h\}} \delta_j^h = \theta^h \qquad h \in C \quad (5.12)$$

$$x_{ij} \leq \theta^i + \theta^j \qquad (i,j) \in A \quad (5.13)$$

$$x_{ij} \leq \omega_i^j + \delta_j^i \qquad (i,j) \in A \quad (5.14)$$

$$\sum_{j:(i,j)\in A} x_{ij} = \sum_{h\in C\setminus\{i\}} \omega_i^h + \theta^i \leq 1 \qquad i \in C \quad (5.15)$$

$$\sum_{i:(i,j)\in A} x_{ij} = \sum_{h\in C\setminus\{j\}} \delta_j^h + \theta^j \leq 1 \qquad j \in C \quad (5.16)$$

Constraints (5.11) ensure the consistency between the arcs traveled by the truck during a sortie and the origin-destination truck path. Moreover, if a customer

h is served by the drone ($\theta^h = 1$), constraints (5.12) impose that the corresponding sortie must have a launch and a rendezvous node. Constraints (5.13) guarantee that if the drone travels the arc $(i,j),(i,j) \in A$, then either node i or node j is a customer served by the drone. Constraints (5.14) ensure that if the drone travels the arc $(i,j),(i,j) \in A$, then either node i is the launch node of the sortie serving node j or node j is the rendezvous node of the sortie serving node i. Constraints (5.15) guarantee that an arc coming out of a node i can be traveled by the drone only if the drone serves i or it is the launch node of a sortie. Finally, constraints (5.16) ensure that an arc entering a node j can be traveled by the drone only if the drone serves j or it is the rendezvous node of a sortie.

Drone endurance constraints

$$\sum_{(i,j)\in A} t_{ij}\gamma_{ij}^h \leq (Dtl - SR)\theta^h \qquad h \in C \quad (5.17)$$

$$\sum_{i\in V\setminus\{t\}} d_{ih}\omega_i^h + \sum_{j\in V\setminus\{s\}} d_{hj}\delta_j^h \leq (Dtl - SR)\theta^h \qquad h \in C \quad (5.18)$$

Constraints (5.17) impose an upper bound on the duration of the truck path of the sortie serving customer h, while constraints (5.18) impose the same upper bound on the duration of the drone path of the sortie.

Waiting time constraints

$$\sum_{i\in V\setminus\{t\}} d_{ih}\omega_i^h + \sum_{j\in V\setminus\{s\}} d_{hj}\delta_j^h - \sum_{(i,j)\in A} t_{ij}\gamma_{ij}^h \leq \sigma^h \qquad h \in C \quad (5.19)$$

Constraints (5.19) are needed to consider the truck waiting time in the objective function.

Based on the model just described, it is possible to highlight the main differences with the formulations previously described in the literature review. Specifically, the proposed model differs in three main aspects: the expression of the objective function, the subtour elimination constraints, and the constraints that ensure synchronization between the two vehicles.

Formulations that use time-variables [85, 247, 263] have such variables in each of these elements of the formulation. Specifically, the subtour elimination constraints and synchronization constraints are Big-M constraints. Therefore, these formulations are generally characterized by a weak lower bound at the root node. In contrast to the previous ones, the formulations proposed in [83, 263] use time-variables, but these are only used for synchronization while the objective function is expressed as the sum of travel times and waiting times at the rendezvous nodes. This still involves the presence of Big-M constraints, but in a smaller number and therefore a better value of the lower bound.

Finally, the main difference with the formulation proposed in [47] is represented by the number of variables and constraints. Specifically, this formulation involved the use of an exponential number of variables for the sorties, as opposed to the polynomial order γ_{ij}^{h} variables of the proposed formulation. It is clear that the presence of an exponential number of variables also entails the presence of an exponential number of constraints to ensure consistency with the rest of the model. Such constraints are not necessary in the proposed model, and therefore the only exponential number of constraints are the subtour elimination constraints present also in the formulation reported in [47].

5.4 Branch-and-cut approach

The proposed formulation contains an exponential number of constraints due to the subtour elimination constraints (5.3). Therefore, we developed a Branch-and-Cut (*B&C*) procedure that solves the problem without constraints (5.3) and then adds to the formulation only the *lazy constraints* (5.3) violated each time an integer solution is found. Moreover, the *B&C* effectiveness is further improved by the integration of several valid inequalities and variable fixings. The valid inequalities are introduced to strengthen the proposed formulation, increasing the quality of the lower bound. The variable fixings are implemented to reduce the solution space. In this section, we begin by introducing the valid inequalities, then discuss the variable fixing strategies. Finally, we present the implementation settings for the *B&C*.

5.4.1 Valid inequalities

This subsection presents four families of valid inequalities based on the coordination of the two vehicles (i.e., the coherency between the truck and the drone movements when operating in tandem).

Cut inequalities

Proposition 1. *Given the graph $G(V, A)$ and two nodes p and q, $p,\ q \in V$, let $Cut(p, q)$ be a generic cut on the graph G between the nodes p and q, and* *let $h, h \in C$, be a generic customer. Then, the cut inequalities:*

$$\sum_{(i,j)\in Cut(s,h)} y_{ij} \geq 1 - \theta^h \tag{5.20}$$

and

$$\sum_{(i,j)\in Cut(h,t)} y_{ij} \geq 1 - \theta^h \tag{5.21}$$

are valid for the FS-TSP formulation.

Proof. If the drone does not serve customer h ($\theta^h = 0$), there must be a truck path from the origin node s to h and from h to the destination node t in any feasible solution. Therefore, an arc must cross any cut from s to h and from h to t. □

It is also simple to prove that constraints (5.20) and (5.21) can be used as *subtour elimination* constraints. Indeed, they guarantee that the set of the arcs traveled by the truck, $\{(i,j) \in A \mid y_{ij} = 1\}$, define a connected graph and constraints (5.2) prevent the truck from coming back to a previously visited node. So, any solution presenting a subtour violates at least an inequality in (5.20), (5.21) or (5.2). The cut inequalities (5.20) and (5.21) are separated and added to the formulation at each node belonging to the first three levels of the enumeration tree. The separation procedure solves the *minimum cut problem* [150], from the origin node s to h and from h to the destination node t on the weighted graph $G'(V, A, W)$ where the weight w_{ij} on the generic arc $(i,j), (i,j) \in A$, is given by the value of the corresponding y_{ij} variable. If the value of the minimum cut is less than $1 - \theta^h$, then the violated cut inequality is added to the formulation.

Proposition 2. *Let h and k be two customer nodes, $h,\ k \in C$. Let $Cut(h,k)$ and $Cut(k,h)$ be a cut between the nodes h and k and nodes k and h, respectively. The $2-$cut inequality:*

$$\sum_{(i,j)\in Cut(h,k)} y_{ij} + \sum_{(i,j)\in Cut(k,h)} y_{ij} \geq 1 - \theta^h - \theta^k \tag{5.22}$$

is valid for the FS-TSP formulation.

Proof. If customers h and k are not served by the drone ($\theta^h = \theta^k = 0$), then the truck either travels from h to k or it travels from from k to h. Therefore, a cut from h to k or from k to h must be crossed by an arc traveled by the truck. □

As the previous set of cut inequalities, the 2-cut inequalities (5.22) are separated and added to the formulation at each node belonging to the first three levels of the enumeration tree. The separation procedure solves the *minimum cut problem* from the k to h and from h to k on the weighted graph $G'(V, A, W)$ where the weight w_{ij} on the generic arc $(i, j), (i, j) \in A$, is given by the value of the corresponding y_{ij} variable. If the value of the minimum cut is less than $1 - \theta^h - \theta^k$, then the violated cut inequality is added to the formulation.

Finally, the following proposition defines another family of cut valid inequalities by considering the sets of ω and δ variables.

Proposition 3. *Given a couple of nodes* p *and* $q,\ p,\ q \in V$, *and a customer node* $h,\ h \in C$, *the following inequality:*

$$\sum_{(i,j)\in Cut(p,q)} y_{ij} \geq \omega_p^h + \delta_q^h - 1 \tag{5.23}$$

is valid for the FS-TSP formulation.

Proof. If customer h is served by a drone launched from node p and retrieved at node q ($\omega_p^h = \delta_q^h = 1$), then the truck travels from p to q. Therefore, a cut from p to q must be crossed by an arc traveled by a truck. □

The cut inequalities (5.23) are separated and added to the formulation at each node belonging to the first three levels of the enumeration tree. The separation procedure solves the *minimum cut problem* from the origin node p to the destination node q on the weighted graph $G'(V, A, W)$ where the weight w_{ij} on the generic arc $(i, j), (i, j) \in A$, is given by the value of the corresponding y_{ij} variable. If the value of the minimum cut is less than $\omega_p^h + \delta_q^h - 1$, then the violated cut inequality is added to the formulation.

Drone sortie inequalities

Let $i - k - j$ be a triple of nodes in the set V, if $d_{ik} + d_{kj} \geq Dtl - Sr$, then the sortie is infeasible and the inequality:

$$\omega_i^k + \delta_j^k \leq 1 \tag{5.24}$$

is valid for the *FS-TSP* formulation. Valid inequalities (5.24) are polynomial in number since the number of triple of nodes is $|V|^3$. As a result, these inequalities can be included at the beginning of the formulation.

Truck sortie inequalities

Let A' be a subset of arcs, $A' = \{(i,j),(i,j) \in A : t_{ij} \geq Dtl - SR\}$, if $\omega_i^k = 1$, then $\sum_{l:(l,j)\in A'} \gamma_{lj}^k = 0$ for each $k \in C$ and for each $l \in V$. Therefore, the inequality:

$$\sum_{l:(l,j)\in A} \gamma_{lj}^k \leq 1 - \omega_i^k \tag{5.25}$$

is valid for the *FS-TSP* formulation.

Likewise, if $\delta_j^k = 1$ then $\sum_{l:(i,l)\in A'} \gamma_{il}^k = 0$ for each $k \in C$ and for each $i \in V$. Therefore, the inequality:

$$\sum_{l:(i,l)\in A} \gamma_{il}^k \leq 1 - \delta_j^k \tag{5.26}$$

is valid for the *FS-TSP* formulation.

Valid inequalities (5.25,5.26) are polynomial in number since the number of arcs is approximately $|V|^2$. As a result, these inequalities can be included as user cuts at the beginning of the formulation.

Let $P(i,j)$ be a path from i to $j, (P(i,j) = \{(i,s_1),(s_1,s_2),...,(s_{n-1},j)\}$, and $D(P(i,j))$ be its track travel time, $D(P(i,j)) = \sum_{(i,j)\in P(i,j)} t_{ij}$. If $D(P(i,j)) > Dtl - SR$ then the inequalities:

$$\sum_{(i,j)\in P(i,j)} \gamma_{ij}^k \leq |P(i,j)| - 1 \quad k \in C \tag{5.27}$$

are valid for the *FS-TSP* formulation.

Valid inequalities (5.27) are exponential in number since the number of possible paths is exponential to the number of nodes. Therefore, we add to the formulation only those for which $|P(i,j)| \leq 2$, where $|P(i,j)|$ denotes the number of arcs in the path $P(i,j)$.

Variable upper bounds

Let (i,k) and (k,j) be two arcs traveled by the drone from the launch node i to the customer k and from the customer k to the rendezvous node j, then the variable upper bounds:

$$\omega_i^k \leq x_{ik} \quad (i,k) \in A \tag{5.28}$$

and

$$\delta_j^k \leq x_{kj} \quad (k,j) \in A \tag{5.29}$$

are valid for the *FS-TSP* formulation. Variable upper bounds (5.28-5.29) are polynomial in number since they are related to the number of arcs. As a result, these inequalities can be included at the beginning of the formulation.

5.4.2 Variable fixing strategies

This subsection presents two families of variable fixings based on the drone endurance. In particular, for each arc $(i,j),(i,j) \in A$, we can set:

$$x_{ij} = 0 \quad (i,j) \in A: \; d_{ij} > Dtl - SR \tag{5.30}$$

and

$$\gamma_{ij}^k = 0 \quad k \in C, \; (i,j) \in A: \; t_{ij} > Dtl - SR \tag{5.31}$$

Moreover, if the triangular inequality with respect to the drone travel times and the truck travel times holds and $d_{ij} = d_{ji} \leq t_{ij} = t_{ji}$, for each arc $(i,j),(i,j) \in A$, we can set to zero also other γ_{ij}^k variables with $t_{ij} \leq Dtl - SR$ by considering the following proposition.

Proposition 4. *If $d_{ik} + d_{kj} > 2 \times (Dtl - SR) - t_{ij}$, then $\gamma_{ij}^k = 0$ in any feasible solution.*

Proof. Any sortie where the truck travels the arc (i,j) cannot be in any feasible solution.

Case 1: *Sortie $i - k - j$.*
If $d_{ik} + d_{kj} + t_{ij} > 2 \times (Dtl - SR)$, then either $d_{ik} + d_{kj} > Dtl - SR$ or $t_{ij} > Dtl - SR$ violate an endurance constraint.

Case 2: *Sortie $p - k - q$ with $p \neq i$ or $q \neq j$.*
If the truck travels the arc (i,j), since the triangular inequality holds, the truck time of the sortie is greater than or equal to $t_{pi} + t_{ij} + t_{jq}$ while the drone time of the sortie is equal to $d_{pk} + d_{kq}$. If $d_{ij} \leq t_{ij}$ for each $(i,j),(i,j) \in A$, the sum of the truck time and the drone time of the sortie $d_{pk} + t_{pi} + t_{ij} + t_{jq} + d_{kq}$ can be decreased by considering the sum $d_{ik} + t_{ij} + d_{kj}$. So, as in **Case 1**, the sortie is not feasible.

□

5.4.3 Implementation settings

The implementation of the *B&C* is based on a classical framework, envisaging the solution of a relaxed problem and the introduction of violated cuts. Thus, it starts with the solution of the linear programming relaxation obtained considering:

- the set of x, y, γ, θ, δ, and ω variables;
- the subset of truck routing constraints (5.1-5.2);
- the truck path and launch/rendezvous linking constraints (5.4-5.6);
- the single assignment constraints (5.7-5.10);
- the consistency constraints (5.11-5.16);
- the drone endurance constraints (5.17-5.18);
- the waiting time constraints (5.19).

The relaxation is eventually strengthened with the integration of the the valid inequalities (5.24), (5.25), (5.26), (5.28), and (5.29), the subset of the valid inequalities (5.27) with $|P_{ij}| \leq 2$, the variable fixing strategies (5.30) and (5.31). Then, at each node of the first three levels of the enumeration tree we separate the valid inequalities (5.20-5.23) by using a max flow separation procedure for both integer and fractional solutions [44]. These cuts are separated if the violation exceeds a threshold equal to 0.01. Successively, at each other node of the enumeration tree, if an integer solution is found, we separate the truck routing constraints (5.3).

5.5 Computational results

This section presents and discusses the computational results of the experimentation performed to evaluate and validate the proposed formulation and the Branch-and-Cut (*B&C*) algorithm. The *B&C* algorithm has been coded in C language using Cplex 12.7 with default setting as *MILP* solver, imposing a computation time limit of 1 hour. The experiments are performed on an Intel(R) Core(TM) i7-8700, 3.20 GHz, 16.00 GB of RAM.

The experimentation results are presented in two different subsections. The first subsection shows the impact of valid inequalities and variable fixing strategies on the performance of the *B&C* algorithm. In the second subsection, the performance of the proposed *B&C* algorithm is assessed by comparing it with state-of-the-art solution approaches for the *FS-TSP*.

These three settings are tested on the *M10-*, *M20-* and *POI-instances* described in Section 5.2.

5.5.1 Impact of valid inequalities and variable fixing strategies

In this subsection, we assess the impact of variable fixing strategies and valid inequalities from different perspectives: number of instances optimally solved, gap, running time and quality of the continuous relaxation at the root node. To this end, we tested three different settings:

- *MILP*: this setting involves solving the *MILP* formulation as described in Section 5.3.
- *MILP-PVI*: this setting involves solving the *MILP* formulation integrated with all the polynomial valid inequalities and variable fixing strategies, i.e., constraints (5.24-5.31). This setting considers all the inequalities that can be added upfront to the model.
- *B&C*: this setting involves solving the *MILP* formulation integrated with valid inequalities and variable fixing strategies. This setting considers all the proposed improvements, including the ones requiring a separation procedure, i.e., constraints (5.20-5.23).

We emphasize that we do not consider the *POI-instances* for evaluating the impact of the variable inequalities and variable fixing strategies. This is because, for the *FS-TSP*, the travel time of these instances is rounded up. Therefore, the integer nature of the travel time of these instances can lead to some bias in the results.

Results on *M10-instances*

The three settings are able to solve all the instances within the time limit. Therefore, Table 5.2 provides a comparison of the different settings in terms of running times on the *M10-instances*. The instances are grouped based on three different criteria: the position of the depot (Dep), the speed of the drone (s_d), and the drone endurance. For the depot position, the letter "a" indicates that the depot is located near the center of gravity of the customers, while "b," "c," and "d" represent the following (x, y) coordinates: (4.0, 2.7), (4.0, 0.0), and (4, -2.7), respectively. For each instance group, we indicate the number of

Table 5.2: Impact of valid inequalities and variable fixing strategies on running times for *M10-instances*

		MILP		MILP - PVI		B&C	
	# Ins	Avg time	Max time	Avg time	Max time	Avg time	Max time
Dep = a	18	1.52	3.52	2.56	6.58	2.2	4.74
Dep = b	18	1.29	2.60	2.20	4.75	2.00	3.60
Dep = c	18	1.15	2.26	1.63	4.93	1.69	3.30
Dep = d	18	1.24	3.59	1.44	3.14	1.61	3.72
s_d = 15	24	1.04	2.13	1.40	3.06	1.55	3.72
s_d = 25	24	1.63	3.59	2.28	6.58	2.20	3.42
s_d = 35	24	1.23	2.74	2.22	4.93	1.88	4.74
Dtl = 20	36	1.3	3.59	2.41	6.58	1.34	3.42
Dtl = 40	36	1.30	2.39	1.53	4.93	2.42	4.74
All	72	1.30	3.59	1.97	4.93	1.88	4.74

instances in the column "# Ins". Then, for each setting we report the average and the maximum running time in the columns "Avg time" and "Max time", respectively.

The results demonstrate that all three settings can solve each instance of this test bed within a matter of seconds. Furthermore, it is observed that the running times of the two settings that incorporate the proposed enhancements (*MILP-PVI* and *B&C*) are either comparable or slightly higher than those of the basic setting (*MILP*). This can be attributed to the fact that the formulation associated with the *MILP* setting is already efficient for these small instances, and therefore, the valid inequalities and variable fixing strategies only augment the size of the formulation that needs to be solved, without having any other significant impact. However, the *B&C* setting displays superior performance to the *MILP-PVI* setting, demonstrating that the combination of all the enhancements produces better outcomes in terms of running times, even for the small instances, than just the polynomial inequalities.

Furthermore, analyzing the variation of the computation times depending on the instance characteristics, we can observe that all settings have similar trends. Indeed, concerning the depot position, the "a" position, corresponding to the instance center of gravity, is the group of instances that, on average, require more time. This can be explained by considering that a baricentric position of the depot does not address the possible vehicle routes compared to a more specific position. Regarding the drone speed, the most complex instances are those in which the drone has a speed of 25 km/h. It should be noted that for all these instances, the truck speed is always 25 km/h. Therefore, instances where the two vehicles have similar speeds are more complex, as they cannot prefer one vehicle over the other for deliveries. Finally, regarding the last criterion, a higher endurance corresponds to higher computation times since the solution space is larger.

Table 5.3 presents a comparison of the different settings in terms of continuous relaxation at the root node. Specifically, we solve the corresponding continuous relaxation for each setting to obtain a lower bound for the problem. The percentage gap at the root node *RGap* is then computed as *(BUB-LB)/LB·100*, where *BUB* is the best upper bound known and *LB* is the lower bound obtained by the continuous relaxation. The columns "Avg RGap" and "Max RGap" show the average and maximum *RGap* for each setting. For the *MILP-PVI* and *B&C* settings, we also compute the percentage difference between the *LB* obtained with these settings and the one obtained with the *MILP* setting. This percentage difference is calculated as *(LB-*LB_{MILP}*)/*LB_{MILP}*·100*, where *LB* is the lower bound computed with the *MILP-PVI* or *B&C* setting, while LB_{MILP} is the lower bound obtained with the *MILP* setting. The average and maximum difference percentage are reported in the columns "Avg Diff" and "Max Diff", respectively.

The results regarding the root node relaxation demonstrate that both the *MILP-PVI* and *B&C* settings lead to an improvement of the lower bound, even on the small instances. Specifically, it can be observed that, on average, the *MILP-PVI* setting improves the lower bound by approximately 1%, with some instances showing improvements of up to 12%. As expected, the *B&C* setting yields even more significant improvements, with an average increase in lower bound of around 31% and a maximum increase of approximately 60%. Furthermore, it is worth noting that, for the *B&C* setting, the average *Rgap* is below 10%, which is lower than the average *Rgap* of the other two settings, and the maximum *Rgap* is even lower than the average *Rgap* of the other settings.

Table 5.3: Impact of valid inequalities and variable fixing strategies on root node relaxation for *M10-instances*

		MILP		MILP - PVI				B&C			
	# Ins	Avg RGap	Max Rgap	Avg Rgap	Max Rgap	Avg Diff	Max Diff	Avg Rgap	Max Rgap	Avg Diff	Max Diff
Dep = a	18	32.49	40.19	31.86	40.19	0.99	11.95	10.60	21.12	32.90	54.66
Dep = b	18	33.52	43.88	33.01	43.88	0.83	9.13	9.48	24.32	37.23	60.31
Dep = c	18	29.17	42.87	28.38	42.87	1.20	11.42	7.22	24.30	32.10	59.55
Dep = d	18	23.96	35.56	23.07	35.56	1.20	5.49	6.18	16.32	23.96	40.76
v_d = 15	24	31.30	42.87	29.97	42.87	2.07	11.95	8.43	24.30	33.99	59.55
v_d = 25	24	30.52	43.88	30.09	43.88	0.57	5.47	10.99	24.32	28.96	52.39
v_d =35	24	27.55	37.87	27.18	37.87	0.51	4.66	5.69	14.90	31.69	60.31
Dtl = 20	36	30.59	43.88	29.74	43.88	1.35	11.95	7.51	24.32	34.29	59.87
Dtl = 40	36	28.98	42.87	28.42	42.87	0.75	5.47	9.23	24.30	28.81	60.31
All	72	29.79	43.88	29.08	43.88	1.05	11.95	8.37	24.32	31.55	60.31

Results on *M20-instances*

Table 5.4 presents a comparison of the different settings in terms of number of instances optimally solved, gap, and running times on the *M20-instances*. The instances are grouped based on two criteria: the position of the depot (Dep) and the drone endurance. The letter "C" indicates that the depot is located near the center of gravity of the customers, while "E" and "O" represent the average of the customer x-coordinate with y-coordinate of zero and the southwest corner of the region (0,0), respectively. The column "# Ins" indicates the number of instances in each group. For each setting, we report the number of instances optimally solved within the time limit ("# Opt"), the average and maximum percentage gap ("Avg Gap" and "Max Gap"), and the average running time ("Avg time"). The percentage gap is calculated as *(UB-LB)/LB·100*, where *UB* and *LB* represent the upper and lower bounds obtained by the different settings.

The results indicate that the running times of the two settings incorporating the proposed enhancements (*MILP-PVI* and *B&C*) are significantly lower than those of the basic setting (*MILP*). Specifically, the average running times of *MILP-PVI* and *B&C* are about 60% and 50% of the average running time of the *MILP* setting, respectively. This is due to the fact that while *MILP* setting can optimally solve 226 out of 240 instances, the *MILP-PVI* and *B&C* settings solve 232 and 235 instances within the time limit, respectively. Furthermore, all the settings have an average gap lower than 0.30%. However, the *MILP-PVI* and *B&C* settings have an average gap that is half and one-third of the *MILP* setting, respectively. Finally, with regard to the maximum percentage gap, it can be observed that *MILP-PVI* yields a slight decrease in the maximum percentage gap compared to the *MILP* setting. On the other hand, the *B&C* setting reduces the maximum percentage gap by about 4%.

Analyzing the variation in the number of instances optimally solved, percentage gap, and computation times depending on the instance characteristics, we can observe that all settings exhibit similar trends on the *M20-instances*. Regarding the depot position, the "E" position, which corresponds to the bottom side of the instance, requires, on average, more time and has the lowest number of instances optimally solved. However, it is worth noting that only 8 instances out of 80 are not solved by the *MILP* setting, while the *B&C* setting is unable to solve just 4 instances of this group. As for the drone endurance, a higher endurance leads to higher computation times and a larger number of instances that are not optimally solved within the time limit, similar to the previous instances. However, unlike the depot position, the drone endurance has a more significant impact on the performance of all settings.

Table 5.4: Impact of valid inequalities and variable fixing strategies on number of instances optimally solved, gap, and running times for *M20-instances*

		MILP				MILP - PVI				B&C			
	# Ins	# Opt	Avg Gap	Max Gap	Avg time	# Opt	Avg Gap	Max Gap	Avg time	# Opt	Avg Gap	Max Gap	Avg time
Dep = C	80	77	0.12	4.15	466.87	79	0.02	1.80	270.73	80	0.00	0.00	190.31
Dep = E	80	72	0.58	14.33	891.21	75	0.36	13.98	544.87	76	0.28	10.30	441.54
Dep = O	80	77	0.12	4.03	541.46	78	0.04	2.19	380.78	79	0.01	0.91	287.49
Dtl = 20	120	117	0.09	4.43	265.58	119	0.05	6.47	93.91	119	0.04	5.31	77.43
Dtl = 40	120	109	0.46	14.33	980.11	113	0.23	13.98	703.68	116	0.15	10.3	535.46
All	240	226	0.27	14.33	622.85	232	0.14	13.98	398.8	235	0.10	10.30	306.44

Table 5.5 compares the different settings in terms of continuous relaxation at the root node for the *M20-instances*. For each setting, we provide the average and maximum percentage root gap, as well as the average and maximum percentage difference.

The root node relaxation results indicate that both the *MILP-PVI* and *B&C* settings provide a significant improvement in the lower bound for the *M20-instances*. On average, the *MILP-PVI* setting leads to a 4% improvement in the lower bound, with some instances showing improvements of up to 18%. Similarly, as observed in the *M10-instances*, the *B&C* setting demonstrates more significant improvements, with an average increase in the lower bound of around 36% and a maximum increase of approximately 97%. It is worth noting that the average *Rgap* for the *B&C* setting is below 20%, which is lower than the average *Rgap* for the other two settings, while the maximum *Rgap* is comparable to the average *Rgap* for the other settings.

Table 5.5: Impact of valid inequalities and variable fixing strategies on root node relaxation for *M20-instances*

		MILP		MILP - PVI				B&C			
	# Ins	Avg RGap	Max Rgap	Avg RGap	Max Rgap	Avg Diff	Max Diff	Avg RGap	Max Rgap	Avg Diff	Max Diff
Dep = C	80	34.66	57.81	31.91	54.83	4.38	14.71	14.78	34.93	32.14	97.60
Dep = E	80	44.09	58.61	41.95	56.44	3.95	17.44	22.87	42.91	39.78	93.22
Dep = O	80	40.17	59.22	38.11	57.66	3.71	16.23	19.12	39.07	39.78	96.13
Dtl = 20	120	38.83	58.61	36.73	56.33	3.67	17.44	16.42	42.91	39.06	96.13
Dtl = 40	120	40.45	59.22	37.91	57.66	4.36	15.08	21.43	39.07	34.01	97.60
All	240	39.64	59.22	37.32	57.66	4.02	17.44	18.92	42.91	36.54	97.60

5.5.2 Comparison with the state-of-the-art solvers for the *FS-TSP*

The results of the comparison with state-of-the-art solution approaches are presented in the following three sub-sections, one for each of the test beds described in Section 5.2. Specifically, the performance of the proposed *B&C* method is evaluated by comparing it with the state-of-the-art solution approaches for each test bed. We would like to note that the results are presented in an aggregated form for the sake of readability. Nevertheless, detailed results can be found in the Appendix for those interested.

The state-of-the-art solution approaches for the *M10-instances* and the *M20-instances* are represented by the Column-and-Row Generation method proposed in [47] and the Branch-and-Cut method proposed in [83], which we refer to as *BMS21* and *DMN22*, respectively. On the other hand, the state-of-the-art approach for the *FS-TSP* on the *POI-instances* is represented by the solution method proposed in [247], which we refer to as *RR21*.

It is important to note that the results related to *BMS21*, *DMN22*, and *RR21* are taken from [47], [83], and [246], respectively. Therefore, a direct comparison in terms of running time is not possible since the different solution methods were tested on different computing environments. However, it is worth mentioning that the CPU used in this work is similar to the one used in *RR21*, and slightly better than those used in *DMN22* and *BMS21*, as indicated by a study reported in [212]. This study ranks processors based on single-thread performance, with a range of scores from 77 to 126,045. In particular, the marks of the CPU used for the *B&C*, *DMN22*, *BMS21*, and *RR21* are 12958, 1843, 3275, and 11490, respectively.

Results on *M10-instances*

All the methods considered for the *M10-instances* were able to obtain the optimal solution for all the instances in this set. Thus, the only way to compare their performance is in terms of running times. To this end, we computed the speed-up related to the *B&C* approach in comparison to *DMN22* and *BMS21*. The speed-up is computed by calculating the ratio of the running time of the literature method and our running time. A speed-up lower than 1 indicates that the *B&C* approach is slower than the literature method, while a speed-up greater than 1 means that it results in the same running time or is faster.

Table 5.6 reports the results of this comparison. We grouped the instances based on the depot location, drone speed, and drone endurance. For each group, we indicate the number of instances ("# Ins"). Then, for each liter-

Table 5.6: Comparison of the speed-up with respect to *DMN22* and *BMS21*

		Speed-up B&C vs DMN21				Speed-up B&C vs vs BMS21			
	# Ins	# Faster	Average	Min	Max	# Faster	Average	Min	Max
Dep = a	18	18	640.46	1.28	6448.85	18	190.39	2.03	1254.99
Dep = b	18	18	119.43	1.10	691.19	18	71.28	1.64	383.96
Dep = c	18	17	33.01	0.48	172.50	18	40.20	1.82	260.01
Dep = d	18	18	28.71	1.57	102.26	17	42.36	0.68	220.85
s_d = 15	24	24	153.34	1.10	1240.90	23	166.44	0.68	1254.99
s_d = 25	24	23	357.05	0.48	6448.85	24	69.63	2.58	592.12
s_d = 35	24	24	105.82	1.74	1112.39	24	22.10	1.43	108.46
Dtl = 20	36	36	36.81	1.10	287.46	35	24.44	0.68	195.81
Dtl = 40	36	35	374.00	0.48	6448.85	36	147.68	1.43	1254.99
All	72	71	205.40	0.48	6448.85	71	86.06	0.68	1254.99

ature method, we indicate the number of instances that our *B&C* approach solves faster ("# Faster"), the average, the minimum, and the maximum speed-up achieved.

The results show that the *B&C* is significantly faster than the two literature approaches. In fact, we can observe that on average with respect to *DMN22*, the *B&C* shows a speed-up of about 200. Furthermore, we can also observe that the *B&C* is slower than *DMN22* only on one instance, where *DMN22* is twice as fast. However, on the instance where *DMN22* performs worse, the *B&C* is about 6000 times faster. Similar considerations can be made for the comparison with *BMS21*. In summary, we can observe that the average speed-up is about 80 while the maximum speed-up is about 1200. In conclusion, although a direct comparison is not possible due to the different environments used for the tests, the speed-up values are so high that they cannot be justified only by the different computing power involved, demonstrating the effectiveness of the proposed solution method.

Results on *M20-instances*

The state-of-the-art solution approaches are not capable of obtaining the optimal solution for all the *M20-instances* within the given time limit of 1 hour. Furthermore, the running times of *DMN22* are not available for this set of instances, making it impossible to compare the running times of the three solution methods. Hence, we assess the methods based on the percentage optimality gap. Specifically, we provide the number of optimal solutions achieved, along with the average and maximum percentage gap of the two literature approaches. For each instance, we select the best result among *DMN22* and *BMS21*. To avoid redundancy, we do not report again the gap for our *B&C*

Table 5.7: Comparison of quality of the solution with respect to *DMN22* and *BMS21*

		DMN22/BMS21			B&C vs DMN22/BMS21		
	# Ins	#O pt	Avg Gap	Max Gap	# Worse	# Equal	# Better
Dep = C	80	27	10.42	39.00	0	27	53
Dep = E	80	20	11.24	45.63	1	20	59
Dep = O	80	41	10.15	49.66	0	41	39
Dtl = 20	120	82	2.54	23.99	1	82	37
Dtl = 40	120	6	18.67	49.66	0	6	114
All	240	88	10.60	49.66	1	88	151

present the number of instances where our *B&C* approach exhibits a higher, equal, or lower gap in the columns "# Worse", "# Equal", "# Better", respectively.

The results show how, even considering the best result between the two state-of-the-art approaches, the results of our *B&C* approach are significantly better. In fact, we can observe that the average percentage gap is 10.60, which is higher than the maximum gap of our *B&C* approach. Moreover, by comparing the instances where the *B&C* approach obtains a lower gap, it can be noticed that in 151 out of 240 instances, our *B&C* approach outperforms the literature approaches, which is about 60% of the instances. Furthermore, it can be observed that only in one instance, the literature approaches perform better than our *B&C* approach. Finally, considering the different characteristics of the instances, it can be observed that the main improvement is related to the ability of the *B&C* approach to solve instances with a drone endurance of 40. In fact, in 114 out of 120 instances, the *B&C* approach obtains a lower gap.

Results on *POI-instances*

The *POI-instances* can be grouped on the basis of two different features: the number of customer and the ratio between the drone and the truck speed (α). Table 5.8 shows the comparison between *RR21* and the *B&C* on the different subsets of *POI-instances*. For each subset, we report the number of instances. Then, for each solution method we report the number of instances solved to optimality within the time limit, the average and the maximum percentage gap, and the average running time.

The performance comparison of the two methods based on instance characteristics reveals that their performance tends to decrease with an increase in the number of clients and drone speed. In the former case, the instance size increases, while in the latter case, the solution space grows as more clients can

Table 5.8: Comparison with *RR21* in terms of number of instances optimally solved, gap, and running time.

		RR21				B&C			
	#Ins	# Opt	Avg Gap	Max Gap	Avg Time	# Opt	Avg Gap	Max Gap	Avg Time
$\lvert C \rvert = 9$	75	75	0.00	0.00	0.13	75	0.00	0.00	0.18
$\lvert C \rvert = 19$	75	75	0.00	0.00	3.18	75	0.00	0.00	6.07
$\lvert C \rvert = 29$	75	75	0.00	0.00	92.29	74	0.03	2.38	198.94
$\lvert C \rvert = 39$	75	61	5.48	51.01	1475.88	57	0.58	6.22	1482.69
$\alpha = 1$	100	94	0.88	33.57	405.01	97	0.05	1.85	238.75
$\alpha = 2$	100	98	0.82	43.51	279.89	94	0.16	6.22	424.72
$\alpha = 3$	100	94	2.42	51.01	493.70	90	0.25	5.43	602.43
All	300	286	1.37	51.01	392.87	281	0.15	6.22	421.97

Table 5.9: Results on *POI-instances* unsolved by *RR21*

	RR21		*B&C*	
Id	UB	% Gap	UB	%Gap
poi-40-11-1	266	0.75	266	0.00
poi-40-13-3	363	42.15	210	0.95
poi-40-15-3	346	33.53	231	4.33
poi-40-17-3	359	32.59	242	0.00
poi-40-18-1	391	25.32	297	1.68
poi-40-22-3	339	34.22	224	0.00
poi-40-23-1	395	26.84	302	0.00
poi-40-23-2	395	38.48	244	0.00
poi-40-25-1	417	33.57	281	0.00
poi-40-3-2	416	43.51	235	0.00
poi-40-3-3	416	48.08	216	3.24
poi-40-4-1	284	0.35	288	0.00
poi-40-7-1	271	0.74	271	0.00
poi-40-7-3	396	51.01	194	1.55

be served by the drone since the endurance is fixed at 20 for all speeds. Furthermore, the results demonstrate that the two approaches yield similar results in terms of the number of optimal solutions. Specifically, *RR21* solves 286 out of 300 instances, while the *B&C* solves 281. However, the *B&C* generally produces significantly lower gaps than *RR21* in terms of percentage. In particular, the average and maximum gaps are 0.15% and 6.22%, respectively, for the *B&C*, while they are 1.37% and 51.01% for *RR21*.

Based on these results, we can conclude that the proposed *B&C* is more robust than *RR21*. Although the *B&C* solves, on average, slightly fewer instances to optimality than *RR21*, the instances that the *B&C* is unable to solve have significantly lower gaps compared to *RR21*, which presents particularly high gaps for unsolved instances. To further support this conclusion, we reported a detailed comparison on the unsolved instances by *RR21* and *B&C* in Tables 5.9 and 5.10, respectively.

From Table 5.9, we can observe that, on the instances unsolved by *RR21*, its average gap is equal to 29.37% while our *B&C* determines an average gap

Table 5.10: Results on *POI-instances* unsolved by *B&C*

	RR21		*B&C*	
Id	UB	% Gap	UB	%Gap
poi-30-20-3	168	0.00	168	2.38
poi-40-1-2	213	0.00	214	0.93
poi-40-13-3	363	42.15	210	0.95
poi-40-15-1	279	0.00	279	1.43
poi-40-15-3	346	33.53	231	4.33
poi-40-16-1	270	0.00	270	1.85
poi-40-16-2	210	0.00	210	0.95
poi-40-18-1	391	25.32	297	1.68
poi-40-18-2	239	0.00	241	6.22
poi-40-18-3	221	0.00	221	5.43
poi-40-21-3	218	0.00	218	1.38
poi-40-24-2	215	0.00	215	2.33
poi-40-25-3	190	0.00	190	1.05
poi-40-3-3	416	48.08	216	3.24
poi-40-4-2	234	0.00	234	1.71
poi-40-5-3	203	0.00	205	3.41
poi-40-6-2	192	0.00	192	3.65
poi-40-6-3	181	0.00	181	1.10
poi-40-7-3	396	51.01	194	1.55

of 0.84%. On the other hand, as shown in Table 5.10, on the instances unsolved by the *B&C*, its average gap is 2.40% while the one of *RR21* is 10.53%. Therefore, on average, our *B&C* performs significantly better than *RR21* on both subsets of unsolved instances.

5.6 Conclusions

In this Chapter, we propose an original *MILP* formulation for the *FS-TSP*, which unlike the ones present in the literature, uses a polynomial number of variables and does not use *big-M* constraints.

We solve the proposed formulation through a *B&C* approach integrated with drone endurance based variable fixing strategies and four different families of valid inequalities based on the vehicle coordination.

Computational results on different sets of benchmark instances confirm the robustness of the proposed *MILP* formulation and the effectiveness of the *B&C*. Indeed, the experimentation shows that the proposed method is competitive or outperforms the state-of-the-art approaches, providing either the optimal solution or improved bounds for several instances unsolved before.

Future research directions naturally include the extension of the proposed formulation to different variants of the *FS-TSP*. Moreover, we are interested in developing a matheuristic framework that includes the proposed formulation to solve larger and harder instances. We will investigate the possibility of combining the proposed solution approach with machine learning and data science techniques to be able to compute good solutions with a limited computational burden. Finally, we plan to apply this methodology to a healthcare-related topic, which, as highlighted at the beginning of the chapter, is already widely benefiting from the use of drones. This application aims to further explore and enhance the potential of drones in critical sectors such as healthcare, where timely and efficient delivery can have significant impacts.

Conclusions

This thesis presents a comprehensive investigation into the application of Operational Research in the healthcare sector, highlighting its critical role in enhancing decision-making processes, patient outcomes, and system efficiency. Across its chapters, the thesis addresses specific Operational Research methodologies and their practical applications, presenting innovative solutions to healthcare challenges.

In Chapter 1, we introduced the link between Operational Research and healthcare, providing an overview of the significant contributions made in this context. This chapter highlighted the key areas of Location Analysis, Scheduling Optimization, and Routing, setting the context for deeper investigations in later chapters. Through this introduction, we arrange the landscape of Operational Research applications within healthcare, showcasing how these methodologies have been effectively employed to address complex healthcare challenges.

In Chapter 2, we explored the strategic management of regional Blood Supply Chain, an essential issue of healthcare logistics. The chapter introduced two MILP formulations aimed at balancing the need for donor self-sufficiency with the requirement to minimize system management costs. This chapter is focused on a strategic problem, offering valuable insights into policy implications and illustrating the role of Operational Research in guiding high-level decision-making. The practical application and relevance of these models is underscored by the validation in real-world regional contexts, affirming their practicality and significance in the field.

Chapters 3 and 4 presented an in-depth analysis of a *IORPS* faced in a hospital in Naples. These chapters introduced original ILP formulations aimed at optimizing surgical schedules while accommodating emergencies and unforeseen surgery delays. The flexibility and robustness of these models were demonstrated through various experiments, highlighting their potential utility in other hospital settings. These chapters not only addressed operational efficiency but also underlines the potential integration of OR tools within broader digitalization efforts in healthcare, recognizing the interplay between technology and operational optimization.

Chapter 5 mark a shift towards drone logistics, an emerging field with profound implications for healthcare delivery. The novel MILP formulation presented for the FS-TSP, distinct in its polynomial number of variables and avoidance of big-M constraints, represented a significant advancement in the field. The robust computational results, outperforming existing approaches or achieving equivalent outcomes, illustrate the effectiveness of the proposed method. This chapter not only contributed a novel OR methodology but also opened avenues for future research in healthcare logistics, particularly in the application of drones for timely and efficient delivery.

Overall, this thesis traverses the multi-faceted landscape of OR applications in healthcare, from strategic planning and policy formulation to operational and logistical optimizations. Each chapter contributes valuable methodologies, insights, and potential applications, underlining the fundamental role of OR in transforming healthcare processes.

Appendix

This appendix includes, for Chapter 5, the comprehensive results of the comparison between the proposed *B&C* method and the state-of-the-art approaches for the *FS-TSP*.

M10-instances
Tables 11 and 12 report the results for the instances with a *Dtl* equal to 20 and 40 minutes, respectively. For each instance, we report the objective function value of the optimal solution and the running times of each solution method.

Concerning the results on the instances with *Dtl* equal to 20 minutes, we can observe that the proposed *B&C* is faster than the other two methods. Indeed, it shows an average computation time of 1.34 seconds. On the other hand, the other two approaches are more than ten times slower than the proposed one, with an average computation time equal to 45.27 and 75.56 seconds for *BMS21* and *DMN22*, respectively. Moreover, our approach performs better than the others, even in the worst-case scenario, i.e., by considering the largest computation time required to solve an instance. In particular, the largest computation time for the *B&C* is 3.42 seconds versus the 329.71 seconds for *BMS21* and the 707.32 seconds for *DMN22*.

Concerning, instead, the instances with Dtl equal to 40 minutes, we can observe that the performance of the *B&C* seems to be unaffected by the larger Dtl. Indeed, the average and the largest computation times are similar to the previous ones. In particular, the average and the largest computation times are slightly larger (2.42 and 4.74 seconds, respectively). Instead, the performance of the other two solution methods decreases when the Dtl increases. Indeed, on the one hand, the average and the largest computation times for *BMS21* are 361.08 and 2762.93 seconds, respectively. On the other hand, *DMN22* shows an average computation time of 975.44 and a largest computation time of about 4 hours. In conclusion, the *B&C* outperforms the state-of-the-art algorithms on the *M10-instances*, being faster than the other approaches by

Table 11: Results on *M10-instances* instances with $Dtl = 20$

Id	OPT	DMN22 time	BMS21 time	B&C time	Id	OPT	DMN22 time	BMS21 time	B&C time	Id	OPT	DMN22 time	BMS21 time	B&C time
M37V1	57.45	1.23	3.80	0.96	M40V1	49.43	4.03	12.14	0.81	M43V1	69.59	0.53	0.3	0.15
M37V2	53.79	0.79	1.96	0.72	M40V2	51.71	12.09	27.93	0.84	M43V2	72.15	0.58	0.3	0.18
M37V3	54.66	6.21	2.78	0.41	M40V3	57.1	6.9	9.11	0.91	M43V3	77.34	0.43	0.24	0.12
M37V4	67.46	5.09	1.03	0.33	M40V4	69.9	1.98	11.26	0.80	M43V4	90.14	0.53	0.23	0.34
M37V5	51.78	406.43	190.25	3.42	M40V5	45.46	132.01	70.81	2.45	M43V5	58.71	282.87	192.68	0.98
M37V6	48.6	128.31	80.29	2.65	M40V6	44.51	9.48	35.44	1.31	M43V6	59.09	69.65	55.03	1.54
M37V7	49.58	18.38	32.77	2.36	M40V7	49.9	5.27	7.78	1.15	M43V7	65.52	8.98	7.7	1.24
M37V8	62.38	43.12	35.87	1.99	M40V8	62.7	5.84	8.97	0.93	M43V8	84.81	103.32	20.98	1.66
M37V9	43.48	361.49	329.71	3.04	M40V9	42.53	14.47	9.57	1.25	M43V9	46.93	707.32	239.28	2.94
M37V10	41.91	211.62	170.37	2.03	M40V10	43.08	3.64	3.78	2.05	M43V10	47.93	72.36	51.35	2.69
M37V11	42.9	37.47	16.02	1.45	M40V11	49.2	1.19	1.26	0.68	M43V11	57.38	11.91	5.52	0.85
M37V12	56.85	38.31	25.85	1.44	M40V12	62	1.12	1.13	0.62	M43V12	69.2	5.28	2.09	0.78

about a couple of orders of magnitude.

M20-instances

The two literature algorithms are not able to solve to optimality several instances even with *Dtl* equal to 20. Therefore, we report the best upper bound and best percentage optimality gap among the two methods for each instance. Then, we report the upper bound, the percentage gap, and the computation time obtained by our $B\&C$. The percentage optimality gap is calculated as $\%Gap = (UB - LB)/LB \cdot 100$. We highlight that we do not report a detailed comparison of the running times of the different solution methods since the running times are not available in [83] for *DMN22*.

The results on the *M20-instances* with *Dtl* equal to 20 and 40 minutes are reported in Tables 13 and 14, respectively.

Table 12: Results on *M10-instances* with $Dtl = 40$

Id	OPT	DMN22 time	BMS21 time	B&C time	Id	OPT	DMN22 time	BMS21 time	B&C time	Id	OPT	DMN22 time	BMS21 time	B&C time
M37V1	50.50	2731.91	2762.93	2.20	M40V1	46.89	252.66	365.46	3.16	M43V1	57.01	2312.62	1643.65	2.35
M37V2	47.31	284.46	550.88	1.43	M40V2	46.42	70.15	123.30	2.31	M43V2	58.05	1446.23	792.02	2.24
M37V3	53.69	190.49	618.77	2.38	M40V3	53.93	150.67	502.94	3.20	M43V3	69.40	219.80	222.88	2.33
M37V4	67.46	148.43	568.48	2.57	M40V4	68.40	252.36	608.64	3.72	M43V4	83.70	224.88	397.32	2.82
M37V5	45.84	644.23	369.05	2.75	M40V5	43.53	53.75	48.96	2.62	M43V5	52.09	15702.75	1441.79	2.43
M37V6	44.60	253.15	178.55	3.17	M40V6	44.08	40.61	13.23	1.92	M43V6	52.33	1921.06	374.47	2.78
M37V7	47.62	267.10	349.93	3.30	M40V7	49.23	1.13	6.05	2.35	M43V7	61.88	452.45	77.90	2.62
M37V8	60.42	241.61	299.12	2.36	M40V8	62.03	6.76	5.62	1.83	M43V8	73.73	123.31	56.58	2.86
M37V9	42.42	1167.62	62.31	2.08	M40V9	42.53	5.50	7.63	1.25	M43V9	46.93	5271.24	362.06	4.74
M37V10	41.91	406.15	76.01	2.22	M40V10	43.08	5.53	4.94	2.34	M43V10	47.93	169.56	82.85	3.61
M37V11	42.90	51.09	9.15	2.44	M40V11	49.20	3.38	1.70	0.80	M43V11	56.40	11.37	3.30	1.82
M37V12	55.70	26.62	6.69	2.20	M40V12	62.00	1.16	0.88	0.62	M43V12	69.20	4.21	2.94	1.17

Table 13: Results on *M20-instances* with $Dtl = 20$

	DMN22/BMS21		*B&C*				*DMN22/BMS21*		*B&C*				*DMN22/BMS21*		*B&C*		
Id	Best UB	Best % Gap	UB	%Gap	Time	Id	Best UB	Best % Gap	UB	% Gap	Time	Id	Best UB	Best % Gap	UB	%Gap	Time
4847	267.05	0.00	267.05	0.00	5.39	5025	131.43	0.00	131.43	0.00	9.87	5154	123.34	9.84	121.90	0.00	49.95
4849	248.30	0.00	248.30	0.00	0.84	5027	114.31	0.00	114.31	0.00	14.32	5156	124.46	0.00	124.46	0.00	21.18
4853	232.87	0.00	232.87	0.00	1.01	5030	116.1	0.00	116.10	0.00	44.69	5159	145.79	0.00	145.79	0.00	18.46
4856	253.33	0.00	253.33	0.00	1.47	5032	117.55	1.88	117.55	0.00	13.28	5201	148.02	0.00	148.02	0.00	9.25
4858	240.63	0.00	240.63	0.00	1.39	5034	105.10	1.18	105.10	0.00	42.98	5203	138.59	0.00	138.59	0.00	21.24
4902	242.32	0.00	242.32	0.00	1.25	5036	124.33	3.65	124.33	0.00	22.99	5205	134.78	0.00	134.78	0.00	49.77
4907	239.28	0.00	239.28	0.00	0.91	5039	130.91	3.67	130.91	0.00	7.50	5207	121.47	0.00	121.47	0.00	23.18
4909	222.88	0.00	222.88	0.00	1.15	5041	125.32	7.77	125.32	0.00	10.83	5209	135.92	0.00	135.92	0.00	15.72
4912	267.62	0.00	267.62	0.00	0.87	5044	121.87	11.51	120.77	0.00	25.80	5212	137.67	0.00	137.67	0.00	18.77
4915	259.39	0.00	259.39	0.00	0.66	5047	112.85	0.00	112.85	0.00	9.61	5214	126.25	0.00	126.25	0.00	27.93
4917	173.97	0.00	173.97	0.00	3.60	5049	197.66	0.00	197.76	0.00	3.31	5216	101.07	18.27	101.07	0.00	46.56
4920	170.05	0.00	170.05	0.00	4.92	5051	180.62	0.00	180.62	0.00	9.92	5218	115.82	4.14	115.62	0.00	27.93
4922	169.60	0.00	169.60	0.00	2.66	5053	176.51	0.00	176.51	0.00	2.01	5220	119.04	0.00	119.04	0.00	28.49
4924	159.57	0.00	159.57	0.00	2.56	5055	177.29	0.00	177.29	0.00	4.23	5223	94.59	3.64	94.59	0.00	42.48
4926	155.98	0.00	155.98	0.00	13.54	5057	180.7	0.00	180.70	0.00	6.02	5225	129.72	0.00	129.72	0.00	12.36
4928	166	0.00	166.00	0.00	18.64	5059	150.82	0.00	150.82	0.00	5.64	5227	116.19	0.00	116.19	0.00	75.77
4931	172.49	0.00	172.49	0.00	14.91	5102	165.49	0.00	165.49	0.00	5.24	5229	94.26	13.69	92.87	0.00	70.17
4933	159.39	0.00	159.39	0.00	13.83	5104	181.61	0.00	181.61	0.00	4.93	5231	98.93	0.00	98.36	0.00	38.46
4935	176.69	0.00	176.69	0.00	18.86	5106	158.49	0.00	158.49	0.00	12.78	5233	111.62	7.85	111.62	0.00	29.04
4937	173.55	0.00	172.50	0.00	5.51	5108	172.12	0.00	172.12	0.00	1.96	5235	118.89	0.00	118.89	0.00	12.20
4939	201.03	0.00	201.03	0.00	5.34	5110	135.43	0.00	135.43	0.00	17.93	5238	79.39	10.33	78.93	0.00	562.17
4941	253.08	0.00	253.08	0.00	1.27	5112	131.23	0.00	131.23	0.00	13.20	5240	87.46	11.84	84.51	0.00	183.79
4944	247.03	0.00	247.03	0.00	2.98	5115	127.39	0.00	127.39	0.00	19.19	5242	85.65	3.25	85.65	0.00	10.29
4946	237.21	0.00	237.21	0.00	1.36	5117	130.36	10.31	130.36	0.00	46.49	5244	86.81	0.00	86.81	0.00	33.92
4948	258.06	0.00	258.06	0.00	1.38	5119	118.97	0.00	118.97	0.00	28.65	5246	74.56	5.12	74.19	5.31	3600
4950	240.99	0.00	240.99	0.00	1.44	5121	131.84	8.90	131.19	0.00	25.25	5248	83.1	23.36	83.10	0.00	1760.19
4952	218.09	0.00	218.09	0.00	4.77	5123	121.95	1.61	121.95	0.00	16.46	5250	81.93	0.00	81.93	0.00	42.83
4954	261.06	0.00	261.06	0.00	1.40	5125	130.96	7.38	130.96	0.00	13.52	5252	86.38	2.42	86.15	0.00	110.90
4957	252.28	0.00	252.28	0.00	1.20	5127	132.24	0.00	132.24	0.00	7.97	5255	82.5	2.83	82.50	0.00	89.69
4959	249.92	0.00	249.92	0.00	1.11	5130	126.49	9.37	126.49	0.00	45.19	5257	79.43	23.99	78.27	0.00	329.14
5001	115.5	0.00	115.50	0.00	17.58	5132	106.36	14.92	106.36	0.00	50.90	5306	100.46	0.00	100.46	0.00	19.14
5003	173.54	3.47	172.73	0.00	8.03	5134	102.57	0.76	102.57	0.00	21.86	5310	92.41	0.00	92.41	0.00	47.10
5006	155.39	0.00	155.39	0.00	5.90	5136	98.91	0.00	98.91	0.00	32.02	5312	83.59	0.00	83.59	0.00	20.19
5008	159.74	0.00	159.74	0.00	5.87	5138	91.83	3.81	91.61	0.00	73.10	5321	101.49	0.00	101.49	0.00	11.89
5010	145.48	0.00	145.48	0.00	15.00	5141	95.53	16.82	95.53	0.00	68.49	5324	104.03	8.98	101.55	0.00	77.96
5012	172.4	0.00	171.38	0.00	8.02	5143	97.69	0.00	97.69	0.00	26.96	5330	118.45	0.00	118.45	0.00	7.55
5015	172.67	0.00	172.67	0.00	3.43	5145	95.09	5.89	95.09	0.00	43.67	5334	102.02	0.00	102.02	0.00	26.40
5017	166.47	7.13	166.47	0.00	8.86	5148	90.58	7.12	90.58	0.00	52.35	5336	104.46	0.00	104.46	0.00	12.53
5020	155.92	0.00	155.92	0.00	3.46	5150	82.19	9.99	82.19	0.00	172.37	5345	114.19	0.00	114.19	0.00	7.88
5022	146.21	3.41	146.21	0.00	13.71	5152	90.58	14.38	90.58	0.00	384.83	5351	115.21	0.44	115.10	0.00	53.96

As shown in Table 13, the two literature approaches can solve only 82 out of 120 instances to optimality. Instead, our approach solves to optimality almost all the instances (119 out of 120). Clearly, our approach shows better performance also in terms of average and maximum optimality gap. Indeed, our optimality gap is equal to 0.04% versus the 2.54% of *BMS21*/*DMN22*. Furthermore, the maximum gap obtained by our *B&C* and the other two methods are 5.31% and 23.99%, respectively. Moreover, considering the 37 instances solved to optimality for the first time in literature by our *B&C*, we can observe that our algorithm determined a new upper bound on 13 of them and proved the optimality of the previously known upper bound on the other 24. For the sake of completeness, we point out that our approach is faster than *BMS21* also on this set of instances with an average running time equal to 77.43 and 1765.25 seconds, respectively.

Table 14: Results on *M20-instances* with $Dtl = 40$

	DMN22/BMS21		*B&C*				*DMN22/BMS21*		*B&C*				*DMN22/BMS21*		*B&C*		
Id	Best UB	Best % Gap	UB	%Gap	Time	Id	Best UB	Best % Gap	UB	% Gap	Time	Id	Best UB	Best % Gap	UB	%Gap	Time
4847	255.60	10.70	255.60	0.00	51.14	5025	119.20	32.26	118.43	0.00	533.62	5154	106.64	20.26	104.28	0.00	296.76
4849	225.15	22.37	225.15	0.00	108.81	5027	112.15	11.51	111.81	0.00	333.10	5156	108.92	16.20	107.12	0.00	120.46
4853	218.60	18.68	211.38	0.00	115.55	5030	102.69	10.35	100.64	0.00	3490.25	5159	125.41	17.76	120.02	0.00	158.89
4856	237.03	21.12	234.63	0.00	113.71	5032	106.69	12.01	103.06	0.00	204.23	5201	140.30	19.87	140.30	0.00	813.36
4858	215.63	13.81	215.63	0.00	52.83	5034	102.63	0.71	102.16	0.00	168.37	5203	124.20	16.32	124.06	0.00	359.09
4902	226.62	23.27	223.84	0.00	48.92	5036	112.30	1.06	112.10	0.00	88.98	5205	119.79	40.06	117.68	0.00	329.91
4907	196.66	17.33	192.87	0.00	41.95	5039	126.96	35.76	115.85	10.30	3600	5207	113.90	30.37	113.90	0.00	318.79
4909	216.04	7.95	216.04	0.00	48.41	5041	114.55	17.12	113.88	0.00	276.09	5209	123.69	49.66	121.24	0.91	3600
4912	238.04	1.83	238.04	0.00	51.64	5044	115.94	15.51	115.09	0.00	510.86	5212	135.82	22.84	133.06	0.00	443.64
4915	234.42	13.26	229.03	0.00	39.29	5047	106.11	20.30	103.86	0.00	354.80	5214	123.03	42.11	121.85	0.00	282.34
4917	165.31	2.17	165.31	0.00	29.56	5049	183.51	20.39	180.71	0.00	1065.02	5216	93.31	14.51	91.94	0.00	146.80
4920	161.89	6.22	157.03	0.00	131.76	5051	165.69	15.32	162.52	0.00	803.47	5218	97.61	21.93	96.06	0.00	161.47
4922	168.10	14.46	164.10	0.00	133.30	5053	145.54	26.26	145.42	0.00	115.44	5220	118.79	16.38	114.93	0.00	171.70
4924	158.98	3.92	158.98	0.00	93.41	5055	172.30	9.87	171.64	0.00	77.22	5223	93.64	22.49	91.22	0.00	503.58
4926	152.24	45.63	151.19	0.00	305.57	5057	173.67	10.18	172.88	0.00	479.58	5225	126.57	7.85	123.58	0.00	84.26
4928	153.88	40.19	146.76	5.00	3600	5059	134.83	14.06	134.83	0.00	261.99	5227	105.30	18.49	101.74	0.00	149.88
4931	166.68	11.64	165.22	0.00	162.10	5102	164.01	19.12	163.77	0.00	124.13	5229	94.26	17.15	88.67	0.00	70.36
4933	157.19	25.17	155.13	0.00	401.80	5104	178.67	17.80	176.70	0.00	104.39	5231	99.17	29.33	94.02	0.00	1583.46
4935	160.97	22.86	155.77	0.00	1860.91	5106	144.37	33.42	141.77	0.00	196.89	5233	108.03	16.81	107.9	0.00	138.11
4937	159.44	8.93	159.18	0.00	490.75	5108	162.21	10.00	162.21	0.00	935.15	5235	98.96	19.80	98.96	0.00	357.60
4939	180.92	14.17	179.10	0.00	2986.07	5110	129.07	22.71	126.20	0.00	1322.61	5238	79.09	17.67	77.99	0.00	281.27
4941	241.90	4.71	241.90	0.00	26.95	5112	129.56	24.30	123.13	0.00	287.31	5240	83.27	31.17	81.02	0.00	134.01
4944	233.75	0.00	233.75	0.00	54.47	5115	126.90	21.09	123.08	0.00	3115.85	5242	85.65	6.77	84.52	0.00	81.70
4946	220.39	0.00	220.39	0.00	75.87	5117	124.10	27.64	118.87	0.00	505.66	5244	85.55	36.27	85.32	0.00	33.78
4948	244.13	0.91	244.13	0.00	18.35	5119	114.18	28.82	112.03	0.00	223.39	5246	63.74	18.28	63.25	0.00	125.96
4950	225.61	0.00	225.61	0.00	28.95	5121	118.22	7.72	118.22	0.00	144.85	5248	83.97	39.64	77.55	0.00	199.20
4952	211.67	6.97	210.67	0.00	346.84	5123	115.40	18.73	113.75	0.00	229.14	5250	81.73	24.78	81.73	0.00	2120.91
4954	240.59	9.46	239.67	0.00	40.83	5125	120.17	25.42	118.86	0.00	592.08	5252	85.33	15.15	85.33	1.94	3600
4957	231.61	9.78	225.85	0.00	88.06	5127	125.95	13.07	124.57	0.00	219.70	5255	72.76	27.81	71.13	0.00	345.29
4959	231.95	0.00	231.95	0.00	36.42	5130	120.53	25.07	119.01	0.00	405.15	5257	79.35	27.36	75.17	0.00	753.53
5001	114.09	4.37	114.09	0.00	129.04	5132	107.47	25.96	102.77	0.00	471.40	5306	72.37	17.01	72.37	0.00	353.28
5003	166.92	5.21	162.39	0.00	40.76	5134	102.13	6.27	101.24	0.00	117.72	5310	91.87	41.22	88.33	0.00	2028.71
5006	135.92	12.90	135.92	0.00	94.63	5136	93.24	0.00	93.24	0.00	111.43	5312	82.20	39.29	82.20	0.00	90.37
5008	155.42	19.72	152.11	0.00	249.34	5138	87.76	18.90	87.53	0.00	645.75	5321	79.60	6.69	79.60	0.00	92.65
5010	135.72	19.88	133.85	0.00	377.62	5141	96.13	22.29	92.44	2.49	592.25	5324	91.68	41.47	79.75	0.00	1108.16
5012	160.20	22.70	154.34	0.00	232.78	5143	97.61	15.01	93.87	0.00	463.49	5330	96.35	31.96	96.35	0.00	632.45
5015	158.97	15.27	154.38	0.00	300.89	5145	88.13	0.00	88.13	0.00	71.65	5334	89.68	7.45	88.89	0.00	98.94
5017	157.33	18.69	156.18	0.00	189.40	5148	91.54	9.25	865.29	0.00	128.27	5336	87.41	43.82	86.39	0.00	3080.09
5020	146.01	26.95	141.92	0.00	146.23	5150	77.99	9.01	77.42	0.00	160.82	5345	101.45	21.97	99.63	0.00	83.02
5022	133.36	39.00	128.16	0.00	1657.48	5152	80.82	23.49	79.17	0.00	163.00	5351	96.86	44.29	90.10	0.00	221.93

Considering the results on the instances with *Dtl* equal to 40 minutes, we can observe that the state-of-the-art algorithms can optimally solve just 6 instances out of 120 with an average and maximum optimality gap equal to 18.66% and 49.66%, respectively. On the other hand, the performance of the *B&C* seems again to be not significantly affected by the larger *Dtl*. Indeed, it is able to solve 115 instances with an average and maximum optimality gap equal to 0.15% and 10.30%, respectively.

Concerning the 109 instances solved to optimality for the first time in literature by our *B&C*, we can observe that it determined a new upper bound for 87 instances and proved the optimality of the previously known upper bound for the other 22 . Finally, concerning the 5 instances that are still open, the *B&C* improves the best known upper bound for almost all of them (4 out of 5).

In terms of computation times, a comparison with *BMS21* is possible but not significant since it is not able to solve almost all the instances within the time limit of 1 hour. Nevertheless, for the sake of completeness, we highlight that the average computation times for *BMS21* is 3509.08 seconds. On the other hand, the average computation times of our *B&C* is 535.46 seconds.

Based on these results, it is clear that the proposed *B&C* could be considered the new state-of-the-art algorithm also for the *M20-instances*.

POI-instances

The *POI-instances* can be grouped into four different subsets, of 75 instances each, on the basis of the number of customers (9, 19, 29, and 39 customers). Tables 15, 16, 17, and 18 report the detailed results of the comparison of our method with *RR21* on the four different subsets. To be more precise, each table corresponds to a subset of instances with the same number of customers. Then, for each instance, each table shows the upper bound, the running time, and the percentage optimality gap corresponding to each solution method.

Concerning the results on the instance with 9 customers reported in Table 15, we can observe that both methods are able to optimally solve all the instances. Moreover, the two solution methods show similar performance also in terms of running times. Indeed, *RR21* and *B&C* show an average running time of 0.13 and 0.18 seconds, respectively.

The two methods behave similarly also on the subset of instances with 19 customers, as can be observed from Table 16. Indeed, both methods get the optimal solution on all the instances with an average running time of 3.18 and

Table 15: Results on *POI-instances* with $|C| = 9$

	RR21			*B&C*				*RR21*			*B&C*		
Id	UB	Time	% Gap	UB	Time	%Gap	Id	UB	Time	% Gap	UB	Time	% Gap
poi-10-1-1	173	0.07	0.00	173	0.13	0.00	poi-10-20-3	128	0.1	0.00	128	0.11	0.00
poi-10-1-2	166	0.07	0.00	166	0.07	0.00	poi-10-21-1	132	0.12	0.00	132	0.06	0.00
poi-10-1-3	166	0.1	0.00	166	0.24	0.00	poi-10-21-2	79	0.22	0.00	79	0.31	0.00
poi-10-10-1	175	0.12	0.00	175	0.08	0.00	poi-10-21-3	78	0.24	0.00	78	0.25	0.00
poi-10-10-2	138	0.16	0.00	138	0.28	0.00	poi-10-22-1	153	0.12	0.00	153	0.13	0.00
poi-10-10-3	110	0.3	0.00	110	0.27	0.00	poi-10-22-2	150	0.14	0.00	150	0.39	0.00
poi-10-11-1	169	0.09	0.00	169	0.10	0.00	poi-10-22-3	147	0.15	0.00	147	0.18	0.00
poi-10-11-2	154	0.09	0.00	154	0.08	0.00	poi-10-23-1	182	0.12	0.00	182	0.06	0.00
poi-10-11-3	140	0.15	0.00	140	0.13	0.00	poi-10-23-2	182	0.13	0.00	182	0.07	0.00
poi-10-12-1	182	0.1	0.00	182	0.08	0.00	poi-10-23-3	182	0.15	0.00	182	0.12	0.00
poi-10-12-2	171	0.13	0.00	171	0.28	0.00	poi-10-24-1	168	0.09	0.00	168	0.09	0.00
poi-10-12-3	169	0.12	0.00	169	0.37	0.00	poi-10-24-2	122	0.12	0.00	122	0.23	0.00
poi-10-13-1	170	0.1	0.00	170	0.09	0.00	poi-10-24-3	118	0.12	0.00	118	0.25	0.00
poi-10-13-2	149	0.1	0.00	149	0.22	0.00	poi-10-25-1	169	0.13	0.00	169	0.15	0.00
poi-10-13-3	100	0.19	0.00	100	0.38	0.00	poi-10-25-2	164	0.11	0.00	164	0.34	0.00
poi-10-14-1	162	0.12	0.00	162	0.10	0.00	poi-10-25-3	148	0.1	0.00	148	0.31	0.00
poi-10-14-2	138	0.14	0.00	138	0.22	0.00	poi-10-3-1	178	0.14	0.00	178	0.08	0.00
poi-10-14-3	106	0.17	0.00	106	0.31	0.00	poi-10-3-2	177	0.1	0.00	177	0.14	0.00
poi-10-15-1	195	0.13	0.00	195	0.09	0.00	poi-10-3-3	165	0.08	0.00	165	0.10	0.00
poi-10-15-2	179	0.15	0.00	179	0.16	0.00	poi-10-4-1	178	0.13	0.00	178	0.10	0.00
poi-10-15-3	143	0.14	0.00	143	0.16	0.00	poi-10-4-2	149	0.09	0.00	149	0.29	0.00
poi-10-16-1	174	0.11	0.00	174	0.07	0.00	poi-10-4-3	145	0.29	0.00	145	0.32	0.00
poi-10-16-2	151	0.1	0.00	151	0.07	0.00	poi-10-5-1	150	0.09	0.00	150	0.09	0.00
poi-10-16-3	113	0.12	0.00	113	0.12	0.00	poi-10-5-2	147	0.09	0.00	147	0.14	0.00
poi-10-17-1	201	0.1	0.00	201	0.06	0.00	poi-10-5-3	147	0.14	0.00	147	0.42	0.00
poi-10-17-2	185	0.1	0.00	185	0.09	0.00	poi-10-6-1	178	0.11	0.00	178	0.11	0.00
poi-10-17-3	129	0.13	0.00	129	0.12	0.00	poi-10-6-2	153	0.11	0.00	153	0.13	0.00
poi-10-18-1	166	0.1	0.00	166	0.07	0.00	poi-10-6-3	148	0.21	0.00	148	0.52	0.00
poi-10-18-2	141	0.11	0.00	141	0.10	0.00	poi-10-7-1	158	0.12	0.00	158	0.09	0.00
poi-10-18-3	135	0.24	0.00	135	0.24	0.00	poi-10-7-2	131	0.12	0.00	131	0.42	0.00
poi-10-19-1	174	0.12	0.00	174	0.09	0.00	poi-10-7-3	123	0.26	0.00	123	0.70	0.00
poi-10-19-2	163	0.1	0.00	163	0.10	0.00	poi-10-8-1	133	0.1	0.00	133	0.18	0.00
poi-10-19-3	159	0.1	0.00	159	0.20	0.00	poi-10-8-2	113	0.12	0.00	113	0.18	0.00
poi-10-2-1	175	0.1	0.00	175	0.08	0.00	poi-10-8-3	97	0.37	0.00	97	0.20	0.00
poi-10-2-2	162	0.1	0.00	162	0.15	0.00	poi-10-9-1	189	0.1	0.00	189	0.07	0.00
poi-10-2-3	162	0.14	0.00	162	0.21	0.00	poi-10-9-2	168	0.11	0.00	168	0.10	0.00
poi-10-20-1	155	0.07	0.00	155	0.09	0.00	poi-10-9-3	159	0.17	0.00	165	0.26	0.00
poi-10-20-2	140	0.1	0.00	140	0.16	0.00							

Table 16: Results on *POI-instances* with $|C| = 19$

	RR21			*B&C*				*RR21*			*B&C*		
Id	UB	Time	% Gap	UB	Time	%Gap	Id	UB	Time	% Gap	UB	Time	% Gap
poi-20-1-1	226	0.4	0.00	226	1.06	0.00	poi-20-20-3	184	2.02	0.00	184	4.47	0.00
poi-20-1-2	212	0.71	0.00	212	3.39	0.00	poi-20-21-1	205	0.27	0.00	205	1.15	0.00
poi-20-1-3	185	3.67	0.00	185	9.70	0.00	poi-20-21-2	179	0.99	0.00	179	5.04	0.00
poi-20-10-1	213	0.23	0.00	213	1.42	0.00	poi-20-21-3	166	5.93	0.00	166	23.09	0.00
poi-20-10-2	188	2.05	0.00	188	5.61	0.00	poi-20-22-1	235	0.68	0.00	235	1.51	0.00
poi-20-10-3	178	1.27	0.00	178	4.44	0.00	poi-20-22-2	198	7.09	0.00	198	8.11	0.00
poi-20-11-1	205	0.32	0.00	205	1.18	0.00	poi-20-22-3	167	2.91	0.00	167	10.30	0.00
poi-20-11-2	180	1.08	0.00	180	3.20	0.00	poi-20-23-1	227	1.11	0.00	228	4.51	0.00
poi-20-11-3	175	3.17	0.00	175	10.74	0.00	poi-20-23-2	207	3.08	0.00	207	8.23	0.00
poi-20-12-1	225	15.03	0.00	225	2.91	0.00	poi-20-23-3	169	10.15	0.00	169	6.43	0.00
poi-20-12-2	172	7.96	0.00	172	2.86	0.00	poi-20-24-1	234	1.2	0.00	234	3.19	0.00
poi-20-12-3	158	3.46	0.00	158	6.96	0.00	poi-20-24-2	154	0.94	0.00	154	5.24	0.00
poi-20-13-1	208	0.41	0.00	208	0.88	0.00	poi-20-24-3	128	3.89	0.00	128	10.29	0.00
poi-20-13-2	200	3.51	0.00	200	3.88	0.00	poi-20-25-1	230	0.55	0.00	230	2.15	0.00
poi-20-13-3	162	11	0.00	169	5.06	0.00	poi-20-25-2	185	2.43	0.00	185	3.44	0.00
poi-20-14-1	189	1.12	0.00	189	2.96	0.00	poi-20-25-3	170	4.71	0.00	171	10.00	0.00
poi-20-14-2	174	16.54	0.00	174	12.08	0.00	poi-20-3-1	255	0.6	0.00	255	2.46	0.00
poi-20-14-3	159	8.84	0.00	159	23.85	0.00	poi-20-3-2	221	1.41	0.00	221	2.58	0.00
poi-20-15-1	251	1.12	0.00	251	2.80	0.00	poi-20-3-3	198	0.8	0.00	198	7.98	0.00
poi-20-15-2	170	1.67	0.00	170	2.58	0.00	poi-20-4-1	228	0.78	0.00	228	1.24	0.00
poi-20-15-3	157	7.3	0.00	157	6.35	0.00	poi-20-4-2	166	5.24	0.00	166	5.90	0.00
poi-20-16-1	212	0.49	0.00	212	1.48	0.00	poi-20-4-3	146	8.66	0.00	146	6.69	0.00
poi-20-16-2	172	1.1	0.00	172	2.99	0.00	poi-20-5-1	217	2.1	0.00	217	3.09	0.00
poi-20-16-3	171	3.22	0.00	171	4.35	0.00	poi-20-5-2	166	5.91	0.00	166	5.60	0.00
poi-20-17-1	192	2.23	0.00	192	2.79	0.00	poi-20-5-3	153	10.33	0.00	153	11.73	0.00
poi-20-17-2	137	1.34	0.00	137	11.97	0.00	poi-20-6-1	208	0.86	0.00	208	1.49	0.00
poi-20-17-3	134	6.41	0.00	134	15.82	0.00	poi-20-6-2	144	1.19	0.00	144	6.68	0.00
poi-20-18-1	223	0.57	0.00	225	2.60	0.00	poi-20-6-3	130	3.46	0.00	130	11.69	0.00
poi-20-18-2	201	1.05	0.00	201	6.84	0.00	poi-20-7-1	220	0.76	0.00	220	3.54	0.00
poi-20-18-3	178	2.82	0.00	178	22.70	0.00	poi-20-7-2	200	1.07	0.00	200	2.41	0.00
poi-20-19-1	215	1.12	0.00	215	1.04	0.00	poi-20-7-3	200	4.27	0.00	200	8.16	0.00
poi-20-19-2	155	0.92	0.00	155	4.25	0.00	poi-20-8-1	217	1.56	0.00	217	3.66	0.00
poi-20-19-3	146	1.8	0.00	146	7.61	0.00	poi-20-8-2	196	3.19	0.00	196	5.15	0.00
poi-20-2-1	233	0.75	0.00	233	1.83	0.00	poi-20-8-3	178	4.05	0.00	178	14.23	0.00
poi-20-2-2	195	2.27	0.00	196	4.69	0.00	poi-20-9-1	188	0.3	0.00	188	2.19	0.00
poi-20-2-3	172	3.93	0.00	172	6.86	0.00	poi-20-9-2	167	2.79	0.00	167	10.26	0.00
poi-20-20-1	208	0.65	0.00	208	2.33	0.00	poi-20-9-3	158	7.12	0.00	158	13.27	0.00
poi-20-20-2	186	2.25	0.00	186	3.82	0.00							

6.07 seconds for *RR21* and *B&C*, respectively.

Concerning the instances instead with 29 customers, we can observe from Table 17 that *RR21* is still able to solve to optimality all the instances. At the same time, our *B&C* is able to solve 74 out of 75 instances. However, the maximum optimality gap is equal to 2.38%. In terms of running time, we can observe that the average running time is equal to 92.29 seconds for *RR21* and 198.94 seconds for *B&C* (including the unsolved instance within the time limit).

Finally, we can observe from Table 18 that none of the two algorithms is able to optimally solve all the 75 instances with 39 customers within the time limit. In particular, *RR21* is able to solve 61 instances while our *B&C* determines the optimal solution on 57 instances. However, despite solving fewer instances to optimality, the *B&C* shows a lower average percentage gap (0.58%) than *RR21* (5.48%). Moreover, the *B&C* performs better than *RR21* also in terms of the largest percentage gap (6.22% vs. 51.01%). Instead, the running times are similar, with an average running time of 1475.88 and 1482.69 seconds for *RR21* and *B&C*, respectively.

Table 17: Results on *POI-instances* with $|C| = 29$

	RR21			*B&C*				*RR21*			*B&C*		
Id	UB	Time	% Gap	UB	Time	%Gap	Id	UB	Time	% Gap	UB	Time	% Gap
poi-30-1-1	246	21.16	0.00	246	14.66	0.00	poi-30-20-3	168	482.93	0.00	168	3600	2.38
poi-30-1-2	204	110.71	0.00	204	129.05	0.00	poi-30-21-1	233	6.53	0.00	233	11.25	0.00
poi-30-1-3	185	21.82	0.00	185	321.78	0.00	poi-30-21-2	213	90.81	0.00	213	128.80	0.00
poi-30-10-1	248	47.25	0.00	248	36.72	0.00	poi-30-21-3	201	28.63	0.00	201	199.27	0.00
poi-30-10-2	204	74.63	0.00	204	543.56	0.00	poi-30-22-1	255	5.38	0.00	255	13.30	0.00
poi-30-10-3	189	175.5	0.00	189	379.35	0.00	poi-30-22-2	217	101.55	0.00	217	43.38	0.00
poi-30-11-1	249	60.31	0.00	249	38.03	0.00	poi-30-22-3	209	160.06	0.00	209	107.71	0.00
poi-30-11-2	200	73.34	0.00	200	106.18	0.00	poi-30-23-1	232	181.92	0.00	232	99.48	0.00
poi-30-11-3	189	139.91	0.00	189	124.91	0.00	poi-30-23-2	197	709.05	0.00	197	133.59	0.00
poi-30-12-1	258	1.88	0.00	258	19.63	0.00	poi-30-23-3	180	39.12	0.00	180	264.68	0.00
poi-30-12-2	190	16.21	0.00	190	66.34	0.00	poi-30-24-1	219	21.66	0.00	219	20.64	0.00
poi-30-12-3	180	73.77	0.00	180	527.78	0.00	poi-30-24-2	185	37.68	0.00	185	345.63	0.00
poi-30-13-1	256	21.42	0.00	256	10.76	0.00	poi-30-24-3	177	63.39	0.00	177	443.47	0.00
poi-30-13-2	205	52.99	0.00	205	61.21	0.00	poi-30-25-1	259	2.39	0.00	259	23.12	0.00
poi-30-13-3	199	181.04	0.00	199	386.96	0.00	poi-30-25-2	234	46.16	0.00	234	84.33	0.00
poi-30-14-1	229	7.17	0.00	229	8.06	0.00	poi-30-25-3	227	197.12	0.00	227	290.78	0.00
poi-30-14-2	186	117.55	0.00	186	55.74	0.00	poi-30-3-1	224	75.84	0.00	224	31.92	0.00
poi-30-14-3	177	200.46	0.00	177	246.43	0.00	poi-30-3-2	192	39.39	0.00	192	42.93	0.00
poi-30-15-1	214	57.29	0.00	214	26.37	0.00	poi-30-3-3	192	245.5	0.00	192	55.33	0.00
poi-30-15-2	189	73.75	0.00	189	112.67	0.00	poi-30-4-1	260	6.67	0.00	260	34.37	0.00
poi-30-15-3	184	44.88	0.00	184	141.59	0.00	poi-30-4-2	206	25.13	0.00	206	378.35	0.00
poi-30-16-1	246	4.99	0.00	246	11.58	0.00	poi-30-4-3	158	25.69	0.00	158	156.68	0.00
poi-30-16-2	202	50.72	0.00	202	74.26	0.00	poi-30-5-1	230	9.56	0.00	230	14.32	0.00
poi-30-16-3	196	83.64	0.00	196	101.33	0.00	poi-30-5-2	193	54.89	0.00	193	311.91	0.00
poi-30-17-1	255	33.82	0.00	255	22.71	0.00	poi-30-5-3	188	80.25	0.00	188	217.22	0.00
poi-30-17-2	192	26.28	0.00	192	47.16	0.00	poi-30-6-1	251	101.66	0.00	251	35.21	0.00
poi-30-17-3	174	32.11	0.00	174	102.18	0.00	poi-30-6-2	208	76.1	0.00	208	256.35	0.00
poi-30-18-1	257	28.8	0.00	257	32.05	0.00	poi-30-6-3	194	125.45	0.00	194	52.24	0.00
poi-30-18-2	219	159.4	0.00	219	46.71	0.00	poi-30-7-1	211	12.19	0.00	211	18.97	0.00
poi-30-18-3	179	53.9	0.00	179	954.68	0.00	poi-30-7-2	190	14.73	0.00	190	116.52	0.00
poi-30-19-1	232	6.59	0.00	232	32.95	0.00	poi-30-7-3	182	72.03	0.00	182	148.21	0.00
poi-30-19-2	203	9.2	0.00	203	149.00	0.00	poi-30-8-1	239	5.93	0.00	239	11.02	0.00
poi-30-19-3	202	53.92	0.00	202	63.95	0.00	poi-30-8-2	200	14.55	0.00	200	111.55	0.00
poi-30-2-1	263	55.31	0.00	263	27.75	0.00	poi-30-8-3	195	109.96	0.00	195	256.07	0.00
poi-30-2-2	200	78.26	0.00	200	105.56	0.00	poi-30-9-1	259	66.9	0.00	259	25.39	0.00
poi-30-2-3	180	120.08	0.00	180	214.48	0.00	poi-30-9-2	210	154.03	0.00	210	185.48	0.00
poi-30-20-1	228	16.14	0.00	228	39.31	0.00	poi-30-9-3	193	886.05	0.00	193	498.22	0.00
poi-30-20-2	178	158.56	0.00	178	768.54	0.00							

Table 18: Results on *POI-instances* with $|C| = 39$

	RR21			*B&C*				*RR21*			*B&C*		
Id	UB	Time	% Gap	UB	Time	%Gap	Id	UB	Time	% Gap	UB	Time	% Gap
poi-40-1-1	283	1364.79	0.00	283	115.45	0.00	poi-40-20-3	216	2545.92	0.00	216	719.56	0.00
poi-40-1-2	213	563.64	0.00	214	3600	0.93	poi-40-21-1	269	2148.13	0.00	269	117.65	0.00
poi-40-1-3	194	3267.08	0.00	194	1002.15	0.00	poi-40-21-2	227	309.95	0.00	227	3077.10	0.00
poi-40-10-1	259	403.45	0.00	259	216.47	0.00	poi-40-21-3	218	802.95	0.00	218	3600	1.38
poi-40-10-2	220	74.69	0.00	220	402.12	0.00	poi-40-22-1	286	50.98	0.00	286	163.69	0.00
poi-40-10-3	218	696.9	0.00	218	971.05	0.00	poi-40-22-2	243	413.37	0.00	243	1017.52	0.00
poi-40-11-1	266	3600	0.75	266	351.62	0.00	poi-40-22-3	339	3600	34.22	224	1302.18	0.00
poi-40-11-2	222	527.99	0.00	222	513.92	0.00	poi-40-23-1	395	3600	26.84	302	609.36	0.00
poi-40-11-3	210	378.64	0.00	210	328.44	0.00	poi-40-23-2	395	3600	38.48	244	480.90	0.00
poi-40-12-1	262	35.93	0.00	262	47.25	0.00	poi-40-23-3	230	1297.49	0.00	230	546.49	0.00
poi-40-12-2	222	1405.59	0.00	222	1164.56	0.00	poi-40-24-1	244	3227.38	0.00	244	1726.32	0.00
poi-40-12-3	209	2534	0.00	209	1162.17	0.00	poi-40-24-2	215	125.84	0.00	215	391.89	0.00
poi-40-13-1	266	988.14	0.00	266	1408.90	0.00	poi-40-24-3	212	1537.43	0.00	212	1076.67	0.00
poi-40-13-2	222	326.15	0.00	222	363.11	0.00	poi-40-25-1	417	3600	33.57	281	1478.26	0.00
poi-40-13-3	363	3600	42.15	210	3600	0.95	poi-40-25-2	210	361.2	0.00	210	667.90	0.00
poi-40-14-1	287	64.94	0.00	287	104.26	0.00	poi-40-25-3	190	2450.66	0.00	190	3600	1.05
poi-40-14-2	229	673.8	0.00	229	1686.05	0.00	poi-40-3-1	282	438.14	0.00	282	51.12	0.00
poi-40-14-3	210	254.43	0.00	210	1063.50	0.00	poi-40-3-2	416	3600	43.51	235	1412.65	0.00
poi-40-15-1	279	215.33	0.00	279	3600	1.43	poi-40-3-3	416	3600	48.08	216	3600	3.24
poi-40-15-2	234	302.24	0.00	234	550.87	0.00	poi-40-4-1	284	3600	0.35	288	1132.86	0.00
poi-40-15-3	346	3600	33.53	231	3600	4.33	poi-40-4-2	234	933.35	0.00	234	3600	1.71
poi-40-16-1	270	1918.15	0.00	270	3600	1.85	poi-40-4-3	210	565.64	0.00	210	1602.15	0.00
poi-40-16-2	210	952.09	0.00	210	3600	0.95	poi-40-5-1	285	810.61	0.00	285	508.73	0.00
poi-40-16-3	201	229.68	0.00	201	3202.26	0.00	poi-40-5-2	231	2302.31	0.00	231	1401.09	0.00
poi-40-17-1	285	187.93	0.00	285	752.74	0.00	poi-40-5-3	203	450.09	0.00	203	3600	3.41
poi-40-17-2	246	79.75	0.00	246	707.20	0.00	poi-40-6-1	240	2325.71	0.00	240	921.17	0.00
poi-40-17-3	359	3600	32.59	242	1142.52	0.00	poi-40-6-2	192	1270.8	0.00	192	3600	3.65
poi-40-18-1	391	3600	25.32	297	3600	1.68	poi-40-6-3	181	1351.85	0.00	181	3600	1.10
poi-40-18-2	239	487.04	0.00	241	3600	6.22	poi-40-7-1	271	3600	0.74	271	593.41	0.00
poi-40-18-3	221	68.9	0.00	221	3600	5.43	poi-40-7-2	206	506.34	0.00	206	2602.22	0.00
poi-40-19-1	236	560.59	0.00	236	715.17	0.00	poi-40-7-3	396	3600	51.01	194	3600	1.55
poi-40-19-2	204	2748.63	0.00	204	400.35	0.00	poi-40-8-1	277	494.89	0.00	277	594.32	0.00
poi-40-19-3	204	2452.33	0.00	204	260.69	0.00	poi-40-8-2	228	628.96	0.00	228	622.08	0.00
poi-40-2-1	269	60.64	0.00	269	175.37	0.00	poi-40-8-3	220	972.64	0.00	220	667.60	0.00
poi-40-2-2	218	1092.7	0.00	218	1128.46	0.00	poi-40-9-1	289	422.85	0.00	289	202.98	0.00
poi-40-2-3	204	834.17	0.00	204	2146.83	0.00	poi-40-9-2	226	924.52	0.00	226	924.76	0.00
poi-40-20-1	262	2285.6	0.00	262	363.29	0.00	poi-40-9-3	208	1252.64	0.00	208	503.69	0.00
poi-40-20-2	220	1332.06	0.00	220	398.77	0.00							

Bibliography

[1] K. Aaby, R. L. Abbey, J. W. Herrmann, M. Treadwell, C. S. Jordan, and K. Wood. Embracing computer modeling to address pandemic influenza in the 21st century. *Journal of Public Health Management and Practice*, 12(4):365–372, 2006.

[2] K. Aaby, J. W. Herrmann, C. S. Jordan, M. Treadwell, and K. Wood. Montgomery county's public health service uses operations research to plan emergency mass dispensing and vaccination clinics. *Interfaces*, 36(6):569–579, 2006.

[3] b. I. Abdulraheem, A. Olapipo, M. Amodu, et al. Primary health care services in nigeria: Critical issues and strategies for enhancing the use by the rural communities. *Journal of public health and epidemiology*, 4(1):5–13, 2012.

[4] I. Adan, J. Bekkers, N. Dellaert, J. Vissers, and X. Yu. Patient mix optimisation and stochastic resource requirements: A case study in cardiothoracic surgery planning. *Health care management science*, 12:129–141, 2009.

[5] B. Addis, G. Carello, A. Grosso, and E. Tànfani. Operating room scheduling and rescheduling: a rolling horizon approach. *Flexible Services and Manufacturing Journal*, 28(1-2):206–232, 2016.

[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*, 17:49–59, 2014.

[7] A. Ahmadi-Javid and N. Ramshe. A stochastic location model for designing primary healthcare networks integrated with workforce cross-training. *Operations Research for Health Care*, 24:100226, 2020.

[8] A. Ahmadi-Javid, P. Seyedi, and S. S. Syam. A survey of healthcare facility location. *Computers & Operations Research*, 79:223–263, 2017.

[9] U. Aickelin, E. K. Burke, and J. Li. An estimation of distribution algorithm with intelligent local search for rule-based nurse rostering. *Journal of the Operational Research Society*, 58(12):1574–1585, 2007.

[10] U. Aickelin and K. A. Dowsland. Exploiting problem structure in a genetic algorithm approach to a nurse rostering problem. *Journal of scheduling*, 3(3):139–153, 2000.

[11] U. Aickelin and K. A. Dowsland. An indirect genetic algorithm for a nurse-scheduling problem. *Computers & operations research*, 31(5):761–778, 2004.

[12] American Society of Anesthesiologists. Asa physical status classification system. https://www.asahq.org/standards-and-guidelines/asa-physical-status-classification-system, 2021.

[13] M. Arani, Y. Chan, X. Liu, and M. Momenitabar. A lateral resupply blood supply chain network design under uncertainties. *Applied Mathematical Modelling*, 93:165–187, 2021.

[14] R. Aringhieri, M. E. Bruni, S. Khodaparasti, and J. T. van Essen. Emergency medical services and beyond: Addressing new challenges through a wide literature review. *Computers & Operations Research*, 78:349–368, 2017.

[15] B. Armbruster and M. L. Brandeau. Contact tracing to control infectious disease: when enough is enough. *Health care management science*, 10:341–355, 2007.

[16] N. Aslani, O. Kuzgunkaya, N. Vidyarthi, and D. Terekhov. A robust optimization model for tactical capacity planning in an outpatient setting. *Health Care Management Science*, 24:26–40, 2021.

[17] M. Y. N. Attari, S. H. R. Pasandideh, A. Aghaie, and S. T. A. Niaki. A bi-objective robust optimization model for a blood collection and testing

problem: an accelerated stochastic benders decomposition. *Annals of Operations Research*, pages 1–39, 2018.

[18] P. Avella, M. Boccia, A. Sforza, and I. Vasil'Ev. A branch-and-cut algorithm for the median-path problem. *Computational Optimization and Applications*, 32(3):215–230, 2005.

[19] AVIS. Il sistema trasfusionale campano e la raccolta associativa. `https://www.avis.it/wp-content/uploads/attachments/4305_documento.pdf`, 2021.

[20] A. Aydemir-Karadag. Bi-objective adaptive large neighborhood search algorithm for the healthcare waste periodic location inventory routing problem. *Arabian Journal for Science and Engineering*, 47(3):3861–3876, 2022.

[21] M. Bagheri, A. G. Devin, and A. Izanloo. An application of stochastic programming method for nurse scheduling problem in real word hospital. *Computers & industrial engineering*, 96:192–200, 2016.

[22] M. O. Ball and F. L. Lin. A reliability model applied to emergency service vehicle location. *Operations research*, 41(1):18–36, 1993.

[23] F. Ballestín, Á. Pérez, and S. Quintanilla. Scheduling and rescheduling elective patients in operating rooms to minimise the percentage of tardy patients. *Journal of Scheduling*, 22(1):107–118, 2019.

[24] J. F. Bard and H. W. Purnomo. Preference scheduling for nurses using column generation. *European Journal of Operational Research*, 164(2):510–534, 2005.

[25] R. Bargetto, T. Garaix, and X. Xie. Dynamic insertion of emergency surgeries with different waiting time targets. *IEEE Transactions on Automation Science and Engineering*, 16(1):87–99, 2018.

[26] R. Bargetto, T. Garaix, and X. Xie. Dynamic insertion of emergency surgeries with different waiting time targets. *IEEE Transactions on Automation Science and Engineering*, 16(1):87–99, 2018.

[27] R. Bargetto, T. Garaix, and X. Xie. A branch-and-price-and-cut algorithm for operating room scheduling under human resource constraints. *Computers & Operations Research*, 152:106136, 2023.

[28] O. Barros, R. Weber, and C. Reveco. Demand analysis and capacity management for hospital emergencies using advanced forecasting models and stochastic simulation. *Operations Research Perspectives*, 8:100208, 2021.

[29] A. Başar, B. Çatay, and T. Ünlüyurt. A taxonomy for emergency service station location problem. *Optimization letters*, 6:1147–1160, 2012.

[30] R. Batta, J. M. Dolan, and N. N. Krishnamurthy. The maximal expected covering location problem: Revisited. *Transportation science*, 23(4):277–287, 1989.

[31] S. Batun and M. A. Begen. Optimization in healthcare delivery modeling: Methods and applications. In *Handbook of Healthcare Operations Management: Methods and Applications*, pages 75–119. Springer, 2013.

[32] H. Beaulieu, J. A. Ferland, B. Gendron, and P. Michelon. A mathematical programming approach for scheduling physicians in the emergency room. *Health care management science*, 3:193–200, 2000.

[33] A. J. Beech. Market-based demand forecasting promotes informed strategic financial planning. *Healthcare Financial Management*, 55(11):46–56, 2001.

[34] V. Bélanger, A. Ruiz, and P. Soriano. Recent optimization models and trends in location, relocation, and dispatching of emergency medical vehicles. *European Journal of Operational Research*, 272(1):1–23, 2019.

[35] J. Beliën and E. Demeulemeester. Building cyclic master surgery schedules with leveled resulting bed occupancy. *European journal of operational research*, 176(2):1185–1204, 2007.

[36] J. Beliën and E. Demeulemeester. A branch-and-price approach for integrating nurse and surgery scheduling. *European journal of operational research*, 189(3):652–668, 2008.

[37] J. Beliën and H. Forcé. Supply chain management of blood products: A literature review. *European Journal of Operational Research*, 217(1):1–16, 2012.

[38] F. Bellanti, G. Carello, F. Della Croce, and R. Tadei. A greedy-based neighborhood search approach to a nurse rostering problem. *European Journal of Operational Research*, 153(1):28–40, 2004.

[39] I. Berrada, J. A. Ferland, and P. Michelon. A multi-objective approach to nurse scheduling with both hard and soft constraints. *Socio-economic planning sciences*, 30(3):183–193, 1996.

[40] S. Bertels and T. Fahle. A hybrid setup for a hybrid scenario: combining heuristics for the home health care problem. *Computers & operations research*, 33(10):2866–2890, 2006.

[41] M. Bester, I. Nieuwoudt, and J. H. Van Vuuren. Finding good nurse duty schedules: a case study. *Journal of scheduling*, 10:387–405, 2007.

[42] J. T. Blake, F. Dexter, and J. Donald. Operating room managers' use of integer programming for assigning block time to surgical groups: a case study. *Anesthesia & Analgesia*, 94(1):143–148, 2002.

[43] J. T. Blake and J. Donald. Mount sinai hospital uses integer programming to allocate operating room time. *Interfaces*, 32(2):63–73, 2002.

[44] M. Boccia, T. G. Crainic, A. Sforza, and C. Sterle. Multi-commodity location-routing: Flow intercepting formulation and branch-and-cut algorithm. *Computers & Operations Research*, 89:94–112, 2018.

[45] M. Boccia, A. Mancuso, A. Masone, and C. Sterle. A feature based solution approach for the flying sidekick traveling salesman problem. In *International Conference on Mathematical Optimization Theory and Operations Research*, pages 131–146. Springer, 2021.

[46] M. Boccia, A. Mancuso, A. Masone, and C. Sterle. Optimization for surgery department management: an application to a hospital in naples. 2023.

[47] M. Boccia, A. Masone, A. Sforza, and C. Sterle. A column-and-row generation approach for the flying sidekick travelling salesman problem. *Transportation Research Part C: Emerging Technologies*, 124:102913, 2021.

[48] M. Boccia, A. Masone, A. Sforza, and C. Sterle. An exact approach for a variant of the fs-tsp. *Transportation Research Procedia*, 52:51–58, 2021.

[49] A. Bon and T. Ng. An optimization of inventory demand forecasting in university healthcare centre. In *IOP Conference Series: Materials Science and Engineering*, volume 166, page 012035. IOP Publishing, 2017.

[50] M. L. Brandeau, F. Sainfort, and W. P. Pierskalla. *Operations research and health care: a handbook of methods and applications*, volume 70. Springer Science & Business Media, 2004.

[51] L. Brotcorne, G. Laporte, and F. Semet. Ambulance location and relocation models. *European journal of operational research*, 147(3):451–463, 2003.

[52] M. E. Bruni, D. Conforti, N. Sicilia, and S. Trotta. A new organ transplantation location–allocation policy: a case study of italy. *Health care management science*, 9:125–142, 2006.

[53] G. Bruno, A. Diglio, C. Piccolo, and L. Cannavacciuolo. Territorial reorganization of regional blood management systems: Evidences from an italian case study. *Omega*, 89:54–70, 2019.

[54] E. K. Burke, P. D. Causmaecker, S. Petrovic, and G. V. Berghe. Metaheuristics for handling time interval coverage constraints in nurse scheduling. *Applied Artificial Intelligence*, 20(9):743–766, 2006.

[55] E. K. Burke, T. Curtois, G. Post, R. Qu, and B. Veltman. A hybrid heuristic ordering and variable neighbourhood search for the nurse rostering problem. *European journal of operational research*, 188(2):330–341, 2008.

[56] E. K. Burke, T. Curtois, R. Qu, and G. Vanden Berghe. A scatter search methodology for the nurse rostering problem. *Journal of the Operational Research Society*, 61(11):1667–1679, 2010.

[57] V. Capocasale, D. Gotta, S. Musso, and G. Perboli. A blockchain, 5g and iot-based transaction management system for smart logistics: an hyperledger framework. In *2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*, pages 1285–1290. IEEE, 2021.

[58] V. Capocasale and G. Perboli. Standardizing smart contracts. *IEEE Access*, 2022.

[59] B. Cardoen, E. Demeulemeester, and J. Beliën. Optimizing a multiple objective surgical case sequencing problem. *International Journal of Production Economics*, 119(2):354–366, 2009.

[60] B. Cardoen, E. Demeulemeester, and J. Beliën. Optimizing a multiple objective surgical case sequencing problem. *International Journal of Production Economics*, 119(2):354–366, 2009.

[61] B. Cardoen, E. Demeulemeester, and J. Beliën. Sequencing surgical cases in a day-care environment: an exact branch-and-price approach. *Computers & Operations Research*, 36(9):2660–2669, 2009.

[62] B. Cardoen, E. Demeulemeester, and J. Beliën. Operating room planning and scheduling: A literature review. *European journal of operational research*, 201(3):921–932, 2010.

[63] B. Cardoen, E. Demeulemeester, and J. Beliën. Operating room planning and scheduling: A literature review. *European journal of operational research*, 201(3):921–932, 2010.

[64] L. Catalano, V. Piccinini, I. Pati, F. Masiello, G, G. Marano, S. Pupella, and V. De Angelis. Italian blood system 2020: activity data, haemovigilance and epidemiological surveillance. *Rapporti ISTISAN 21/14*, 2021.

[65] S. Cavani, M. Iori, and R. Roberti. Exact methods for the traveling salesman problem with multiple drones. *Transportation Research Part C: Emerging Technologies*, 130:103280, 2021.

[66] E. Cetin and L. S. Sarul. A blood bank location model: A multiobjective approach. *European Journal of Pure and Applied Mathematics*, 2(1):112–124, 2009.

[67] P. Chaiwuttisak, H. Smith, Y. Wu, C. Potts, T. Sakuldamrongpanich, and S. Pathomsiri. Location of low-cost blood collection and distribution centres in thailand. *Operations Research for Health Care*, 9:7–15, 2016.

[68] I.-M. Chao. A tabu search method for the truck and trailer routing problem. *Computers & Operations Research*, 29(1):33–51, 2002.

[69] S. E. Chick, H. Mamani, and D. Simchi-Levi. Supply chain coordination and influenza vaccination. *Operations Research*, 56(6):1493–1506, 2008.

[70] C.-F. Chien, F.-P. Tseng, and C.-H. Chen. An evolutionary approach to rehabilitation patient scheduling: A case study. *European Journal of Operational Research*, 189(3):1234–1253, 2008.

[71] C. Childers and M. Maggard-Gibbons. Understanding costs of care in the operating room. *JAMA Surgery*, 153(4):1–7, 2018.

[72] S. C. Chu and L. Chu. A modeling framework for hospital location and service allocation. *International transactions in operational research*, 7(6):539–568, 2000.

[73] M. Cissé, S. Yalçındağ, Y. Kergosien, E. Şahin, C. Lenté, and A. Matta. Or problems related to home health care: A review of relevant routing and scheduling problems. *Operations research for health care*, 13:1–22, 2017.

[74] I. Correia and F. Saldanha-da Gama. Facility location under uncertainty. *Location Science*, pages 185–213, 2019.

[75] M. J. Côté, S. S. Syam, W. B. Vogel, and D. C. Cowper. A mixed integer programming model to locate traumatic brain injury treatment units in the department of veterans affairs: A case study. *Health care management science*, 10:253–267, 2007.

[76] M. J. Cote and S. L. Tucker. Four methodologies to improve healthcare demand forecasting. *Healthcare Financial Management*, 55(5):54–54, 2001.

[77] M. Daskin. Location, dispatching, and routing models for emergency services with stochastic travel times. spatial analysis and location-allocation models, a. ghosh and g. rushton, eds, 1987.

[78] M. S. Daskin. Application of an expected covering model to emergency medical service system design. *Decision Sciences*, 13(3):416–439, 1982.

[79] M. S. Daskin. A maximum expected covering location model: formulation, properties and heuristic solution. *Transportation science*, 17(1):48–70, 1983.

[80] M. S. Daskin and E. H. Stern. A hierarchical objective set covering model for emergency medical service vehicle deployment. *Transportation Science*, 15(2):137–152, 1981.

[81] K. De Boeck, C. Decouttere, and N. Vandaele. Vaccine distribution chains in low-and middle-income countries: A literature review. *Omega*, 97:102097, 2020.

[82] J. C. de Freitas and P. H. V. Penna. A variable neighborhood search for flying sidekick traveling salesman problem. *International Transactions in Operational Research*, 27(1):267–290, 2020.

[83] M. Dell'Amico, R. Montemanni, and S. Novellani. Exact models for the flying sidekick traveling salesman problem. *International Transactions in Operational Research*, 29(3):1360–1393, 2022.

[84] M. Dell'Amico, R. Montemanni, and S. Novellani. Algorithms based on branch and bound for the flying sidekick traveling salesman problem. *Omega*, 104:102493, 2021.

[85] M. Dell'Amico, R. Montemanni, and S. Novellani. Drone-assisted deliveries: New formulations for the flying sidekick traveling salesman problem. *Optimization Letters*, 15(5):1617–1648, 2021.

[86] F. Dexter, A. Macario, and R. D. Traub. Which algorithm for scheduling add-on elective cases maximizes operating room utilization? use of bin packing algorithms and fuzzy constraints in operating room management. *The Journal of the American Society of Anesthesiologists*, 91(5):1491–1491, 1999.

[87] M. Di Mascolo, C. Martinez, and M.-L. Espinouse. Routing and scheduling in home health care: A literature survey and bibliometric analysis. *Computers & Industrial Engineering*, 158:107255, 2021.

[88] A. Diglio, A. Mancuso, A. Masone, C. Piccolo, and C. Sterle. A milp formulation for the reorganization of the blood supply chain in italian regions. *Optimization and Data Science: Trends and Applications, AIRO Springer Series*, 6:51–66, 2021.

[89] K. A. Dowsland. Nurse scheduling with tabu search and strategic oscillation. *European journal of operational research*, 106(2-3):393–407, 1998.

[90] K. A. Dowsland and J. M. Thompson. Solving a nurse scheduling problem with knapsacks, networks and tabu search. *Journal of the operational research society*, 51:825–833, 2000.

[91] G. Du, X. Liang, and C. Sun. Scheduling optimization of home health care service considering patients' priorities and time windows. *Sustainability*, 9(2):253, 2017.

[92] Y. Dudko, D. E. Robey, E. Kruger, and M. Tennant. Selecting a location for a primary healthcare facility: combining a mathematical approach with a geographic information system to rank areas of relative need. *Australian Journal of Primary Health*, 24(2):130–134, 2018.

[93] D. Duma and R. Aringhieri. An online optimization approach for the real time management of operating rooms. *Operations Research for Health Care*, 7:40–51, 2015.

[94] D. Duma and R. Aringhieri. The management of non-elective patients: shared vs. dedicated policies. *Omega*, 83:199–212, 2019.

[95] G. Durán, P. A. Rey, and P. Wolff. Solving the operating room scheduling problem with prioritized lists of patients. *Annals of Operations Research*, 258(2):395–414, 2017.

[96] S. R. Earnshaw, K. Hicks, A. Richter, and A. Honeycutt. A linear programming model for allocating hiv prevention funds with state agencies: a pilot study. *Health Care Management Science*, 10(3):239–252, 2007.

[97] S. R. Earnshaw, A. Richter, S. W. Sorensen, T. J. Hoerger, K. A. Hicks, M. Engelgau, T. Thompson, K. Venkat Narayan, D. F. Williamson, E. Gregg, et al. Optimal allocation of resources across four interventions for type 2 diabetes. *Medical Decision Making*, 22(1_suppl):80–91, 2002.

[98] D. J. Eaton, M. S. Daskin, D. Simmons, B. Bulloch, and G. Jansma. Determining emergency medical service vehicle deployment in austin, texas. *Interfaces*, 15(1):96–108, 1985.

[99] A. Elalouf, S. Hovav, D. Tsadikovich, and L. Yedidsion. Minimizing operational costs by restructuring the blood sample collection chain. *Operations Research for Health Care*, 7:81–93, 2015.

[100] S. Enayati and O. Y. Özaltın. Optimal influenza vaccine distribution with equity. *European Journal of Operational Research*, 283(2):714–725, 2020.

[101] E. Erdem, X. Qu, and J. Shi. Rescheduling of elective patients upon the arrival of emergency patients. *Decision Support Systems*, 54(1):551–563, 2012.

[102] E. Erkut, A. Ingolfsson, T. Sim, and G. Erdoğan. Computational comparison of five maximal covering models for locating ambulances. *Geographical Analysis*, 41(1):43–65, 2009.

[103] M. M. Ershadi and M. S. Ershadi. Logistic planning for pharmaceutical supply chain using multi-objective optimization model. *International Journal of Pharmaceutical and Healthcare Marketing*, 16(1):75–100, 2022.

[104] J. Euchi, M. Masmoudi, and P. Siarry. Home health care routing and scheduling problems: a literature review. *4OR*, 20(3):351–389, 2022.

[105] M. Fairley, D. Scheinker, and M. L. Brandeau. Improving the efficiency of the operating room environment with an optimization and machine learning model. *Health care management science*, 22(4):756–767, 2019.

[106] A. M. Fathollahi-Fard, A. Ahmadi, F. Goodarzian, and N. Cheikhrouhou. A bi-objective home healthcare routing and scheduling problem considering patients' satisfaction in a fuzzy environment. *Applied soft computing*, 93:106385, 2020.

[107] M. Fattahi, E. Keyvanshokooh, D. Kannan, and K. Govindan. Resource planning strategies for healthcare systems during a pandemic. *European Journal of Operational Research*, 304(1):192–206, 2023.

[108] H. Fei, C. Chu, and N. Meskens. Solving a tactical operating room planning problem by a column-generation-based heuristic procedure with four criteria. *Annals of Operations Research*, 166:91–108, 2009.

[109] H. Fei, N. Meskens, and C. Chu. A planning and scheduling problem for an operating theatre using an open scheduling strategy. *Computers & Industrial Engineering*, 58(2):221–230, 2010.

[110] G. Felici and C. Gentile. A polyhedral approach for the staff rostering problem. *Management Science*, 50(3):381–393, 2004.

[111] H. J. Finarelli Jr and T. Johnson. Effective demand forecasting in 9 steps: shifts in demand for a hospital's services can occur unexpectedly. demand forecasting can help you prepare for these shifts and avoid strategic missteps. *Healthcare Financial Management*, 58(11):52–58, 2004.

[112] N. K. Freeman, S. H. Melouk, and J. Mittenthal. A scenario-based approach for operating theater scheduling under uncertainty. *Manufacturing & Service Operations Management*, 18(2):245–261, 2016.

[113] R. D. Galvão and C. ReVelle. A lagrangean heuristic for the maximal covering location problem. *European Journal of Operational Research*, 88(1):114–123, 1996.

[114] Gazzetta Ufficiale. Linee guida per l'accreditamento dei servizi trasfusionali e delle unità di raccolta del sangue e degli emocomponenti (allegato A). `https://www.gazzettaufficiale.it/atto/serie_generale/caricaArticolo?art.progressivo=0&art.idArticolo=1&art.versione=1&art.codiceRedazionale=13A03965&art.dataPubblicazioneGazzetta=2013-05-09&art.idGruppo=0&art.idSottoArticolo1=10&art.idSottoArticolo=1&art.flagTipoArticolo=1`, May 2013.

[115] M. Gendreau, G. Laporte, and F. Semet. Solving an ambulance location model by tabu search. *Location science*, 5(2):75–88, 1997.

[116] H. Gilani and H. Sahebi. A data-driven robust optimization model by cutting hyperplanes on vaccine access uncertainty in covid-19 vaccine supply chain. *Omega*, 110:102637, 2022.

[117] J. Goddard and M. Tavakoli. Efficiency and welfare implications of managed public sector hospital waiting lists. *European journal of operational research*, 184(2):778–792, 2008.

[118] J. Goldberg, R. Dietrich, J. M. Chen, M. G. Mitwasi, T. Valenzuela, and E. Criss. Validating and applying a model for locating emergency medical vehicles in tuczon, az. *European Journal of Operational Research*, 49(3):308–324, 1990.

[119] A. Goli, A. Ala, and S. Mirjalili. A robust possibilistic programming framework for designing an organ transplant supply chain under uncertainty. *Annals of Operations Research*, 328(1):493–530, 2023.

[120] P. R. Gould and T. R. Leinbach. Approach to the geographic assignment of hospital services. *Tijdschrift voor Economische en Sociale Geografie*, 57(5):203–206, 1966.

[121] K. Govindan, M. Fattahi, and E. Keyvanshokooh. Supply chain network design under uncertainty: A comprehensive review and future research directions. *European Journal of Operational Research*, 263(1):108–141, 2017.

[122] L. V. Green. Capacity planning and management in hospitals. *Operations research and health care: A handbook of methods and applications*, pages 15–41, 2004.

[123] L. V. Green and S. Savin. Reducing delays for medical appointments: A queueing approach. *Operations Research*, 56(6):1526–1538, 2008.

[124] F. Guerriero and R. Guido. Operational research in the management of the operating theatre: a survey. *Health care management science*, 14:89–114, 2011.

[125] A. Guinet and S. Chaabane. Operating theatre planning. *International Journal of Production Economics*, 85(1):69–81, 2003.

[126] A. Gunawan, H. C. Lau, and P. Vansteenwegen. Orienteering problem: A survey of recent variants, solution approaches and applications. *European Journal of Operational Research*, 255(2):315–332, 2016.

[127] W. J. Gutjahr and M. S. Rauner. An aco algorithm for a dynamic regional nurse-scheduling problem in austria. *Computers & Operations Research*, 34(3):642–666, 2007.

[128] S. N. Hall, S. H. Jacobson, and E. C. Sewell. An analysis of pediatric vaccine formulary selection problems. *Operations Research*, 56(6):1348–1365, 2008.

[129] B. Hamdan and A. Diabat. A two-stage multi-echelon stochastic blood supply chain problem. *Computers & Operations Research*, 101:130–143, 2019.

[130] E. Hans, G. Wullink, M. Van Houdenhoven, and G. Kazemier. Robust surgery loading. *European Journal of Operational Research*, 185(3):1038–1050, 2008.

[131] K. Hanson, N. Brikci, D. Erlangga, A. Alebachew, M. De Allegri, D. Balabanova, M. Blecher, C. Cashin, A. Esperato, D. Hipgrave, et al. The lancet global health commission on financing primary health care: putting people at the centre. *The Lancet Global Health*, 10(5):e715–e772, 2022.

[132] P. R. Harper, M. Sayyad, V. de Senna, A. K. Shahani, C. Yajnik, and K. Shelgikar. A systems modelling approach for the prevention and treatment of diabetic retinopathy. *European Journal of Operational Research*, 150(1):81–91, 2003.

[133] P. R. Harper, A. Shahani, J. Gallagher, and C. Bowie. Planning health services with explicit geographical considerations: a stochastic location–allocation approach. *Omega*, 33(2):141–152, 2005.

[134] P. R. Harper and A. K. Shahani. Modelling for the planning and management of bed capacities in hospitals. *Journal of the Operational research Society*, 53:11–18, 2002.

[135] P. R. Harper and A. K. Shahani. A decision support system for the care of hiv and aids patients in india. *European Journal of Operational Research*, 147(1):187–197, 2003.

[136] S. H. Hashemi Doulabi, L.-M. Rousseau, and G. Pesant. A constraint-programming-based branch-and-price-and-cut approach for operating room planning and scheduling. *INFORMS Journal on Computing*, 28(3):432–448, 2016.

[137] M. J. Hodgson, G. Laporte, and F. Semet. A covering tour model for planning mobile health care facilities in suhumdistrict, ghama. *Journal of Regional Science*, 38(4):621–638, 1998.

[138] K. Hogan and C. ReVelle. Concepts and applications of backup coverage. *Management science*, 32(11):1434–1444, 1986.

[139] A. Ikegami and A. Niwa. A subproblem-centric model and approach to the nurse scheduling problem. *Mathematical programming*, 97:517–541, 2003.

[140] A. Ingolfsson, S. Budge, and E. Erkut. Optimal ambulance location with random delays and travel times. *Health care management science*, 11:262–274, 2008.

[141] ISTAT. Invecchiamento attivo e condizioni di vita degli anziani in italia. https://www.istat.it/it/files//2020/08/Invecchiamento-attivo-e-condizioni-di-vita-degli-anziani-in-Italia.pdf, 2021.

[142] A. Jabbarzadeh, B. Fahimnia, and S. Seuring. Dynamic supply chain network design for the supply of blood in disasters: A robust model with real world application. *Transportation Research Part E: Logistics and Transportation Review*, 70:225–244, 2014.

[143] S. H. Jacobson, E. C. Sewell, R. A. Proano, and J. A. Jokela. Stockpile levels for pediatric vaccines: How much is enough? *Vaccine*, 24(17):3530–3537, 2006.

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

[145] L. Jiang and R. E. Giachetti. A queueing network model to analyze the impact of parallelization of care on patient cycle time. *Health care management science*, 11:248–261, 2008.

[146] S. S. Jones, A. Thomas, R. S. Evans, S. J. Welch, P. J. Haug, and G. L. Snow. Forecasting daily patient volumes in the emergency department. *Academic Emergency Medicine*, 15(2):159–170, 2008.

[147] K. S. Jung, M. Pinedo, C. Sriskandarajah, and V. Tiwari. Scheduling elective surgeries with emergency patients at shared operating rooms. *Production and Operations Management*, 28(6):1407–1430, 2019.

[148] G. C. Kaandorp and G. Koole. Optimal outpatient appointment scheduling. *Health care management science*, 10:217–229, 2007.

[149] E. H. Kaplan and L. M. Wein. Decision making for bioterror preparedness: Examples from smallpox vaccination policy. *Operations Research and Health Care: A Handbook of Methods and Applications*, pages 519–536, 2004.

[150] D. R. Karger and C. Stein. A new approach to the minimum cut problem. *Journal of the ACM (JACM)*, 43(4):601–640, 1996.

[151] R. Khaldi, A. El Afia, R. Chiheb, and R. Faizi. Artificial neural network based approach for blood demand forecasting: Fez transfusion blood center case study. In *Proceedings of the 2nd international Conference on Big Data, Cloud and Applications*, pages 1–6, 2017.

[152] A. Khatoon. A blockchain-based smart contract system for healthcare management. *Electronics*, 9(1):94, 2020.

[153] P. Khodabandeh, V. Kayvanfar, M. Rafiee, and F. Werner. A bi-objective home health care routing and scheduling model with considering nurse downgrading costs. *International Journal of Environmental Research and Public Health*, 18(3):900, 2021.

[154] K. Kim and S. Mehrotra. A two-stage stochastic integer programming approach to integrated staffing and scheduling with application to nurse management. *Operations Research*, 63(6):1431–1451, 2015.

[155] A. Kokangul. A combination of deterministic and stochastic approaches to optimize bed capacity in a hospital unit. *Computer methods and programs in biomedicine*, 90(1):56–65, 2008.

[156] A. Kokangul, S. Akcan, and M. Narli. Optimizing nurse capacity in a teaching hospital neonatal intensive care unit. *Health care management science*, 20:276–285, 2017.

[157] Q. Kong, S. Li, N. Liu, C.-P. Teo, and Z. Yan. Appointment scheduling under time-dependent patient no-show behavior. *Management Science*, 66(8):3480–3500, 2020.

[158] M. L De Grano, D. Medeiros, and D. Eitel. Accommodating individual preferences in nurse scheduling via auctions and optimization. *Health Care Management Science*, 12:228–242, 2009.

[159] M. Lamiri, X. Xie, A. Dolgui, and F. Grimaud. A stochastic model for operating room planning with elective and emergency demand for surgery. *European Journal of Operational Research*, 185(3):1026–1037, 2008.

[160] G. Laporte, S. Nickel, and F. S. da Gama. *Location Science*, volume 528. Springer, 2019.

[161] A. Legrain, H. Bouarab, and N. Lahrichi. The nurse scheduling problem in real-life. *Journal of medical systems*, 39:1–11, 2015.

[162] Y. Li, Y. Zhang, N. Kong, and M. Lawley. Capacity planning for long-term care networks. *IIE Transactions*, 48(12):1098–1111, 2016.

[163] J. Lim, B. A. Norman, and J. Rajgopal. Redesign of vaccine distribution networks. *International Transactions in Operational Research*, 29(1):200–225, 2022.

[164] N. Liu, V.-A. Truong, X. Wang, and B. R. Anderson. Integrated scheduling and capacity planning with considerations for patients' length-of-stays. *Production and operations management*, 28(7):1735–1756, 2019.

[165] W. Liu, G. Y. Ke, J. Chen, and L. Zhang. Scheduling the distribution of blood products: A vendor-managed inventory routing approach. *Transportation Research Part E: Logistics and Transportation Review*, 140:101964, 2020.

[166] Y. Liu, A. M. Roshandeh, Z. Li, K. Kepaptsoglou, H. Patel, and X. Lu. Heuristic approach for optimizing emergency medical services in road safety within large urban networks. *Journal of transportation engineering*, 140(9):04014043, 2014.

[167] W. S. Lovejoy and Y. Li. Hospital operating room capacity expansion. *Management Science*, 48(11):1369–1387, 2002.

[168] Y. Y. Luo and B. Wang. A new method of block allocation used in two-stage operating rooms scheduling. *IEEE Access*, 7:102820–102831, 2019.

[169] X. Ma, A. Sauré, M. L. Puterman, M. Taylor, and S. Tyldesley. Capacity planning and appointment scheduling for new patient oncology consults. *Health care management science*, 19:347–361, 2016.

[170] G. Macrina, L. Di Puglia Pugliese, F. Guerriero, and G. Laporte. Drone-aided routing: A literature review. *Transportation Research Part C: Emerging Technologies*, 120:102762, 2020.

[171] M. B. Mandell. Covering models for two-tiered emergency medical services systems. *Location Science*, 6(1-4):355–368, 1998.

[172] D. S. Mankowska, F. Meisel, and C. Bierwirth. The home health care routing and scheduling problem with interdependent services. *Health care management science*, 17:15–30, 2014.

[173] V. Marianov and C. Revelle. The queuing probabilistic location set covering problem and some extensions. *Socio-Economic Planning Sciences*, 28(3):167–178, 1994.

[174] V. Marianov and C. ReVelle. The queueing maximal availability location problem: a model for the siting of emergency vehicles. *European Journal of Operational Research*, 93(1):110–120, 1996.

[175] V. Marianov and P. Taborga. Optimal location of public health centres which provide free and paid services. *Journal of the Operational Research Society*, 52(4):391–400, 2001.

[176] I. Marques, M. E. Captivo, and M. V. Pato. An integer programming approach to elective surgery scheduling. *OR spectrum*, 34(2):407–427, 2012.

[177] I. Marques, M. E. Captivo, and M. V. Pato. A bicriteria heuristic for an elective surgery scheduling problem. *Health care management science*, 18(3):251–266, 2015.

[178] A. Masone, S. Poikonen, and B. L. Golden. The multivisit drone routing problem with edge launches: An iterative approach with discrete and continuous improvements. *Networks*, 2022.

[179] C. Mateus, I. Marques, and M. E. Captivo. Local search heuristics for a surgical case assignment problem. *Operations Research for Health Care*, 17:71–81, 2018.

[180] L. Mayhew and D. Smith. Using queuing theory to analyse the government's 4-h completion time target in accident and emergency departments. *Health care management science*, 11:11–21, 2008.

[181] D. I. McIsaac, K. Abdulla, H. Yang, S. Sundaresan, P. Doering, S. G. Vaswani, K. Thavorn, and A. J. Forster. Association of delay of urgent or emergency surgery with mortality and use of health care resources: a propensity score–matched observational cohort study. *Cmaj*, 189(27):E905–E912, 2017.

[182] L. A. McLay. A maximum expected covering location model with two types of servers. *IIE Transactions*, 41(8):730–741, 2009.

[183] R. Mendoza-Gómez and R. Z. Ríos-Mercado. Location of primary health care centers for demand coverage of complementary services. *Computers & Industrial Engineering*, 169:108237, 2022.

[184] M. Meneses, D. Santos, and A. Barbosa-Póvoa. Modelling the blood supply chain–from strategic to tactical decisions. *European Journal of Operational Research*, 2022.

[185] N. Meskens, D. Duvivier, and A. Hanset. Multi-objective operating room scheduling considering desiderata of the surgical team. *Decision Support Systems*, 55(2):650–659, 2013.

[186] H. Miao and J.-J. Wang. Scheduling elective and emergency surgeries at shared operating rooms with emergency uncertainty and waiting time limit. *Computers & Industrial Engineering*, 160:107551, 2021.

[187] H. Miao and J.-J. Wang. Scheduling elective and emergency surgeries at shared operating rooms with emergency uncertainty and waiting time limit. *Computers & Industrial Engineering*, 160:107551, 2021.

[188] P. Miç, M. Koyuncu, and J. Hallak. Primary health care center (phcc) location-allocation with multi-objective modelling: a case study in idleb, syria. *International journal of environmental research and public health*, 16(5):811, 2019.

[189] Ministero della Salute. Servizi trasfusionali. `https://www.salute.gov.it/portale/temi/p2_6.jsp?lingua=italiano&id=2949&area=sangueTrasfusioni&menu=trasfusionale`, 2021.

[190] P. Mitropoulos, E. Adamides, and I. Mitropoulos. Redesigning a network of primary healthcare centres using system dynamics simulation and optimisation. *Journal of the Operational Research Society*, 74(2):574–589, 2023.

[191] A. Mokrini, Y. Boulaksil, and A. Berrado. Modelling facility location problems in emerging markets: The case of the public healthcare sector in morocco. *Operations and Supply Chain Management: An International Journal*, 12(2):100–111, 2019.

[192] J. Molema, S. Groothuis, I. Baars, M. Kleinschiphorst, E. Leers, A. Hasman, and G. Van Merode. Healthcare system design and parttime working doctors. *Health Care Management Science*, 10:365–371, 2007.

[193] M. Momenitabar, Z. Dehdari Ebrahimi, M. Arani, and J. Mattson. Robust possibilistic programming to design a closed-loop blood supply chain network considering service-level maximization and lateral resupply. *Annals of Operations Research*, pages 1–43, 2022.

[194] P. M. Morse, G. E. Kimball, and S. I. Gass. *Methods of operations research*. Courier Corporation, 2003.

[195] S. Moslemi and S. H. R. Pasandideh. A location-allocation model for quality-based blood supply chain under ier uncertainty. *RAIRO-Operations Research*, 55:S967–S998, 2021.

[196] M. Moz and M. Pato. A bi-objective network flow approach for nurse rerostering. In *Proceedings of the INOC*, pages 817–824, 2005.

[197] C. Mullinax and M. Lawley. Assigning patients to nurses in neonatal intensive care. *Journal of the operational research society*, 53(1):25–35, 2002.

[198] L. Murawski and R. L. Church. Improving accessibility to rural health services: The maximal covering network improvement problem. *Socio-Economic Planning Sciences*, 43(2):102–110, 2009.

[199] C. C. Murray and A. G. Chu. The flying sidekick traveling salesman problem: Optimization of drone-assisted parcel delivery. *Transportation Research Part C: Emerging Technologies*, 54:86 – 109, 2015.

[200] C. C. Murray and R. Raj. The multiple flying sidekicks traveling salesman problem: Parcel delivery with multiple drones. *Transportation Research Part C: Emerging Technologies*, 110:368–398, 2020.

[201] C. Myers and T. Green. Forecasting demand and capacity requirements. *Healthcare Financial Management*, 58(8):34–38, 2004.

[202] W. Najy, C. Archetti, and A. Diabat. Collaborative truck-and-drone delivery for inventory-routing problems. *Transportation Research Part C: Emerging Technologies*, page 103791, 2022.

[203] National Institute ore Health and Care Excellence. The use of routine preoperative tests for elective surgery. https://www.nice.org.uk/guidance/NG45/documents/preoperative-tests-update-final-scope2, 2021.

[204] M. Ndiaye and H. Alfares. Modeling health care facility location for moving population groups. *Computers & Operations Research*, 35(7):2154–2161, 2008.

[205] F. Niakan and M. Rahimi. A multi-objective healthcare inventory routing problem; a fuzzy possibilistic approach. *Transportation research part E: logistics and transportation review*, 80:74–94, 2015.

[206] S. N. Ogulata, M. Koyuncu, and E. Karakas. Personnel and patient scheduling in the high demanded hospital services: a case study in the physiotherapy service. *Journal of medical systems*, 32:221–228, 2008.

[207] N. Oladzad-Abbasabady and R. Tavakkoli-Moghaddam. Dynamic routing-scheduling problem for home health care considering caregiver-patient compatibility. *Computers & Operations Research*, 148:106000, 2022.

[208] M. D. Oliveira and G. Bevan. Modelling hospital costs to produce evidence for policies that promote equity and efficiency. *European Journal of Operational Research*, 185(3):933–947, 2008.

[209] A. F. Osorio, S. C. Brailsford, and H. K. Smith. A structured review of quantitative models in the blood supply chain: a taxonomic framework for decision-making. *International Journal of Production Research*, 53(24):7191–7212, 2015.

[210] A. F. Osorio, S. C. Brailsford, and H. K. Smith. Whole blood or apheresis donations? a multi-objective stochastic optimization approach. *European Journal of Operational Research*, 266(1):193–204, 2018.

[211] I. Ozkarahan. Allocation of surgeries to operating rooms by goal programing. *Journal of Medical Systems*, 24:339–378, 2000.

[212] PassMark Software. Cpu benchmarks. `http://www.cpubenchmark.net`, April 04, 2023.

[213] M. V. Pato and M. Moz. Solving a bi-objective nurse rerostering problem by using a utopic pareto genetic heuristic. *Journal of Heuristics*, 14:359–374, 2008.

[214] J. Patrick and M. L. Puterman. Improving resource utilization for diagnostic services through flexible inpatient scheduling: A method for

improving resource utilization. *Journal of the Operational Research Society*, 58:235–245, 2007.

[215] J. Patrick, M. L. Puterman, and M. Queyranne. Dynamic multipriority patient scheduling for a diagnostic resource. *Operations research*, 56(6):1507–1525, 2008.

[216] M. Persson and J. A. Persson. Health economic modeling to support surgery management at a swedish hospital. *Omega*, 37(4):853–863, 2009.

[217] D.-N. Pham and A. Klinkert. Surgical case scheduling as a generalized job shop scheduling problem. *European Journal of Operational Research*, 185(3):1011–1025, 2008.

[218] W. P. Pierskalla. Supply chain management of blood banks. *Operations research and health care: A handbook of methods and applications*, pages 103–145, 2004.

[219] W. P. Pierskalla and D. J. Brailer. Applications of operations research in health care delivery. *Handbooks in operations research and management science*, 6:469–505, 1994.

[220] A. Pirabán, W. J. Guerrero, and N. Labadie. Survey on blood supply chain management: Models and methods. *Computers & Operations Research*, 112:104756, 2019.

[221] S. Poikonen and B. Golden. Multi-visit drone routing problem. *Computers & Operations Research*, 113:104802, 2020.

[222] S. Poikonen, A. Masone, and B. Golden. A framework to transform truck-and-drone coordination problems into traveling salesman problems. *Computers & Operations Research (under review)*, 2020.

[223] J. Puente, A. Gómez, I. Fernández, and P. Priore. Medical doctor rostering problem in a hospital emergency department by means of genetic algorithms. *Computers & Industrial Engineering*, 56(4):1232–1242, 2009.

[224] P. Punnakitikashem, J. M. Rosenberger, and D. Buckley Behan. Stochastic programming for nurse assignment. *Computational Optimization and Applications*, 40:321–349, 2008.

[225] X. Quan, Y. Pang, J. Chen, X. Chu, and L. Shangguan. Open pollution routing problem of logistics distribution in medical union based on differential search algorithm. *Scientific Reports*, 12(1):19472, 2022.

[226] I. Rahimi and A. H. Gandomi. A comprehensive review and analysis of operating room and surgery scheduling. *Archives of Computational Methods in Engineering*, 28(3):1667–1688, 2021.

[227] I. Rahimi and A. H. Gandomi. A comprehensive review and analysis of operating room and surgery scheduling. *Archives of Computational Methods in Engineering*, 28(3):1667–1688, 2021.

[228] E. Rahimian, K. Akartunalı, and J. Levine. A hybrid integer and constraint programming approach to solve nurse rostering problems. *Computers & Operations Research*, 82:83–94, 2017.

[229] E. Rahimian, K. Akartunalı, and J. Levine. A hybrid integer programming and variable neighbourhood search algorithm to solve nurse rostering problems. *European Journal of Operational Research*, 258(2):411–423, 2017.

[230] S. Rahman and D. K. Smith. Deployment of rural health facilities in a developing country. *Journal of the Operational Research Society*, 50(9):892–902, 1999.

[231] A. Rais and A. Viana. Operations research in healthcare: a survey. *International transactions in Operational Research*, 18(1):1–31, 2011.

[232] D. Rameshwar, G. Angappa, and P. Thanos. *Application of Operations Research (OR) in Disaster Relief Operations (DRO), Part I and Part II*, volume 283, issue 1-2. Bantam, 2019.

[233] R. Ramezanian and Z. Behboodi. Blood supply chain network design under uncertainties in supply and demand considering social aspects. *Transportation Research Part E: Logistics and Transportation Review*, 104:69–82, 2017.

[234] A. Rasekh and K. Brumbelow. A dynamic simulation–optimization model for adaptive management of urban water distribution system contamination threats. *Applied Soft Computing*, 32:59–71, 2015.

[235] E. Rashidzadeh, S. M. Hadji Molana, R. Soltani, and A. Hafezalkotob. Assessing the sustainability of using drone technology for last-mile delivery in a blood supply chain. *Journal of Modelling in Management*, 16(4):1376–1402, 2021.

[236] M. Rastegar, M. Tavana, A. Meraj, and H. Mina. An inventory-location optimization model for equitable influenza vaccine distribution in developing countries during the covid-19 pandemic. *Vaccine*, 39(3):495–504, 2021.

[237] M. S. Rauner, M. Kraus, and S. Schwarz. Competition under different reimbursement systems: The concept of an internet-based hospital management game. *European Journal of Operational Research*, 185(3):948–963, 2008.

[238] M. Restrepo, S. G. Henderson, and H. Topaloglu. Erlang loss models for the static deployment of ambulances. *Health care management science*, 12:67–79, 2009.

[239] C. ReVelle. Review, extension and prediction in emergency service siting models. *European Journal of Operational Research*, 40(1):58–69, 1989.

[240] C. ReVelle and K. Hogan. A reliability-constrained siting model with local estimates of busy fractions. *Environment and Planning B: Planning and Design*, 15(2):143–152, 1988.

[241] C. ReVelle and K. Hogan. The maximum availability location problem. *Transportation science*, 23(3):192–200, 1989.

[242] J. Ridge, S. Jones, M. Nielsen, and A. Shahani. Capacity planning for intensive care units. *European journal of operational research*, 105(2):346–355, 1998.

[243] L. Righi, S. Trapassi, and C. Ramacciani Isemann. L'implementazione del nuovo sistema di triage in toscana. *L'infermiere*, 57:1:9–12, 2020.

[244] A. Riise and E. K. Burke. Local search for the surgery admission planning problem. *Journal of Heuristics*, 17(4):389–414, 2011.

[245] A. Riise, C. Mannino, and E. K. Burke. Modelling and solving generalised operational surgery scheduling problems. *Computers & Operations Research*, 66:1–11, 2016.

[246] R. Roberti and M. Ruthmair. Exact methods for the traveling salesman problem with drone. *Optimization Online*, 2019.

[247] R. Roberti and M. Ruthmair. Exact methods for the traveling salesman problem with drone. *Transportation Science*, 55(2):315–335, 2021.

[248] D. Rojas Viloria, E. L. Solano-Charris, A. Muñoz-Villamizar, and J. R. Montoya-Torres. Unmanned aerial vehicles/drones in vehicle routing problems: a literature review. *International Transactions in Operational Research*, 2020.

[249] B. Roland, C. Di Martinelly, F. Riane, and Y. Pochet. Scheduling an operating theatre under human resource constraints. *Computers & Industrial Engineering*, 58(2):212–220, 2010.

[250] V. Roshanaei, K. E. Booth, D. M. Aleman, D. R. Urbach, and J. C. Beck. Branch-and-check methods for multi-level operating room planning and scheduling. *International Journal of Production Economics*, 220:107433, 2020.

[251] A. E. Roth, T. Sönmez, and M. U. Ünver. Kidney exchange. *The Quarterly journal of economics*, 119(2):457–488, 2004.

[252] D. H. Rothstein and M. V. Raval. Operating room efficiency. In *Seminars in pediatric surgery*, volume 27, pages 79–85. Elsevier, 2018.

[253] G. Şahin, H. Süral, and S. Meral. Locational analysis for regionalization of turkish red crescent blood services. *Computers & Operations Research*, 34(3):692–704, 2007.

[254] M. R. G. Samani, S.-M. Hosseini-Motlagh, and S. F. Ghannadpour. A multilateral perspective towards blood network design in an uncertain environment: Methodology and implementation. *Computers & Industrial Engineering*, 130:450–471, 2019.

[255] M. Samorani and L. R. LaGanga. Outpatient appointment scheduling given individual day-dependent no-show predictions. *European Journal of Operational Research*, 240(1):245–257, 2015.

[256] M. Samudra, C. Van Riet, E. Demeulemeester, B. Cardoen, N. Vansteenkiste, and F. E. Rademakers. Scheduling operating rooms: achievements, challenges and pitfalls. *Journal of scheduling*, 19:493–525, 2016.

[257] G. D. Sanders, A. M. Bayoumi, V. Sundaram, S. P. Bilir, C. P. Neukermans, C. E. Rydzak, L. R. Douglass, L. C. Lazzeroni, M. Holodniy, and D. K. Owens. Cost-effectiveness of screening for hiv in the era of highly active antiretroviral therapy. *New England Journal of Medicine*, 352(6):570–585, 2005.

[258] B. Sandıkçı, L. M. Maillart, A. J. Schaefer, O. Alagoz, and M. S. Roberts. Estimating the patient's price of privacy in liver transplantation. *Operations Research*, 56(6):1393–1410, 2008.

[259] P. Santibáñez, M. Begen, and D. Atkins. Surgical block scheduling in a system of hospitals: an application to resource and wait list management in a british columbia health authority. *Health care management science*, 10:269–282, 2007.

[260] H. G. Santos, T. A. Toffolo, R. A. Gomes, and S. Ribas. Integer programming techniques for the nurse rostering problem. *Annals of Operations Research*, 239:225–251, 2016.

[261] S. K. Savaşer and B. Y. Kara. Mobile healthcare services in rural areas: an application with periodic location routing problem. *Or Spectrum*, 44(3):875–910, 2022.

[262] D. Schermer, M. Moeini, and O. Wendt. A matheuristic for the vehicle routing problem with drones and its variants. *Transportation Research Part C: Emerging Technologies*, 106:166–204, 2019.

[263] D. Schermer, M. Moeini, and O. Wendt. A branch-and-cut approach and alternative formulations for the traveling salesman problem with drone. *Networks*, 76(2):164–186, 2020.

[264] D. Schilling, D. J. Elzinga, J. Cohon, R. Church, and C. ReVelle. The team/fleet models for simultaneous facility and equipment siting. *Transportation science*, 13(2):163–175, 1979.

[265] D. L. Segev, S. E. Gentry, D. S. Warren, B. Reeb, and R. A. Montgomery. Kidney paired donation and optimizing the use of live donor organs. *Jama*, 293(15):1883–1890, 2005.

[266] X. Shang, G. Zhang, B. Jia, and M. Almanaseer. The healthcare supply location-inventory-routing problem: A robust approach. *Transportation Research Part E: Logistics and Transportation Review*, 158:102588, 2022.

[267] H. D. Sherali, M. H. Ramahi, and Q. J. Saifee. Hospital resident scheduling problem. *Production Planning & Control*, 13(2):220–233, 2002.

[268] Y. Shi, T. Boudouh, and O. Grunder. A robust optimization for a home health care routing and scheduling problem with consideration of uncertain travel and service times. *Transportation Research Part E: Logistics and Transportation Review*, 128:52–95, 2019.

[269] H. Shih and S. Rajendran. Comparison of time series methods and machine learning algorithms for forecasting taiwan blood services foundation's blood supply. *Journal of healthcare engineering*, 2019, 2019.

[270] A. Shmueli, C. L. Sprung, and E. H. Kaplan. Optimizing admissions to an intensive care unit. *Health Care Management Science*, 6:131–136, 2003.

[271] T. A. Silva and M. C. de Souza. Surgical scheduling under uncertainty by approximate dynamic programming. *Omega*, 95:102066, 2020.

[272] T. A. Silva, M. C. de Souza, R. R. Saldanha, and E. K. Burke. Surgical scheduling with simultaneous employment of specialised human resources. *European Journal of Operational Research*, 245(3):719–730, 2015.

[273] M. E. Singer and Z. M. Younossi. Cost effectiveness of screening for hepatitis c virus in asymptomatic, average-risk adults. *The American journal of medicine*, 111(8):614–621, 2001.

[274] H. K. Smith, P. R. Harper, C. N. Potts, and A. Thyle. Planning sustainable community health schemes in rural areas of developing countries. *European Journal of Operational Research*, 193(3):768–777, 2009.

[275] V. L. Smith-Daniels, S. B. Schweikhart, and D. E. Smith-Daniels. Capacity management in health care services: Review and future research directions. *Decision Sciences*, 19(4):889–919, 1988.

[276] S. Srinivas and A. R. Ravindran. Optimizing outpatient appointment system using machine learning algorithms and scheduling rules: A prescriptive analytics framework. *Expert Systems with Applications*, 102:245–261, 2018.

[277] J. Steeg and M. Schröder. A hybrid approach to solve the periodic home health care problem. In *Operations Research Proceedings 2007: Selected Papers of the Annual International Conference of the German Operations Research Society (GOR) Saarbrücken, September 5–7, 2007*, pages 297–302. Springer, 2008.

[278] K. Stein, K. Dalziel, A. Walker, B. Jenkins, A. Round, and P. Royle. Screening for hepatitis c in injecting drug users: a cost utility analysis. *Journal of Public Health*, 26(1):61–71, 2004.

[279] J. E. Storbeck. Slack, natural slack, and location covering. *Socio-Economic Planning Sciences*, 16(3):99–105, 1982.

[280] C. Stummer, K. Doerner, A. Focke, and K. Heidenberger. Determining location and size of medical departments in a hospital network: A multi-objective decision support approach. *Health care management science*, 7:63–71, 2004.

[281] L. Subramanian. Effective demand forecasting in health supply chains: emerging trend, enablers, and blockers. *Logistics*, 5(1):12, 2021.

[282] D. Sundaramoorthi, V. C. Chen, J. M. Rosenberger, S. B. Kim, and D. F. Buckley-Behan. A data-integrated simulation model to evaluate nurse–patient assignments. *Health care management science*, 12:252–268, 2009.

[283] J. Tang and Y. Wang. An adjustable robust optimisation method for elective and emergency surgery capacity allocation with demand uncertainty. *International Journal of Production Research*, 53(24):7317–7328, 2015.

[284] F. Tanser. Methodology for optimising location of new primary health care facilities in rural communities: a case study in kwazulu-natal, south africa. *Journal of Epidemiology & Community Health*, 60(10):846–850, 2006.

[285] E. B. Tirkolaee, H. Golpîra, A. Javanmardan, and R. Maihami. A socio-economic optimization model for blood supply chain network design during the covid-19 pandemic: An interactive possibilistic programming approach for a real case study. *Socio-Economic Planning Sciences*, 85:101439, 2023.

[286] S. Topaloglu. A multi-objective programming model for scheduling emergency medicine residents. *Computers & Industrial Engineering*, 51(3):375–388, 2006.

[287] C. Toregas, R. Swain, C. ReVelle, and L. Bergman. The location of emergency service facilities. *Operations research*, 19(6):1363–1373, 1971.

[288] A. M. Turhan and B. Bilgen. A hybrid fix-and-optimize and simulated annealing approaches for nurse rostering problem. *Computers & Industrial Engineering*, 145:106531, 2020.

[289] United Nations. World population ageing 2015. https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2015_Report.pdf, 2015.

[290] C. Valouxis and E. Housos. Hybrid optimization techniques for the workshift and rest assignment of nursing personnel. *Artificial Intelligence in Medicine*, 20(2):155–175, 2000.

[291] P. Van den Berg, G. Kommer, and B. Zuzáková. Linear formulation for the maximum expected coverage location model with fractional coverage. *Operations Research for Health Care*, 8:33–41, 2016.

[292] J. T. van Essen, E. W. Hans, J. L. Hurink, and A. Oversberg. Minimizing the waiting time for emergency surgery. *Operations research for health care*, 1(2-3):34–44, 2012.

[293] M. L. van Genugten, M.-L. A. Heijnen, and J. C. Jager. Pandemic influenza and healthcare demand in the netherlands: scenario analysis. *Emerging Infectious Diseases*, 9(5):531, 2003.

[294] C. Van Riet and E. Demeulemeester. Trade-offs in operating room planning for electives and emergencies: A review. *Operations Research for Health Care*, 7:52–69, 2015.

[295] P. T. VanBerkel and J. T. Blake. A comprehensive simulation for wait time reduction and capacity planning applied in general surgery. *Health care management Science*, 10:373–385, 2007.

[296] M. Veenstra, K. J. Roodbergen, L. C. Coelho, and S. X. Zhu. A simultaneous facility location and vehicle routing problem arising in health care

logistics in the netherlands. *European Journal of Operational Research*, 268(2):703–715, 2018.

[297] V. Verter and S. D. Lapierre. Location of preventive health care facilities. *Annals of Operations Research*, 110:123–132, 2002.

[298] B. Vijayakumar, P. J. Parikh, R. Scott, A. Barnes, and J. Gallimore. A dual bin-packing approach to scheduling surgical cases at a publicly-funded hospital. *European Journal of Operational Research*, 224(3):583–591, 2013.

[299] C.-W. Wang, L.-M. Sun, M.-H. Jin, C.-J. Fu, L. Liu, C.-H. Chan, and C.-Y. Kao. A genetic algorithm for resident physician scheduling problem. In *Proceedings of the 9th annual conference on Genetic and evolutionary computation*, pages 2203–2210, 2007.

[300] Y. Wang, Z. Wang, X. Hu, G. Xue, and X. Guan. Truck–drone hybrid routing problem with time-dependent road travel time. *Transportation Research Part C: Emerging Technologies*, 144:103901, 2022.

[301] C. Wanying, G. Alain, and R. Angel. Modeling the logistics response to a bioterrorist anthrax attack. *European Journal of Operational Research*, 254(2):458–471, 2016.

[302] World-Health-Organization. Public spending on health: A closer look at global trends. https://apps.who.int/nha/database/DocumentationCentre/GetFile/577 20661/en, 2018.

[303] World Health Organization. 10 facts on blood transfusion. `https://www.who.int/bloodsafety/FactFile2009.pdf`, 2021.

[304] World Health Organization et al. Regional status report on blood safety and availability 2016. Technical report, World Health Organization. Regional Office for the Eastern Mediterranean, 2017.

[305] J. T. Wu, L. M. Wein, and A. S. Perelson. Optimization of influenza vaccine selection. *Operations Research*, 53(3):456–476, 2005.

[306] T.-H. Wu, J.-Y. Yeh, and Y.-M. Lee. A particle swarm optimization approach with refinement procedure for nurse rostering problem. *Computers & Operations Research*, 54:52–63, 2015.

[307] G. Wullink, M. Van Houdenhoven, E. W. Hans, J. M. Van Oostrum, M. Van Der Lans, and G. Kazemier. Closing emergency operating rooms improves efficiency. *Journal of medical systems*, 31:543–546, 2007.

[308] J. L. Xue, J. Z. Ma, T. A. Louis, and A. J. Collins. Forecast of the number of patients with end-stage renal disease in the united states to the year 2010. *Journal of the American Society of Nephrology*, 12(12):2753–2758, 2001.

[309] M. Yang, Y. Ni, and L. Yang. A multi-objective consistent home healthcare routing and scheduling problem in an uncertain environment. *Computers & Industrial Engineering*, 160:107560, 2021.

[310] M. Younespour, A. Atighehchian, K. Kianfar, and E. T. Esfahani. Using mixed integer programming and constraint programming for operating rooms scheduling with modified block strategy. *Operations Research for Health Care*, 23:100220, 2019.

[311] E. Yurek and H. Ozmutlu. A decomposition-based iterative optimization algorithm for traveling salesman problem with drone. *Transportation Research Part C: Emerging Technologies*, 91:249–262, 2018.

[312] B. Zahiri and M. S. Pishvaee. Blood supply chain network design considering blood group compatibility under uncertainty. *International Journal of Production Research*, 55(7):2013–2033, 2017.

[313] B. Zahiri, S. Torabi, M. Mousazadeh, and S. Mansouri. Blood collection management: Methodology and application. *Applied Mathematical Modelling*, 39(23-24):7680–7696, 2015.

[314] S. A. Zenios. Optimal control of a paired-kidney exchange program. *Management Science*, 48(3):328–342, 2002.

[315] S. A. Zenios. Models for kidney allocation. *Operations research and health care: A handbook of methods and applications*, pages 537–554, 2004.

[316] B. Zhang, P. Murali, M. Dessouky, and D. Belson. A mixed integer programming approach for allocating operating room capacity. *Journal of the Operational Research Society*, 60(5):663–673, 2009.

[317] J. Zhang, M. Dridi, and A. El Moudni. Column-generation-based heuristic approaches to stochastic surgery scheduling with downstream capacity constraints. *International Journal of Production Economics*, 229:107764, 2020.

[318] Y. Zhang, O. Berman, and V. Verter. Incorporating congestion in preventive healthcare facility network design. *European Journal of Operational Research*, 198(3):922–935, 2009.

[319] J. Zhao, T. Wang, and T. Monteiro. A bi-objective home health care routing and scheduling problem under uncertainty. *Available at SSRN 4503329*, 2023.

[320] S. Zhu, W. Fan, T. Liu, S. Yang, and P. M. Pardalos. Dynamic three-stage operating room scheduling considering patient waiting time and surgical overtime costs. *Journal of Combinatorial Optimization*, 39(1):185–215, 2020.

[321] S. Zhu, W. Fan, S. Yang, J. Pei, and P. M. Pardalos. Operating room planning and surgical case scheduling: a review of literature. *Journal of Combinatorial Optimization*, 37(3):757–805, 2019.

[322] N. Zinouri, K. M. Taaffe, and D. M. Neyens. Modelling and forecasting daily surgical case volume using time series analysis. *Health Systems*, 7(2):111–119, 2018.