Master Thesis Business Analytics

Optimization of Ward Layouts for Clinical and Day Treatment Departments

Author: Femke Stokkink (2666619)

Company supervisor:	Wouter Veneklaas
First supervisor:	René Bekker
Second reader:	Ger Koole

March 7, 2025

Optimization of Ward Layouts for Clinical and Day Treatment Departments

Femke Stokkink

Master Thesis

Vrije Universiteit Amsterdam

Faculty of Science

Business Analytics

De Boelelaan 1081a

1081 HV Amsterdam

Host Organisation:

ChipSoft

Orlyplein 10

1043 DP Amsterdam

March 7, 2025

Preface

This report is written for the Master Project Business Analytics at the Vrije Universiteit Amsterdam. The purpose of this study is to optimize the ward layouts for clinical and day treatment departments, while considering bed requirements, medical and patient-related constraints, and spatial limitations.

This research is conducted within the Capacity Management department at ChipSoft, a Dutch software company specializing in healthcare technology solutions. The Capacity Management team within ChipSoft aims to provide healthcare providers with data-driven insights that facilitate optimal decision-making on strategic, tactical, and operational levels.

I would like to give a special thank you to Wouter Veneklaas, my supervisor at ChipSoft, for providing me with the incredible opportunity to conduct this research at ChipSoft. His support throughout this project has been invaluable. Not only did he always show a genuine interest in my progress, but he also helped me get the very best out of this research. I am especially grateful for the willingness to help, whether it was providing insight, answering my questions, or simply offering reassurance during moments of doubt. The encouragement and dedication played a significant role in making this research a rewarding and enjoyable experience.

I would also like to extend my gratitude to the entire Capacity Management team at ChipSoft. I want to thank the team for not only supporting me with my research but also for making this experience so enjoyable. The fun distractions made it a pleasure to come into the office every day.

Finally, I would like to thank René Bekker, my supervisor from Vrije Universiteit Amsterdam. I am grateful for our insightful meetings and the valuable feedback. His expertise and knowledge were very helpful in my research. I also appreciate his enthusiasm for the project, which motivated me to explore new ideas.

Summary

This research will focus on optimizing the ward layout for clinical and day treatment departments by assigning specialties to available wards as efficiently as possible.

Research Structure We developed an Integer Linear Program (ILP) to allocate clusters of specialties to the available wards to optimize both clinical and day treatment ward layouts. Since each hospital has specific preferences for an efficient ward layout, we incorporate these preferences into the allocation constraints based on user input. We request the hospital to specify which specialties can be grouped together in a ward to ensure the resulting layout complies with medical restrictions. Based on these compatibilities among specialties, we can create feasible clusters of specialties. This means we compute all possible combinations of specialties that are compatible. Additionally, we inquire about specialties that must be assigned to particular wards and those that should be located near the Operating Room (OR). To minimize walking distances for staff, we also request the physical locations of the wards and the OR to calculate the distances between them.

We also ask the user to specify a desired blocking probability to ensure that sufficient bed capacity is available based on this probability. We compute the number of required beds for each possible cluster of beds. For this, we employed both the Erlang loss model and Discrete Event Simulation. The Erlang loss model assumes stationary arrivals, which does not hold for day treatment department. While there exists little to no research on determining bed requirements for day treatment departments, we developed a time-dependent Erlang loss model using the Modified Offered Load (MOL) approach. To account for seasonality and peak periods, we calculate bed requirements based on the busiest quarter of the year for clinical departments and the busiest hour of the week for day treatment departments.

The bed requirements and user input serve as input for the ILP which determines which specialties should be clustered together and to which wards they should be assigned. The objective is to minimize the total required number of beds while also limiting the size of the largest specialty cluster to prevent the model from grouping all specialties into a single cluster.

Key Findings The bed requirements models showed similar outcomes, with a maximum difference of mostly one or two beds. For both models, we observed that the combination of specialties into a single ward resulted in a reduction in the number of beds needed, attributed to the advantages of bed pooling. There was a significant difference in computation time, where the Erlang loss model is much faster than the Discrete Event Simulation. Despite the computational advantage, our preference lies with Discrete Event Simulation if the number of

clusters is limited as hospitals often have more faith in Discrete Event Simulation because it is more tangible in how the bed requirements are computed.

We conducted a case study with a collaborating hospital. The case hospital experienced particularly high blocking probability on one of their wards. The optimized clinical ward layout was able to decrease the blocking probability on that ward from 10.3% to 3.7% and thereby also reducing the average blocking probability. Additionally, it ensured that patient demand was more evenly distributed across the wards, leading to more balanced bed utilization rates. For the day treatment wards, we assumed full compatibility between specialties, except for oncology, and were able to find a new ward layout with lower blocking probabilities while using the same number of beds. However, ignoring different bed types led to an underestimation of required beds.

Conclusions and Recommendations The case study demonstrated the model's effectiveness in optimizing ward layouts by reducing blocking probabilities and more evenly distributing patient demand across the wards. The user input ensures that the hospital retains full control over the demands and preferences for the ward layout. However, prior to implementation, it is important to account for different bed types to prevent underestimations of bed requirements, particularly in day treatment departments. Future research could also explore integrating day and clinical departments and incorporating nursing staff requirements into the model.

Contents

1 Introduction

Hospitals today are facing an increasing demand for healthcare services, driven by an ageing and growing population. Addressing this increased demand becomes particularly challenging when hospitals are understaffed, underfunded, and lacking essential facilities (Burdett et al. (2024)). Consequently, effective capacity management has become critical for optimizing the use of available beds, staff, and resources. Inpatient care, where patients are admitted to the hospital for treatment, is a central component of this challenge. With limited bed availability, staff shortages, fluctuating patient demand, and complex medical requirements, hospitals must ensure that their ward layouts facilitate optimal patient flow and resource allocation.

1.1 Problem Statement

Given the increasing pressure on healthcare systems, it is crucial to explore strategies that improve the efficiency of hospital resources within clinical wards. A clinical ward is a designated area in a hospital where patients are accommodated, treated, and cared for, typically organized by specialty or medical need. It includes patient rooms and essential facilities, such as nurse stations, storage and medication rooms, and bathrooms. While the physical layout of these spaces is fixed and generally cannot be altered, there is significant potential for optimization in how wards are organized, particularly in the assignment of specialties to these wards.

For example, assigning surgical specialties, such as orthopedics or general surgery, to wards closer to the operating rooms (ORs) reduces patient transport time and enhances operational efficiency. Another way to improve efficiency is by allowing multiple specialties to share a ward. This approach leverages the benefits of pooling variability in patient arrivals and lengths of stay, which can reduce the overall capacity required. By combining specialties, hospitals can meet capacity demands more effectively while creating a more balanced workload, see for example Green and Nguyen (2001). By pooling beds across specialties, hospitals can reduce the risk of patient refusals and lower overall bed requirements. However, this must be done carefully, as placing incompatible specialties together, such as surgical recovery patients with those suffering from infectious diseases, could lead to healthcare risks. Medical limitations must therefore be considered when combining specialties. In some cases, combining specialties cannot be avoided, especially in smaller hospitals where dedicated wards for each specialty are not feasible due to the construction of the hospital.

Specialties can be classified based on the type of care provided, such as distinguishing between day treatment and clinical departments. For example, oncology patients may receive chemotherapy in a day treatment unit, while post-surgical recovery and long-term care occur in clinical departments. Whereas previous research has primarily focused on clinical departments, this study will also include day treatment departments.

This research will focus on optimizing the ward layout for clinical and day treatment departments by assigning specialties to available wards as efficiently as possible. The layout must comply with strict medical and patient-related constraints, as well as factors such as proximity to treatment facilities and department-specific preferences. Moreover, the allocation should ensure sufficient bed availability to meet a predefined blocking probability.

1.2 Motivation

Designing the clinical ward layout of a new hospital or reconfiguring existing ones is a complex task due to the intricate nature of healthcare facility layout design (Li et al. (2023)). This process requires aligning medical care processes with physical resources, while also considering operational efficiency and patient needs. Additionally, there are many constraints and potentially conflicting objectives, making the problem even more challenging.

Currently, clinical ward layouts are largely based on the experience and expertise of hospital planners and architects (Arnolds and Nickel (2013)). These layouts are shaped by daily operations, staffing arrangements, and the current layout of clinical wards, the operating theater, and other treatment facilities. However, Li et al. (2023) argues that relying solely on human experience is limiting, as it is difficult for individuals to fully analyze complex, multidimensional data. This can lead to a misalignment between the anticipated and actual workflows within the hospital. A mathematical model can explore far more configurations than manual methods, leading to more efficient solutions. Moreover, manually creating and evaluating ward layouts is not only time-consuming but also labor-intensive (Li et al. (2023)). Thus, by integrating data-driven insights with the expertise of planners, hospitals can make more informed and time-efficient strategic decisions in the design and reconfiguration of clinical ward layouts. Additionally, the use of capacity management modules within EHR systems can significantly enhance this process.

1.3 Research Questions

This research aims to develop a mathematical model that optimizes the allocation of specialties within clinical and day treatment departments to wards while considering bed requirements, medical and patient-related constraints, and spatial limitations. To this end, five distinct research questions are formulated:

1. How can the required number of beds for (a combination of) specialties within clinical departments be determined while maintaining a predefined blocking probability?

2. How can the required number of beds for (a combination of) specialties

within day treatment departments be determined while maintaining a predefined blocking probability?

3. What are the key constraints and requirements for effective assignment of (a combination of) specialties within clinical and day treatment departments to available wards and how can this be translated into a mathematical framework?

4. How can the defined mathematical framework be optimized to ensure efficient clinical ward layouts?

5. What are the potential applications of implementing a model to optimize clinical ward layout in real-world hospital settings?

1.4 Thesis Outline

The remainder of this report is structured as follows. Section 2 describes the context of the study, followed by a more detailed description of the problem in Section 3. Section 4 presents a review of relevant literature. Section 5 covers data analysis, detailing the patterns in patient arrivals, length of stay, occupancy rates, and the availability of beds and wards. Next, Section 6 explains the methodology used to develop models for determining bed requirements and assigning specialties to available wards. Section 7 outlines the results obtained from applying the proposed methodology. Section 8 includes a case study that applies the model to a specific hospital scenario, demonstrating its practical applications and impact. Finally, this thesis is concluded with the main findings in Section 9 and a discussion in Section 10.

2 Context Description

This section provides an overview of the context of the study. Section 2.1 reflects on the host organization. Section 2.2 addresses the patient admission process, distinguishing elective and emergency patients. This is followed by a description of excluded departments in Section 2.3 and a description of refused admissions in Section 2.4.

2.1 Host Organization

This research is conducted within the Capacity Management department at ChipSoft, a leading software company specializing in healthcare technology solutions. ChipSoft provides a fully integrated Hospital Information System (HIS) and Electronic Health Record (EHR) system for healthcare institutions across the entire healthcare chain, named HiX. By constantly exploring opportunities to make healthcare even more efficient, ChipSoft aims to optimally use scarce time and resources for the benefit of the patient.

The Capacity Management team within ChipSoft arose from the idea that the extensive data stored in HiX holds significant potential for providing insights into efficient hospital management. The team aims to provide healthcare providers with data-driven insights that facilitate optimal decision-making on strategic, tactical, and operational levels (Figure 1). Strategic planning deals with long-term, structural decisions that typically span a time period of 1 to 3 years. At this level, hospitals make high-level decisions that set the foundation for future operations. Tactical planning serves as a bridge between strategic decisions and day-to-day operations. It typically has a time horizon of several months to a year. Tactical decisions involve adjusting and organizing resources based on the strategic framework already established, such as temporarily expanding or reducing capacity in response to seasonal variations in patient arrivals. Operational planning focuses on the short-term, day-to-day execution of hospital activities. This level deals with detailed scheduling and resource allocation, often for weeks or days in advance. It includes specific tasks like scheduling staff shifts, assigning patients to beds, and managing daily admissions and discharges.

Optimizing the layout of clinical wards falls under strategic capacity management as it involves long-term, high-level decisions aimed at improving the overall functionality and efficiency of the hospital (Hans et al. (2012)). By researching the optimization of the clinical ward layout, we take a first step towards developing a practical solution, and this paves the way for future implementation into the strategic planning module in HiX, thereby enhancing the software Capacity Management has to offer.

Figure 1: Different levels of Capacity Management within ChipSoft

One of ChipSoft's clients expressed interest in this topic and was eager to collaborate with us on this research. Their willingness to participate provides valuable real-world insights into their challenges and needs. This partnership will enhance the relevance and applicability of the research findings.

2.2 Patient Admission Process

This section will discuss the patient admission and treatment flow within a hospital setting, as depicted in Figure 2. This research focuses only on inpatient care, so outpatient care is excluded from this figure. Patients arrive as either elective or emergency cases. Elective patients are those scheduled in advance, often for planned surgeries or treatments, allowing hospitals to prepare resources and beds accordingly. Emergency patients, however, arrive without prior notice due to urgent health issues, requiring immediate attention and often placing additional pressure on bed availability and hospital resources.

Elective patients are admitted to either a day treatment ward or a general ward. Day treatment wards are intended for patients needing only short stays, often for minor procedures or treatments that do not require an overnight stay. If a patient in a day treatment ward cannot be safely discharged by the end of the day, they are transferred to a general ward for overnight care. General wards provide longer-term care for patients requiring overnight or extended stays, typically involving a range of specialties and medical treatments. Elective patients admitted to a general ward may also be transferred to other units based on the course of their treatment. If they require surgery, they can be transferred from their ward to the Operating Room (OR) for a procedure. After surgery, they may either return to a general ward for recovery or, if necessary, be transferred to the Intensive Care Unit (ICU) for more intensive monitoring. Elective patients are either discharged from the general ward or transferred to the appropriate facility for further recovery.

Emergency patients enter the hospital through the Emergency Department (ED). If their condition is not severe, they are discharged directly. Since these patients do not require inpatient care, this flow is excluded from consideration. For those needing further treatment, patients from the ED are either admitted to a general ward, transferred to the ICU for critical care, or sent to the OR if immediate surgery is necessary. The path taken depends on the severity of the patient's condition and the urgency of the required medical intervention.

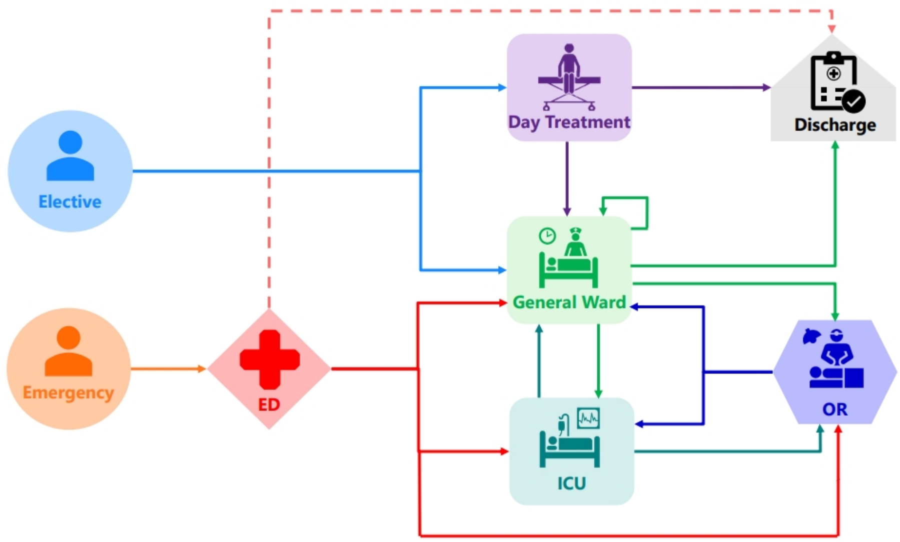


Figure 2: Patient admission and treatment flow within an inpatient hospital setting

Furthermore, in this study, the number of beds at a clinical or day treatment ward refers to the number of operational beds rather than the physical beds. Operational beds refer to the number of beds that are available for patient care, which is primarily determined by the availability of staff. While there may be physical beds present in a ward, if there is insufficient staff to attend to the patients, those beds cannot be utilized.

2.3 Excluded Departments

This study excludes the following departments due to their specialized care requirements, which fall outside the scope of this research:

- **Intensive Care Unit (ICU)**: Provides life-saving, intensive treatment and continuous monitoring for critically ill patients.
- **Acute Admissions Unit (AAU)**: Provides care for patients who require urgent hospital admission. Patients typically stay in the AAU for up to 48 hours. If a longer hospital stay is necessary, they are transferred to the nursing department of the admitting specialist.

- **Elective Admissions Unit**: Here patients are admitted before their scheduled treatment or surgery.
- **Coronary Care Unit (CCU)**: Specialized hospital ward dedicated to the care of patients with heart attacks, unstable angina, cardiac arrhythmias, and other cardiac conditions requiring continuous monitoring and treatment.
- **Dialysis Center**: Dedicated to providing regular dialysis treatments for patients with kidney failure.
- **Obstetric department**: Specialize in the care of patients during pregnancy, childbirth, and the postpartum period.
- **Pediatric department**: Focus on the treatment and care of all patients under 18 years old.
- **Neonatal department**: Provide care for newborns, particularly those born prematurely or with medical complications.

All these departments have distinct operational dynamics, staffing requirements, and resource allocations that set them apart from regular clinical and day treatment wards. For the purpose of this study, they are assumed to operate in fixed locations with a constant number of beds. Their exclusion allows the study to focus exclusively on optimizing the layout of regular clinical and day treatment wards. If a patient's admission includes both ICU and regular clinical ward stays, we exclude only the ICU portion.

2.4 Refused Admissions

Arriving patients are admitted to a ward if there is a free bed available. It is preferred to admit patients to the ward assigned to the department of their responsible medical specialist. However, if all beds are occupied, this is not possible. In this case, the arrival is defined as a refused admission. In practice, a refused admission can result in, for example, diversion to another hospital or transfer to another ward. A patient who is not placed in their preferred ward can be defined as a misallocated patient. Admitting a patient to a ward outside of their specialty poses risks, such as suboptimal care due to the staff's unfamiliarity with the specific medical needs and protocols of the patient's condition. Stylianou et al. (2017) found that patients admitted to an inappropriate ward have double the length of stay of patients admitted to the correct ward. We therefore want to keep the probability of refusing patients as low as possible. In other words, we want to minimize the blocking probability.

Measuring the historic blocking probability from available data is challenging. One approach is to identify misallocated patients and reassign them to the correct ward. This involves adjusting the ward information in the data for patients initially admitted to an incorrect unit, ensuring they are placed in the appropriate ward. The correction process considers factors such as specialism, admission type, hospital location, patient age, and diagnosis. By evaluating whether a patient was properly admitted to their assigned ward, the correction method detects misallocated patients and determines the correct ward based on these factors, ensuring proper placement. This way, the number of misallocated patients is equivalent to the number of refused patients. However, no data is available for patients who were refused admission, not transferred to another ward, but instead had their planned admission postponed or were transferred to a different hospital for example.

In this study, we assume that arriving patients who find all beds occupied are refused and leave the system to simplify the modeling process and focus on internal capacity constraints. This assumption is common in capacity planning and queuing models, where systems are treated as closed environments, and patients unable to access care are considered lost demand. While in reality patients will be transferred to other wards, this approach could slightly underestimate overall demand.

3 Problem Description

The problem at hand is to allocate specialties within clinical and day treatment departments to existing wards while considering bed requirements, medical and patient-related constraints, and spatial limitations. Figure 4 illustrates this challenge using a small example hospital. We split the problem into four components: physical locations, patient groups, bed requirements model, and ward assignment model. This section will discuss each component.

1. Physical Locations

The first crucial aspect of this problem is the physical layout of the hospital, depicted by the blue part of Figure 4. This includes:

- **Location of wards**: This refers to the positioning of wards within the hospital, such as which floor and wing they are located in, allowing for the calculation of distances between wards. When specialties are assigned to multiple wards, these wards should be situated close to each other to minimize walking distances for both staff and patients and thereby increase efficiency.
- **Bed availability**: To optimally assign specialties to wards, it is crucial to consider the available bed capacity in each ward. Each ward must meet a pre-defined blocking probability to ensure sufficient beds for the assigned specialties, minimizing the likelihood of patient refusals. The required number of beds should therefore not exceed the available beds.
- **Proximity to the operating room (OR)**: Surgical specialties should be located near the OR to minimize walking distances and improve operational flows. Therefore, the distances from each ward to the OR must be known.

2. Patient Groups and Restrictions

The second component of our problem is the patient groups, depicted by the pink part of Figure 4. Assigning specialties as a whole to wards is not always practical, as patients within a specialty may be distributed across different wards based on sub-specialties and the type of admission (day treatment or clinical). To clearly define patient groups in this study, we consider the following components:

- **Specialty**: The overarching medical discipline, such as gynecology.
- **Admission type**: Clinical or day treatment admission.
- **Sub-specialty** (optional): A more specific division within the specialty, such as gynecological oncology or obstetrics within the gynecology specialty.

For these patient groups, we need to consider medical and logistic restrictions. These restrictions are crucial because neglecting them can lead to a ward layout that lacks coherence, undermining hospitals' trust in the proposed solutions. Therefore, it is essential to carefully evaluate these elements to develop an effective model that proposes a good clinical ward layout.

Whether patient groups can share the same ward depends on several medical and patient compatibility constraints, which can be expressed in the form of a **compatibility matrix**. Medical compatibility ensures that patients with different medical conditions sharing the same ward do not pose any health risks to one another. For example, immunocompromised patients, such as those undergoing chemotherapy, should not be placed in the same ward as patients with infectious diseases, to avoid the risk of cross-contamination. Patient compatibility refers to social and emotional considerations like variations in the severity of medical conditions or age categories. For instance, merging a specialized ward for geriatric dementia patients with an orthopedic department could create tension in shared wards due to age differences and conflicting therapy plans. These medical and patient compatibility constraints must be carefully considered to ensure patient safety, which in turn contributes to improved quality of care (Mitchell (2008)). Furthermore, when multiple departments share a ward, nurses are expected to care for all patients in that ward, this may necessitate additional qualifications for the nursing staff. To minimize additional qualifications, we want to avoid combinations of specialties that require different types of care. This compatibility matrix enables us to form possible combinations of patient groups, which we call clusters. These clusters exclude all combinations that are incompatible according to the compatibility matrix.

To illustrate the process of cluster formation, Figure 3 presents the possible clusters for three specialties, assuming full compatibility among them. A model will be developed to determine the optimal clustering of specialties, ensuring that each specialty is assigned exactly once.

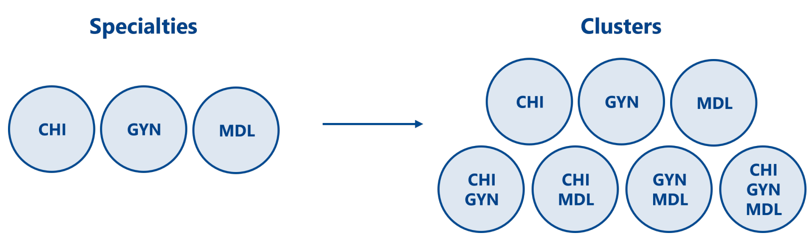


Figure 3: Illustration of possible clusters for three fullly compatible specialties

Patient groups may also have **location preferences**, which refers to the physical locations of the blue part of Figure 4. Surgical specialties like orthopedics need to be close to OR, this is not important for non-surgical specialties like psychology. Additionally, specialties might have a preference for a specific ward.

Cardiology might require a specialized ward for monitoring patients with heart conditions.

3. Bed Requirements Model

The possible combinations of specialties determined in the previous part will be part of the input for the next component, namely the bed requirements model (orange part in Figure 4). For each possible combination, we will determine the required number of beds based on a pre-specified blocking probability. Clusters can only be assigned to wards if the required number of beds does not exceed the available beds. Arrival patterns and length of stay (LoS) distributions, derived from admission data, will serve as inputs for a model to be determined and developed in Section 6.3. The output of the model will be the required number of beds per possible cluster.

4. Ward Assignment Model

Finally, all key components are combined to form the input for a model that determines the optimal clinical ward layout. The input includes the placement of wards, beds, and the OR within the hospital, along with patient groups and their location preferences and the required number of beds per patient group for each possible combination. The model's output will indicate the optimal combinations of patient groups and to which wards they are assigned.

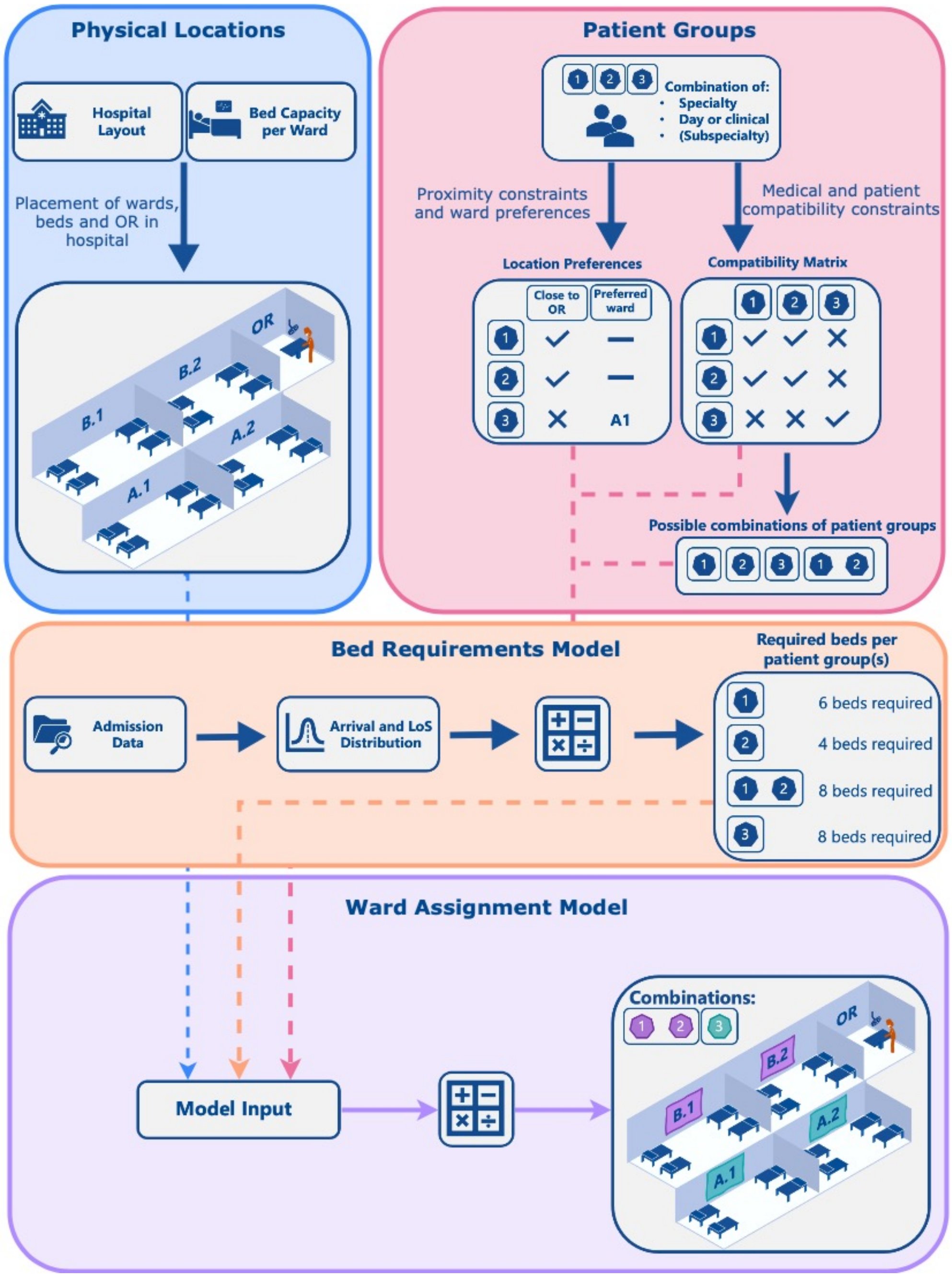


Figure 4: Illustration of assigning specialties for clinical and day treatment departments to wards in a small example hospital

4 Literature Review

This literature review is structured into four sections. Section 4.1 discusses several ways to determine bed requirements for patient groups within both clinical and day treatment departments. This is followed by Section 4.2 which reviews studies on ward pooling. Section 4.3 gives an overview of different models used to assign (combinations of) patient groups to available wards. Lastly, Section 4.4 defines the research gap.

4.1 Determining bed requirements

Determining the optimal number of hospital beds is a complex and challenging task that requires models and techniques capable of accounting for the multi-level, uncertain, and dynamic variables inherent in hospital operations (Ravaghi et al. (2020)). While numerous techniques exist for tactical and operational management, this research focuses on strategic modeling approaches, limiting the scope to the allocation of beds to departments. A shortage of available beds can significantly disrupt hospital operations, serving as a primary cause of admission and surgery cancellations, delays in emergency admissions, premature transfers from intensive care units, delays in inter-unit transfers, and early patient discharges (Green (2002) and Liu (2012)). Conversely, Ravaghi et al. (2020) states that excess bed capacity may result in increased costs and underutilized resources. Striking the right balance between bed availability and demand is crucial for ensuring efficiency, minimizing unnecessary costs, and reducing the blocking probability.

Clinical and day treatment admissions differ significantly in terms of arrival patterns and length of stay, which may lead to differences in the methods used to determine bed requirements. Many studies have explored methods for determining bed requirements in clinical departments; however, there has been limited research on bed requirements for day treatment departments.

A common approach for determining the required number of beds in clinical departments is the application of queuing models. One of the most widely used queuing-based methods for this purpose is the Erlang loss model (Bekker and de Bruin (2010), de Bruin et al. (2010), Green (2006), van Essen et al. (2015), and Veneklaas et al. (2021)). The Erlang loss formula is a mathematical model that calculates the probability of blocking, which represents the likelihood that incoming patients will be denied immediate service when all available beds are occupied (Erlang (1948)).

The Erlang loss model operates under the assumption that patient arrivals follow a Poisson distribution. Studies by Green (2006) and de Bruin et al. (2010) have demonstrated that the number of daily scheduled and unscheduled admissions can often be well-approximated using this distribution. Furthermore, de Bruin et al. (2010) suggests that, for practical applications, strict adherence to a Poisson distribution for the number of daily new admissions is not necessary. Since the variability in the daily number of patient admissions is generally well represented by the Poisson distribution, this assumption is typically valid for practical modeling purposes.

Another assumption of the Erlang loss model is a steady-state system, where patient arrivals and service rates are constant over time. For clinical departments, this assumption is generally true. However, day treatment wards typically experience scheduled admissions with periodic spikes in demand and are often closed during nights and weekends, further disrupting the continuous flow of patients. These factors make the Erlang loss model unsuitable for predicting bed management and patient flow in day treatment wards, as it cannot adequately account for the non-continuous and time-dependent nature of their operations. However, there are approaches that focus on time-dependent loss models. Bekker and de Bruin (2010) examine the effect of a time-dependent arrival pattern on the required number of operational beds and the fraction of refused admissions for clinical wards. They approximate the time-dependent loss system using the Modified-Offered Load (MOL) method. Massey and Whitt (1994) demonstrate that this MOL approximation performs well, particularly when the blocking probability is not too high. This suggests that the MOL approach could be a promising method for day treatment departments.

Another promising approach is the use of simulation modeling. Several researchers have proposed models based on simulation to address the bed management problem, due to its ability to effectively analyze dynamic and complex situations (Berge Holm et al. (2013), Devapriya et al. (2015), and Moengin et al. (2014)). Berge Holm et al. (2013) argues that the complexity that characterizes the healthcare system and patient flow dynamics is better captured with simulation models. Moengin et al. (2014) states that simulation models allow for the evaluation of different scenarios and policy changes, providing more accurate estimates of bed requirements. A study by Hu et al. (2018) found that queuing models tend to oversimplify operations and underestimate congestion levels (especially for smaller systems), and obtain less realistic results than comparable simulation models. However, simulation models tend to be complex and very computationally intensive because of their need to replicate a wide range of variables and interactions within the healthcare system, often requiring significant computational power and time to run multiple scenarios. In contrast, the Erlang loss model, despite its simplifications, provides a more computationally efficient solution, making it a preferred option when the assumptions hold and speed and usability take precedence over the necessity for a detailed system representation.

4.2 Ward pooling

One potential strategy to improve the efficiency of clinical ward layouts is to assign combinations of clinical departments to the same ward, known as ward pooling. Pooling involves combining resources, such as beds, across multiple departments to serve a broader range of patient types. This approach can lead to better utilization of beds by sharing capacity among wards, reducing the total number of required beds while maintaining an acceptable blocking probability (Green and Nguyen (2001) and Koole (2021)). Pooling wards smooths the arrival process of patients across the combined departments, helping to mitigate demand fluctuations that individual wards may experience.

Green and Nguyen (2001) analyze the effects of bed pooling in a large hospital where multiple surgical specialties are combined into a single nursing unit. In their report, they highlight that pooling capacity is particularly effective for smaller departments, which often demonstrate high inefficiency due to underlying stochastic demand. Additionally, they emphasize the importance of considering medical constraints and patient compatibility when pooling bed capacity. Research by Hübner et al. (2018) and van Essen et al. (2015) also underscore the significance of constraints in pooling wards. These constraints vary from one hospital to another. Academic hospitals typically feature larger specialized departments, while smaller regional hospitals often have more combined wards due to their limited patient volume and limitations in infrastructure that prevent each department from having its own dedicated ward. As a result, smaller hospitals might have more pooling of beds across specialties, which helps maximize resource utilization but may introduce challenges related to patient compatibility and specialized care.

Green and Nguyen (2001) also demonstrate that combining clinical departments with varying admission priorities, lengths of stay, and demand can result in an increased number of required beds for a specified blocking probability. In their study, they recommend reserving a certain number of beds for each clinical department. Bekker et al. (2017) suggests that smaller systems particularly benefit from the pooling of clinical departments, while larger systems gain significant advantages from even limited flexibility. They state that the threshold type of control is effective in prioritizing patient types and coping with patients having diverse lengths of stay. In this approach, all beds are fully flexible, but a hierarchy exists in the admission of patient types. The most urgent patients are always admitted when beds are available, while other patient types are only admitted when the number of available beds exceeds a prespecified threshold.

Izady et al. (2024) examined various inpatient bed configurations, ranging from fully dedicated to fully flexible configurations. In a fully dedicated configuration, each specialty is assigned its own ward. This approach offers benefits like highly focused care and low variability in patient conditions, as each ward treats a specific set of patients. However, a key drawback is high slack capacity: patients may wait for admission to their designated ward while beds in other wards remain empty. On the other hand, a fully flexible configuration allows patients from any specialty to be admitted to any available bed, effectively reducing slack capacity through resource pooling. Despite this advantage, the fully flexible setup has its downsides. Mix variability increases as patients with different lengths of stay (LOS) are treated together, potentially leading to a higher average LOS due to less specialized care. Moreover, implementing a fully flexible system requires extensive cross-training of nursing staff, which can be costly.

Both Izady et al. (2024) and Bekker et al. (2017) also explore intermediate bed allocation strategies that fall between fully dedicated and fully flexible configurations. However, such patient assignment strategies are beyond the scope of this research, as they pertain more to tactical-level decision-making. For this study, we focus solely on two configurations: a fully dedicated ward assigned to a single department or a fully flexible joint ward shared by more than one department.

4.3 Models for assigning clinical departments to wards

Both van Essen et al. (2015) and Hübner et al. (2018) formulate models to optimally assign clinical departments to available wards, utilizing a set partitioning approach in their methodologies. A set partitioning approach is a mathematical optimization technique used in operations research, particularly in problems involving resource allocation and scheduling. In this context, it refers to dividing a set of items (clinical and day treatment departments) into distinct, non-overlapping subsets (wards) in such a way that each item belongs to exactly one subset.

The study by van Essen et al. (2015) was one of the first to investigate which clinical departments to cluster such that all departments have enough beds available on the assigned wards. They introduced a methodology to cluster clinical departments and assign wards to these clusters while not exceeding a pre-specified blocking probability. The research builds upon the work of de Bruin et al. (2010) and uses the Erlang loss formula to determine bed requirements. To model the clustering of clinical departments and assignment to available wards, they formulate an Integer Linear Program (ILP). They formulate the problem exact to be able to achieve optimal solutions. To reduce the long computation time, they introduced two heuristic solution approaches. The first heuristic uses the same formulation as the exact model; however, the number of required beds is now approximated by a linear function. The resulting model is again solved by an exact solver. The second heuristic uses a local search approach to determine the assignment of clinical departments to clusters, and the exact model is used to determine the assignment of clusters to wards.

In contrast to van Essen et al. (2015), Hübner et al. (2018) does not use heuristics but introduces a pre-processing step to enhance computational efficiency. They eliminate infeasible and impractical combinations before solving the model. This significantly reduces the solution space while still enabling customization of input parameters. Another key difference is that Hübner et al. (2018) integrates both weekly and monthly seasonality effects into their approach, allowing for more accurate modeling of fluctuating occupancy levels. Their method estimates the bed requirements for each combination of clinical departments by convoluting the probability distribution functions of occupancy levels within each department combination for a specified time period. They then determine the required number of beds based on the day with the highest occupancy demand within that period and the specified blocking probability. This method favors department combinations that help balance overall bed occupancy, as these combinations require fewer beds. However, this approach can be risky, as a single day with unusually high occupancy may distort the results and lead to an overestimation of the required number of beds. They also formulate an ILP to optimize the assignment of clinical departments to wards. Because of the efficient pre-processing step, they handle the problem without having to rely on heuristics.

Several strategies have been proposed to improve the efficiency of clinical ward layouts. However, an equally important consideration is defining the criteria by which one ward layout can be considered superior to another. In existing literature, this is often implicitly captured within a multi-criteria objective function. The study by van Essen et al. (2015) minimizes the maximum number of departments grouped within a single cluster and the walking distances between wards assigned to the same cluster, while also maximizing the proximity of departments to central facilities such as the intensive care unit (ICU) and operating rooms (ORs). Walther (2020) utilizes a utility function to quantify the trade-offs between patient-specific, doctor-specific, and nurse-specific objectives. Furthermore, Hübner et al. (2018) minimizes total costs by selecting the most cost-effective department combinations, while simultaneously reducing the maximum walking distances. This total cost function incorporates both additional pooling costs and costs related to bed allocation for each department combination.

4.4 Research Gap

Existing approaches, such as those by van Essen et al. (2015) and Hübner et al. (2018), have developed models to assign combinations of departments to clinical wards to optimize bed occupancy and minimize patient blocking probabilities. However, current models lack the inclusion of day treatment departments, which have become increasingly essential in hospital operations, and this trend is expected to continue due to the effectiveness and efficiency of the treatment methods. Moreover, there has been little to no research on determining bed requirements for day treatment departments.

This research addresses this gap by developing a model that integrates both clinical and day treatment departments within a single framework. The model will also be tested using real hospital data to ensure its applicability in real-world settings. Furthermore, multiple models will be considered for determining bed requirements, allowing for a comprehensive evaluation of different approaches.

Incorporating day treatment departments alongside clinical departments allows for a more complete solution that better reflects real-world hospital demands, potentially improving both efficiency and patient flow.

5 Data Analysis

In this section, all admissions of the ChipSoft development data recorded in 2019 is analyzed. This data was used to quantify the number of arrivals (Section 5.1), the length of stay (Section 5.2), and the occupancy (Section 5.3). Furthermore, the available wards and beds will also be analyzed (Section 5.4). In the data analysis, we focus on the main specialty and exclude sub-divisions within specialties. Table 1 gives an overview of the specialties within clinical and day treatment departments.

Specialties			
ANE	Anesthesiology	**MDL**	Gastroenterology
CAR	Cardiology	**NCH**	Neurosurgery
CHI	Surgery	**NEU**	Neurology
GER	Geriatrics	**OOG**	Eye surgery
GYN	Gynecology	**ORT**	Orthopaedics
INT	Internal Medicine	**PLA**	Plastic surgery
KAA	Jaw surgery	**REU**	Rheumatology
KNO	Ear, Nose and Throat	**TAN**	Dentistry
LON	Lung surgery	**URO**	Urology

Table 1: Specialties in clinical and day treatment departments

5.1 Arrivals

This section analyzes the number of arrivals in ChipSoft's development data from 2019. We begin by focusing on admissions to specialties within clinical departments and then shift to day treatment departments. Given that weekends typically show lower arrival numbers, our analysis is limited to weekday arrivals.

Clinical departments

To better understand the arrival patterns, Figure 5 below presents the daily number of arrivals in clinical departments by specialty throughout the year. This analysis aims to highlight potential trends, peak periods, and variations in patient admissions across specialties. The figure shows significant variability throughout the year, with some days experiencing higher arrivals than others. However, despite this variability, no clear or consistent patterns emerge throughout the year.

Figure 5: Daily number of arrivals in clinical departments by specialty for the year 2019

For clinical departments, the number of daily arrivals is often assumed to be Poisson. To analyze whether this assumption holds, Figure 6 presents a histogram illustrating the daily arrival counts for the Internal Medicine specialty, with the horizontal axis representing the number of arrivals per day and the y-axis showing the frequency in days. The orange line indicates the Poisson fit, with a rate of $\lambda = 8.26$. This rate is calculated as the mean of daily arrivals for the Internal Medicine specialty, excluding weekends. As observed, the Poisson distribution aligns well with the actual data. However, this observation alone does not establish statistical significance. The histograms for all specialties within clinical departments are provided in Appendix A.

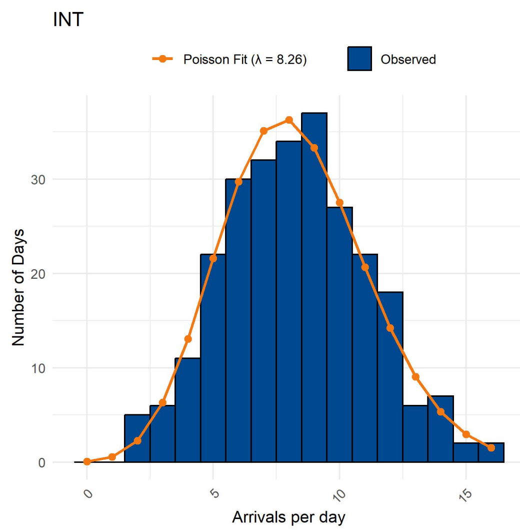


Figure 6: Distribution of admissions for Internal Medicine

To further analyze whether the Poisson assumption holds, we perform a Chi-squared goodness-of-fit test. The Chi-squared goodness-of-fit test is a statistical method used to evaluate whether the observed distribution of data matches a theoretical distribution. For each specialty, we calculate the observed counts of daily arrivals and compare them with the theoretical counts predicted by a Poisson distribution, where the arrival rate (λ) is estimated from the observed data. The test evaluates the similarity between observed and expected values, providing a p-value that indicates whether any observed differences are statistically significant. In this analysis, we take a significance level of 0.05 meaning that a p-value above 0.05 suggests that the data do not significantly deviate from the Poisson distribution. If this is the case, then arrivals could reasonably follow this distribution. Conversely, a p-value below 0.05 suggests that the data deviates from the Poisson assumption. Table 2 shows the results of this test, indicating whether each specialty's arrival pattern aligns with a Poisson distribution. We see that the p-value is larger than 0.05 for 9 specialties, highlighted in bold. This means that, for 9 out of 15 specialties, there is insufficient evidence to reject the Poisson assumption based on a significance level of 0.05. Furthermore, according to de Bruin et al. (2010), strict adherence to the Poisson distribution is often unnecessary in practical applications.

Statistical Analysis of Clinical Arrivals

Specialty	p-value	Mean λ	Variance σ^2	Var/Mean σ^2/λ
ANE	0.01	0.169	0.333	1.975
CAR	0.024	12.115	8.710	0.719
CHI	**0.724**	6.939	6.804	0.981
GER	**0.834**	1.510	1.482	0.981
GYN	**0.283**	2.008	2.369	1.180
INT	**0.932**	8.264	8.041	0.973
KAA	**0.064**	0.636	0.779	1.224
KNO	0.017	1.115	1.364	1.223
LON	**0.377**	6.395	6.694	1.047
MDL	**0.170**	4.781	4.479	0.982
NCH	**0.998**	2.513	2.589	1.030
NEU	**0.400**	6.249	6.134	0.982
ORT	0.001	6.636	7.148	1.078
PLA	0.002	1.701	1.318	0.775
URO	0.001	4.241	7.368	1.737

Table 2: Chi-squared goodness of fit p-values, mean, variance, and variance-to-mean ratio of daily arrivals per specialty for clinical departments

Table 2 also presents the mean (λ), variance (σ^2), and variance-to-mean ratio (σ^2/λ) for each specialty. For a Poisson distribution, the mean and variance are equal, so the variance-to-mean ratio should be close to 1. A variance-to-mean ratio significantly different from 1 indicates a deviation from a Poisson distribution, suggesting either overdispersion (ratio $>$ 1) or underdispersion (ratio $<$ 1). The variance-to-mean ratio across specialties in Table 2 ranges from 0.719 to 1.975, reflecting considerable variability in daily arrivals. For specialties with a p-value greater than 0.05, the variance-to-mean ratio is generally close to 1. For the other specialties, we observe overdispersion for ANE, GYN, KNO, and URO, where the variance exceeds the mean. However, ANE has a low number of arrivals, making it uncertain how accurate the results for this specialty are. Underdispersion is evident for CAR and PLA, where the variance is lower than expected. For ORT, the mean and variance are nearly equal, resulting in a variance-to-mean ratio close to 1. Variability in patient admissions is generally well represented by a Poisson model, even if the variance-to-mean ratio is not exactly 1. Given the histograms of the daily number of arrivals, the variability expressed by the variance-to-mean ratio, and the Chi-squared goodness-of-fit test results—where 9 out of 15 specialties show p-values above 0.05—we can conclude that the Poisson distribution provides a sufficiently adequate approximation to justify its use for modeling the arrival data for specialties within clinical department.

Day treatment departments

The arrival patterns for day treatment departments differ significantly from those of clinical departments, particularly in terms of patient distribution throughout the week. Figure 7 illustrates the mean number of arrivals per hour and weekday, averaged across all clinical departments (top, orange line) and all day treatment departments (bottom, blue line).

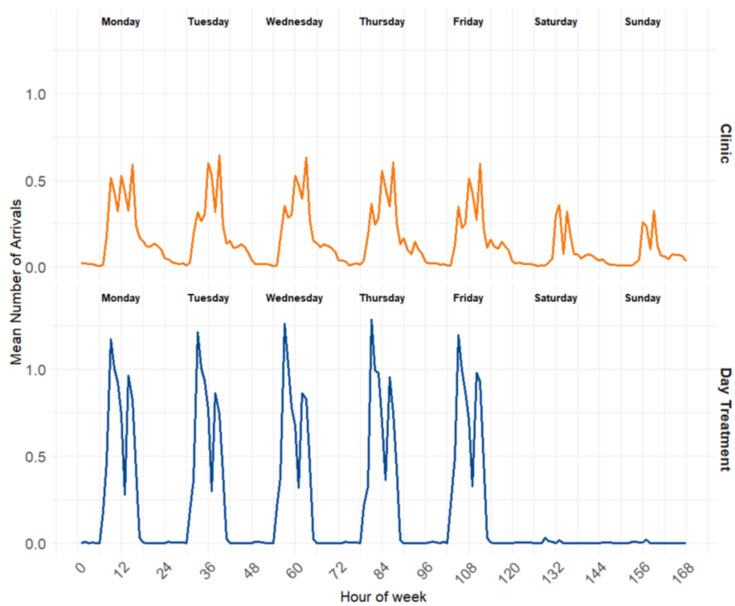


Figure 7: Average number of arrivals per hour and weekday, averaged across all clinical departments (top) and all day treatment departments (bottom)

In clinical departments, patient arrivals are distributed relatively evenly across weekdays. Weekend arrivals show a slight decrease compared to weekdays. The overall pattern remains consistent throughout the week, with no extreme peaks observed, only a lower number of arrivals during nighttime hours. In contrast, day treatment departments exhibit more variable arrival patterns, with strong and consistent peaks observed between 7 AM and 1 PM on weekdays. These peaks suggest a high volume of patients arriving for time-sensitive, same-day procedures. Patient arrivals drop dramatically in the late afternoon, with nearly no arrivals during the night. Additionally, the data shows a significant decrease in arrivals over the weekend, indicating that day treatment departments primarily operate during standard working hours. The sharp peaks and absence of arrivals outside working hours show that these departments are designed for time-sensitive, same-day procedures, whereas the clinical patient care shows a more continuous and less time-sensitive nature of clinical patient care.

In Section 4.1, we discussed the assumption of stationary arrivals in Erlang models, which implies that arrival patterns remain stable over time, allowing for a balanced distribution of resources. In clinical departments, this assumption holds more reasonably. While arrivals are not perfectly evenly distributed throughout the week, their variability remains relatively low, making the overall arrival process more predictable. In contrast, for day treatment departments, arrivals exhibit significant time-dependent fluctuations, with peak demand during weekdays and very little activity during nights and weekends. The strong time dependency in patient arrivals makes it difficult to assume stationarity in these departments, as the system experiences substantial variations in demand over short time periods.

Admission on day treatment departments are often planned admissions, therefore the patterns across the weekdays is roughly the same. An example is shown in Figure 8. This illustrates the average number of arrivals per hour and weekday for the Internal Medicine specialty. The peaks are at approximately the same hours on each day and there are no arrivals outside office hours. Comparable plots for all specialties are provided in Appendix B.

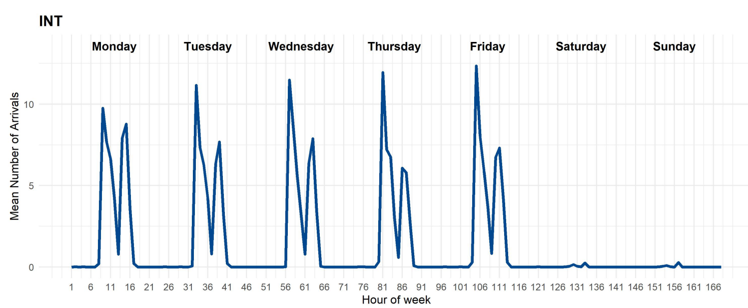


Figure 8: Average number of day treatment arrivals for Cardiology per hour of the week

To better understand the arrival patterns, Figure 9 below presents the daily number of arrivals in day treatment departments by specialty throughout the year. Note that we again exclude weekends for further analysis of day treatment arrivals.

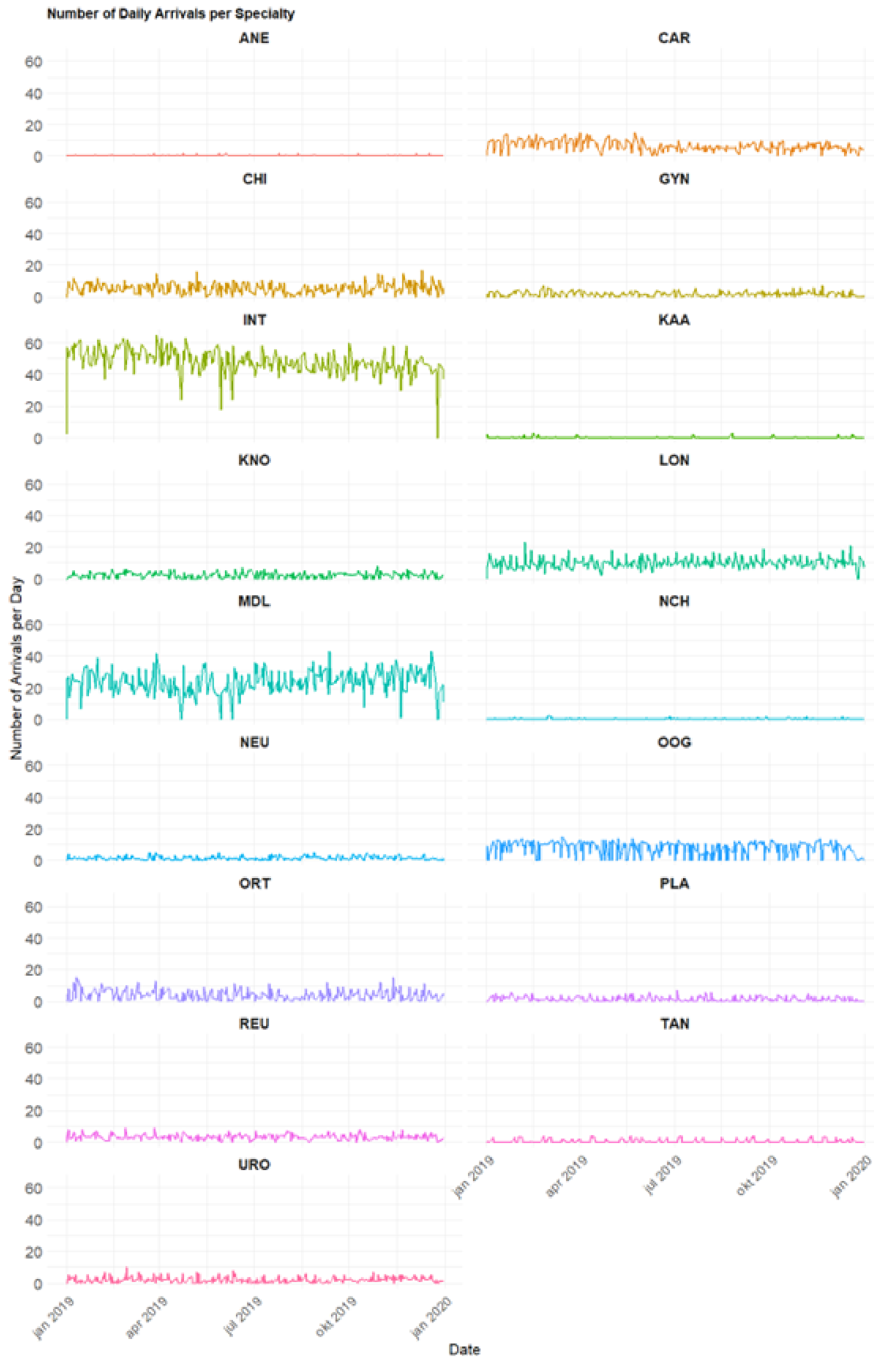


Figure 9: Daily number of arrivals in day treatment departments by specialty for the year 2019

Figure 9 illustrates variability in patient arrivals throughout the year but does not reveal any distinct seasonal patterns. However, for both Internal Medicine (INT) and Gastroenterology (MDL)—the two largest specialties within day treatment— certain days exhibit sudden drops in patient arrivals, occurring consistently on the same dates: 01-01-2019, 22-04-2019, 10-06-2019, 30-05-2019, 25-12-2019, and 26-12-2019. These dates correspond to public holidays in the Netherlands, including New Year's Day, Easter Monday, Whit Monday, Ascension Day, and Christmas (both days). As most admissions to day treatment departments are planned in advance, it is reasonable that fewer admissions are scheduled on these holidays, likely due to reduced staff availability and lower demand during festive periods.

Finally, we analyse the mean, variance, and variance-to-mean ratio of daily arrivals on day treatment departments split on specialty in Table 3. The variance-to-mean ratio for day treatment departments ranges between 0.871 and 2.723. Only 2 ratios are < 1 (NEU, REU), the other 15 are > 1 and thus experience over-dispersion. Compared to clinical departments, day treatment departments exhibit significantly higher variability in patient arrivals.

Statistical Analysis of Clinical Arrivals			
Specialty	**Mean** λ	**Variance** σ^2	**Var/Mean** σ^2/λ
ANE	0.157	0.218	1.385
CAR	6.287	10.167	1.617
CHI	5.590	12.820	2.293
GYN	1.782	2.402	1.348
INT	47.670	67.999	1.426
KAA	0.149	0.235	1.575
KNO	2.276	3.370	1.481
LON	10.126	12.619	1.246
MDL	23.682	62.241	2.628
NCH	0.165	0.184	1.119
NEU	1.602	1.394	0.871
OOG	7.582	20.644	2.723
ORT	4.272	11.114	2.602
PLA	1.567	2.470	1.576
REU	3.333	3.300	0.990
TAN	0.475	1.104	2.324
URO	2.180	3.548	1.628

Table 3: Mean, variance, and variance-to-mean ratio of the number of daily arrivals per specialty for day treatment departments

5.2 Length of Stay

Understanding the length of stay for each specialty is critical, as it directly impacts bed availability and turnover. The box plots in Figure 10 show the distribution of length of stay in days across various specialties within clinical departments, providing insights into the differences in hospital stay durations. The box plots highlight numerous outliers, represented by dots, which correspond to patients with significantly longer stays. These cases likely result from complex medical conditions or other exceptional circumstances. While the inclusion of outliers provides a complete picture of the data distribution, it also compresses the interquartile range and thereby reduces the visual clarity of the typical length of stay for most patients.

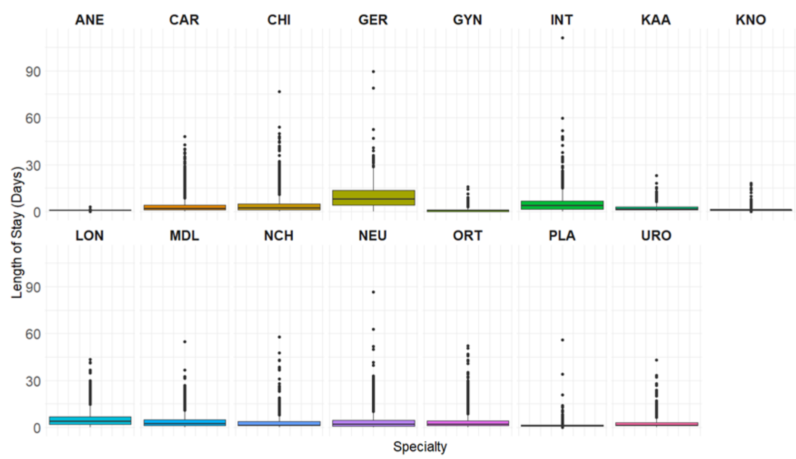


Figure 10: Boxplot of length of stay by specialty for clinical departments

To address this, Figure 11 presents a refined box plot showing only lengths of stay up to 30 days to emphasize the distribution of most observations. A larger box indicates greater variability in the length of stay within a specialty. Specialties such as GER and INT exhibit substantial variation, while specialties like ANE, KNO, and PLA have comparatively less variability, implying a more consistent length of stay among patients in these areas. Additionally, the distributions for several specialties, especially KAA, NCH, NEU, and URO, are right-skewed. This is evident from the box plots, where the median lies close to the lower quartile. This right-skewness indicates that while most patients in these specialties have relatively shorter stays, a smaller subset requires extended hospitalizations. Such extended stays could be associated with more complex treatments or complications. The average length of stay across all specialties is 4 days, whereas the median length of stay is 2 days.

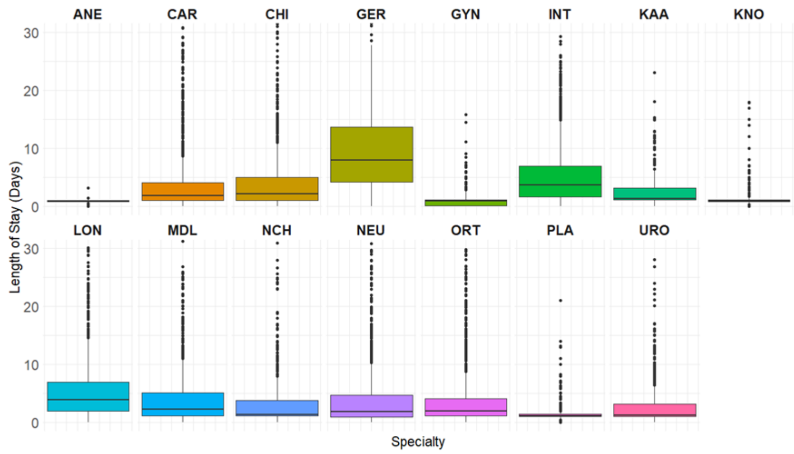


Figure 11: Boxplot of length of stay by specialty for clinical departments truncated at 30 days

Figure 12 presents box plots for specialties within day treatment departments, with the length of stay now expressed in hours. Compared to clinical departments, these box plots show fewer outliers, which likely reflects the more standardized procedures and lower complexity of cases typically handled in day treatment departments. The average length of stay across all specialties is about 3.5 hours, suggesting that efficient combination of day treatment specialties could allow for double occupancy of beds, reducing the total number of beds needed and improving resource utilization.

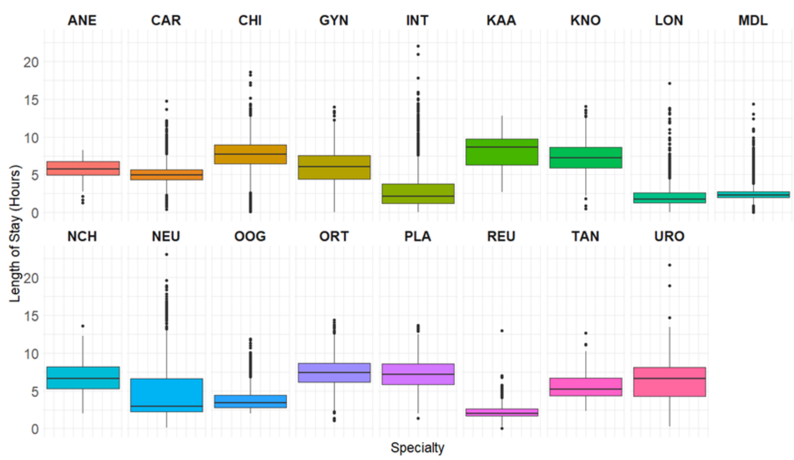


Figure 12: Boxplot of length of stay by specialty for day treatment departments

5.3 Occupancy

After analyzing both arrivals and length of stay, we now shift our focus to bed occupancy, as it offers a more detailed insight into bed utilization. Figure 13 illustrates the bed occupancy per specialty for both clinical and day treatment departments in 2019. It shows for each day the maximum number of occupied beds on that day. While most specialties show admissions to both day treatment and clinical departments, there are some exceptions. Specialties OOG, REU, and TAN only have admissions to day treatment departments, whereas GER admissions are exclusively to clinical departments. The plots also reveal potential seasonal trends; for example, the specialty LON exhibits higher occupancy during the colder months. Additionally, across all specialties, the number of beds occupied for day treatment is consistently lower than that for clinical departments. All these findings provide valuable insights for estimating the required number of beds. Figures 14 and 15 illustrate the average occupancy for clinical and day treatment departments, respectively, with error bars indicating the standard deviation in occupancy. These statistics are based exclusively on weekdays.

These plots offer a first direction towards the required number of beds per specialty by showing the average occupancy. Additionally, the error bars allow us to assess the variation in bed usage, highlighting which specialties experience more stable occupancy and which ones are highly variable. For both clinical and day treatment departments, specialties INT exhibits relatively large error bars, suggesting higher variability in bed usage. In clinical departments, a general trend can be observed: higher average occupancy often corresponds with greater variation. In contrast, for day treatment departments, substantial deviations are also noticeable in smaller specialties like OOG and ORT. Overall, the variation in bed occupancy tends to be relatively higher in day treatment departments compared to clinical departments.

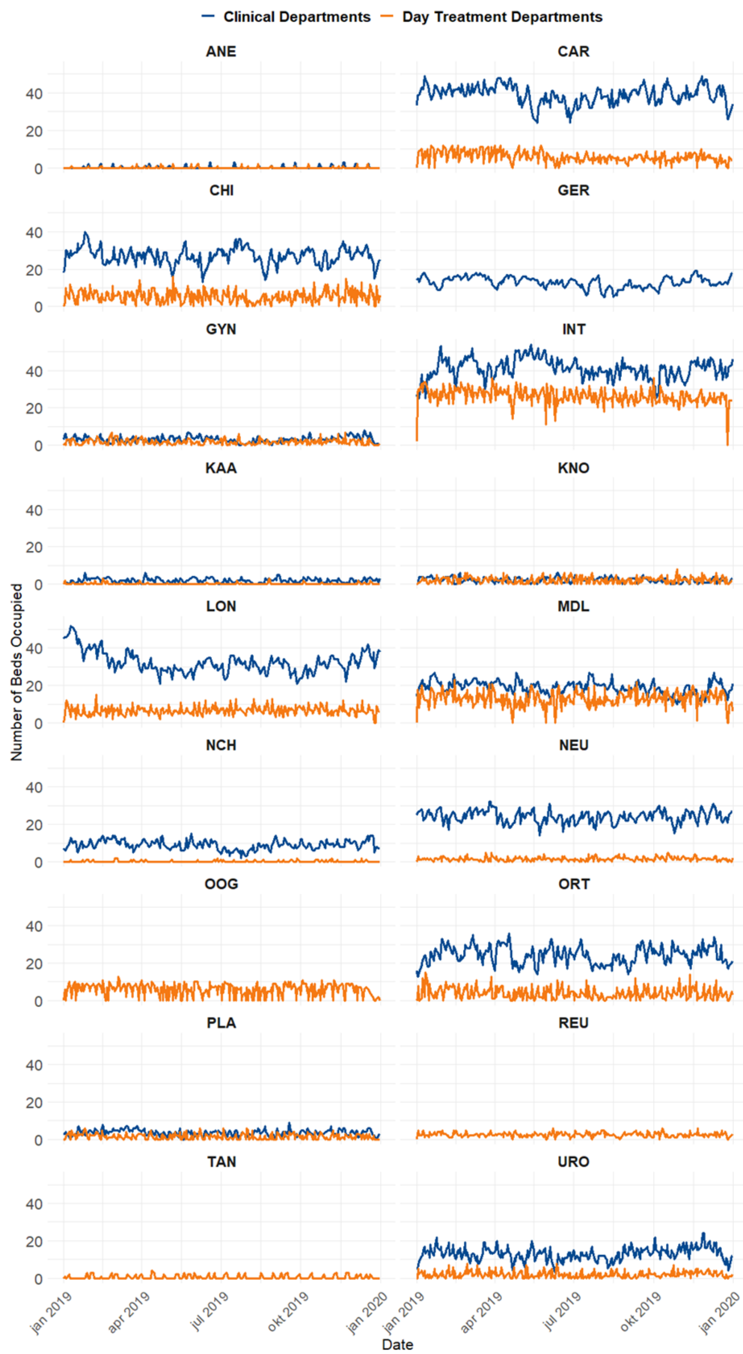


Figure 13: Number of occupied beds per specialty for clinical and day treatment departments throughout 2019

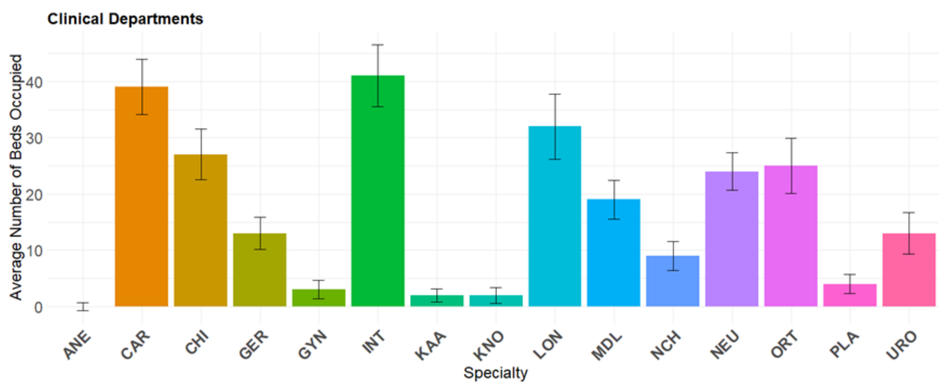


Figure 14: Average number of occupied beds per specialty for clinical departments

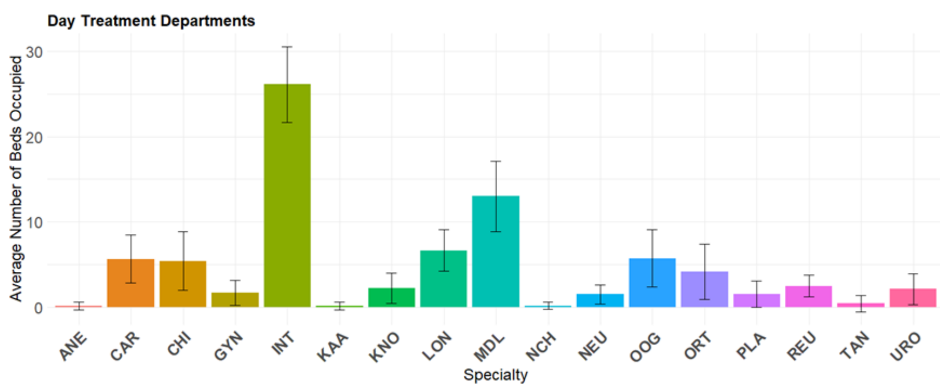


Figure 15: Average number of occupied beds per specialty for day treatment departments

5.4 Available beds and wards

When assigning specialties to wards, we must consider the availability of beds and ward space, as we cannot exceed the capacity provided. To address this, we include all beds registered in HiX that are assigned to one of the wards. The ChipSoft development dataset contains 34 wards, each with its own rules regarding patient admission. One example of a ward is the Brain Care Unit, which admits only patients aged 18 and above, and the allowed specialties are neurology (NEU) and neurosurgery (NCH). This ward is designated for clinical admissions only, with no day treatments, and has 8 beds available. In this study, we set aside these specific rules and the current specialty assignments. Instead, we focus solely on the available spaces and bed capacity in each ward. For instance, for the Brain Care Unit, we consider it simply as a ward with 8 beds available for modeling purposes. The model will then determine which specialties will be assigned to that ward and which patients are allowed to be admitted.

Most of the available beds in the hospital are standard beds. However, certain patients require specialized bed types tailored to their specific medical needs. For example, chairs are available for short-term treatments or procedures like chemotherapy that do not require a full bed. For simplicity reasons, we do not distinguish between different types of beds for now.

In Section 2.3 we discussed several wards that are excluded from this study. We treat them as wards that are in a fixed position with a fixed number of beds. Therefore, we do not want to count these wards and beds as available capacity. After removing these wards from the total of 34 available wards, we have 13 clinical wards and 7 day treatment wards with 493 beds spread across them. Figure 16 shows how these 493 beds are divided among the available wards.

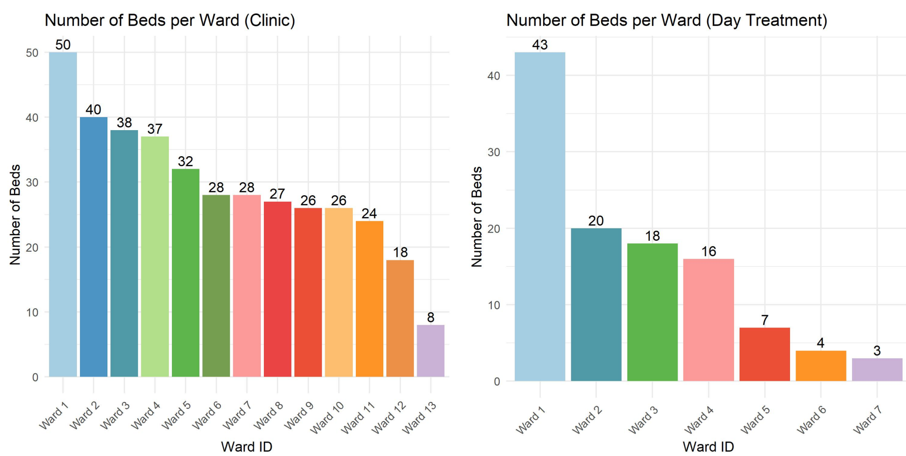


Figure 16: Number of available beds per ward for clinical and day treatment department

6 Methodology

This chapter introduces the model and solution approach for assigning specialties to wards in both clinical and day treatment departments while ensuring a predefined blocking probability and complying with capacity, medical, and location constraints. Figure 17 provides an overview of the methodology. In the figure, red databases represent data retrieved from HiX, while blue databases indicate user-provided input. The orange boxes illustrate the various steps of model preparation. The green box represents the bed requirements model, and the purple box the cluster-to-ward assignment model.

Section 6.1 describes the processing of user input, while Section 6.2 focuses on the preprocessing of data retrieved from the hospital's database. Section 6.3 explains the calculation of bed requirements, and finally, Section 6.4 formulates the assignment of clusters to wards as an Integer Linear Program (ILP).

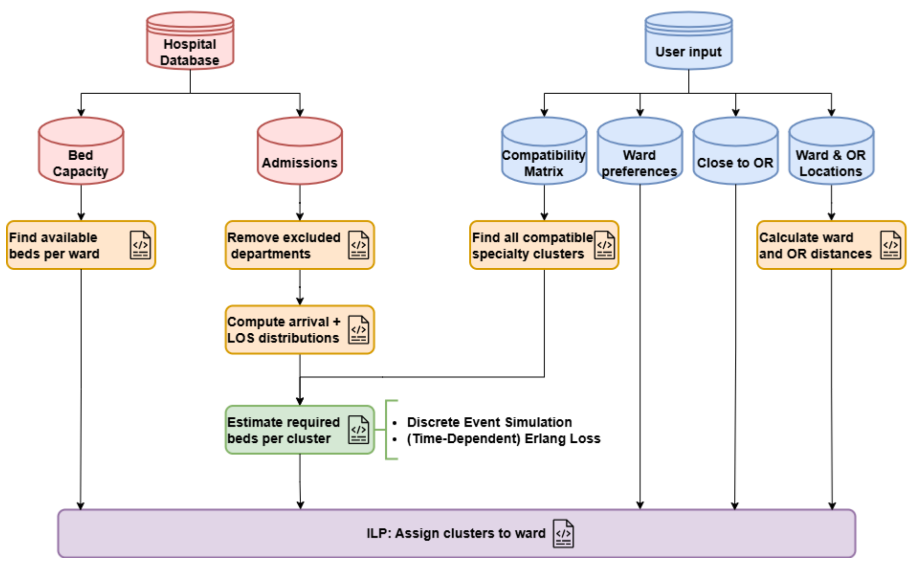


Figure 17: Summary of the Methodology

6.1 Processing User Input

To find an optimal solution for the problem at hand, user input is required, as each hospital has specific preferences for an efficient clinical ward layout. Therefore, we request the following four inputs from the hospital:

1. **Compatibility Matrix**: A matrix indicating which specialties can or cannot be assigned to the same ward.

2. **Ward Preferences**: Users can specify a preferred ward for each specialty, such as maintaining its current assignment. If left unspecified, the specialty may be placed in any available ward.

3. **Close to OR**: Users can indicate whether a specialty should be located near the operating room.

4. **Ward and OR Locations**: Users must provide the floor and wing locations for all wards and the OR. This information enables the calculation of distances between wards and between wards and the OR.

The ward preferences and proximity to the OR can be directly used as input for the cluster-to-ward assignment model without additional preprocessing. However, the compatibility matrix and the ward and OR locations require further preprocessing before they can be utilized in the model. We will elaborate on both aspects, first addressing the preprocessing needed for the compatibility matrix.

An example compatibility matrix is shown in Table 4. The diagonal of the matrix is always 1, as it reflects the compatibility of a specialty with itself. Since the matrix is symmetric, the values below the diagonal are left empty. In the values above the diagonal, a 0 indicates that the two specialties cannot be placed together, while a 1 signifies that there are no restrictions, meaning these specialties can be placed together at the same ward. We expect the user to complete a compatibility matrix because the compatibility of specialties may vary between hospitals. This variation arises from differences in the size of specialties and the level of specialized care available at each hospital. Where one hospital may be highly specialized in neurology and require a dedicated ward, other hospitals might want to combine neurology with neurosurgery, for example.

If all n specialties in a hospital were mutually compatible, there would be $2^n - 1$ possible combinations. For example, in the ChipSoft development data, there are 15 specialties across clinical departments, leading to $2^{15} - 1 = 32{,}767$ potential combinations. However, in practice, only a limited number of specialties can typically be grouped together due to medical and patient care constraints, as discussed in Section 3. We can reduce the 32,767 potential combinations by eliminating those where two specialties are not permitted to be placed in the same ward, as indicated by the compatibility matrix. For each pair of specialties with a "*0*" in the matrix, any combination that includes both specialties is removed. This filtering process leads to a reduced set of valid combinations, referred to as clusters, which represent feasible specialty groupings. Importantly, for each valid cluster, the order of the specialties does not matter, so combinations such as {CHI, GYN} are considered the same as {GYN, CHI}. Using the example compatibility matrix in Table 4, the number of possible combinations is reduced to 595 clusters. These clusters will serve as input for the bed requirements model, discussed in Section 6.3.

	ANE	CAR	CHI	GER	GYN	INT	KAA	KNO	LON	MDL	NCH	NEU	ORT	PLA	URO
ANE	**1**	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CAR		**1**	0	0	0	0	0	0	0	0	0	0	0	0	0
CHI			**1**	0	1	0	1	1	0	1	1	0	1	1	1
GER				**1**	0	0	0	0	0	0	0	0	0	0	0
GYN					**1**	0	1	1	0	1	1	0	1	1	1
INT						**1**	0	0	1	1	0	0	0	0	0
KAA							**1**	1	0	1	1	0	1	1	1
KNO								**1**	1	1	0	1	1	1	1
LON									**1**	0	0	0	0	0	0
MDL										**1**	1	0	1	1	0
NCH											**1**	1	1	1	1
NEU												**1**	0	0	0
ORT													**1**	1	1
PLA														**1**	1
URO															**1**

Table 4: Example compatibility matrix for all clinical specialties

Furthermore, we use separate compatibility matrices for clinical departments and day treatment departments. Since the care provided in day treatment departments is generally less complex, it is more manageable for nursing staff to care for multiple specialties. In clinical departments, the complexity of care often requires specialized knowledge and skills that limit the ability of nursing staff to care for multiple specialties. As a result, clinical departments tend to have stricter compatibility rules to ensure that nurses can effectively manage the specific needs of each specialty.

We also preprocess the physical locations of the wards and the OR. The user is expected to specify the ward along with an additional characteristic that clarifies its exact location on a specific floor, such as a wing. We aim to calculate the distance between all locations in order to set constraints in the cluster-to-ward assignment model on maximum allowed distances. If a cluster is assigned to more than one ward, we want to assign the cluster preferably to two adjacent wards and not to wards that are on the other side of the hospital. We expect the user to fill in for each ward:

- Floor the ward is located at: 0, 1, 2, ...

- Wing the ward is located at: A, B, C, ...

This approach assumes that the user can appropriately categorize wards according to the given input structure. This was developed in consultation with a collaborating hospital, where wards are divided based on this system. However, if another hospital were to implement this, adjustments may be necessary to account for differences in architectural layout and ward organization.

The distance between two wards is computed based on their respective floor and wing locations. The calculation follows these steps:

1. **Floor Difference**: The absolute difference between the floors of two wards is determined. Since movement between floors is generally more time consuming than movement within the same floor, this difference is weighted more heavily. Specifically, the floor difference is multiplied by a factor of 2 to reflect the additional effort required for vertical movement.

2. **Wing Difference**: The absolute difference in wing locations is calculated based on their position in the alphabet. Adjacent wings (e.g., A and B) have a difference of 1, while wings further apart (e.g., A and C) have a difference of 2, and so on.

The total distance between two wards is then given by the formula:

$$\text{distance} = 2 \times |F_1 - F_2| + |W_1 - W_2| \tag{1}$$

where F_1 and F_2 represent the floor numbers of the two wards, and W_1 and W_2 denote their respective wing positions, converted into numerical values based on their alphabetical order. Table 5 shows example user input of ward locations and Table 6 shows the corresponding distance matrix. The same idea is used for calculating distances between wards and the OR.

Ward	Floor	Wing
A1	1	A
B1	1	B
D1	1	D
A2	2	A
B2	2	B

Table 5: Example user input

Ward	A1	B1	D1	A2	B2
A1	0	1	3	2	3
B1		0	2	4	2
D1			0	5	6
A2				0	1
B2					0

Table 6: Example ward distance matrix

6.2 Processing Data from HiX

In this section we discuss how we process the data from HiX. The data is extracted from HiX using an SQL query and then loaded into RStudio for model implementation, following standard practice within ChipSoft's R&D team for capacity management. Two key datasets are retrieved from HiX: bed capacity and admissions.

The bed capacity dataset provides a comprehensive overview of all wards and their associated beds, ensuring a clear understanding of the available resources. This dataset corresponds to the available bed capacity presented in Figure 16 in Section 5.4. Wards that are excluded from the analysis are removed from the available capacity. The same applies to the admissions dataset, where all admissions from excluded wards are filtered out at the start of the preprocessing. The refined admission data serves as input for modeling arrival patterns and length of stay distributions, which are essential for determining bed requirements. These aspects will be discussed in greater detail in the following section.

6.3 Bed Requirements Model

The preceding sections outlined the process of retrieving admission data to analyze arrival patterns and length of stay distributions, as well as the use of the compatibility matrix to identify all feasible specialty clusters. Both components serve as input for the bed requirements model, which estimates the necessary number of beds for each potential cluster based on a predefined blocking probability. In this section, we present two methodological approaches: **Discrete Event Simulation** and the **Erlang loss model**. Additionally, we differentiate between the standard Erlang loss model and its time-dependent variant, which is specifically applied to day treatment departments.

6.3.1 Erlang loss Model

In this section, we focus on determining bed requirements using the Erlang loss model. This model is used for clinical departments only, the time-dependent variant of the Erlang loss model will be discussed in Section 6.3.2. The required number of beds in a ward depends on both the arrival rate and the length of stay of the departments assigned to that ward. We follow the Erlang loss (or $M/G/c/c$) model formulation as stated by de Bruin et al. (2010). In the $M/G/c/c$ model, patients arrive according to a Poisson process with parameter λ, which is denoted by the first M in the model notation. The M stands for *Markovian*, indicating that the inter-arrival times follow an exponential distribution with a memoryless property. This means that the time until the next patient's arrival is independent of previous arrivals, making the process well-suited to represent random arrivals typically seen in hospital settings. In this model formulation, the second letter, G, stands for *General*. This indicates that the service time, so the length of stay, follows a general probability distribution. The third letter c represents the number of servers, which in this context are the operational beds available in a ward. The fourth letter c indicates that if a patient arrives and finds all c beds occupied, the patient is blocked and does not enter the system.

In Section 5.1, we assessed whether the Poisson distribution is a reasonable assumption for daily arrivals. Building on this, we now assume that arrivals follow a homogeneous Poisson process and estimate the arrival rate based on the average number of arrivals per day. We calculate the arrival rate λ by taking the average number of arrivals during weekdays, excluding weekends. Weekends typically experience lower arrival rates, and including them would lead to an underestimation of λ. This could result in fewer beds being allocated, which would not adequately account for peak demand during weekdays. By excluding weekends, we ensure that the computed arrival rate reflects the busier part of the week, providing a more accurate estimate of the required bed capacity.

The assumption that the length of stay (LOS) of patients is independent and identically distributed, following a general distribution may not always hold in real-life scenarios. For instance, when a ward is at full capacity, hospital staff might prioritize discharging patients who are closest to recovery. Nonetheless, for simplicity we assume that LOS is independent and identically distributed. In the context of the Erlang loss model, the average length of stay (ALOS) is represented by the parameter μ, which is calculated by taking the mean of the length of stay for all patients within each cluster.

The blocking probability P_c for a ward with c available beds, according to the Erlang loss formula, is given by:

$$P_c = \frac{\frac{(\lambda\mu)^c}{c!}}{\sum_{k=0}^{c} \frac{(\lambda\mu)^k}{k!}} \tag{2}$$

The average number of occupied beds can then be computed by:

$$\text{Average number of occupied beds} = (1 - P_c)\lambda\mu \tag{3}$$

We can then formulate the occupancy rate as follows:

$$\text{Occupancy rate} = \frac{(1 - P_c)\lambda\mu}{c} \tag{4}$$

So, using Equation 2, we can compute the blocking probability for a number of beds c. The goal is to find the smallest number of beds c for which the blocking probability P_c is smaller or equal to a specified blocking probability $P_{\max}$. This means we can iteratively increase c until $P_c \leq P_{\max}$. This is done for each cluster of specialties.

6.3.2 Time-Dependent Erlang loss

This section will discuss the time-dependent Erlang loss model. We use this model exclusively for day treatment departments. Previous research by Bekker and de Bruin (2010) has demonstrated that incorporating daily arrival patterns for clinical departments has only a limited effect on bed requirements.

A time-dependent model for clinical departments is typically more relevant for operational or tactical levels within clinical departments, but for the strategic level of this research, it is deemed too detailed. So, we apply the time-dependent Erlang loss model only to day treatment departments, where both the weekend-weekday effect and daily patterns have a significant influence.

We model the number of patients present in day treatment departments using an $M_t/G/c/c$ queue. The key difference from the Erlang loss model in Section 6.3.1 is the M_t, which represents arrivals following a non-homogeneous Poisson process with time-dependent arrival rate $\lambda(t)$. However, obtaining exact results for queueing models with time-dependent arrival patterns is generally challenging. Therefore, we introduce an approximation for the $M_t/G/c/c$ model.

Exact results are available for the infinite-server queue ($M_t/G/\infty$), which allows us to directly determine the offered load when bed capacity is not constrained. We define $\lambda(t)$ as the arrival rate at time t, and we define S as the continuous random variable denoting the length of stay of a patient. Then, by leveraging the $M_t/G/\infty$ model, we can estimate the expected bed occupancy at time τ as follows:

$$m(\tau) = \int_0^\tau \lambda(t) P(S > \tau - t) dt \tag{5}$$

The integral is based on a continuous time approximation. In practice, we often work with discrete time points (e.g., per hour) instead of continuous time. Therefore, we transform the integral into a sum:

$$m(\tau) \approx \sum_{t=0}^{\tau} \lambda(t) P(S > \tau - t) \tag{6}$$

In Equation The expected number of arrivals strongly depends on the hour and weekday. We therefore use historical data to compute the average number of arrivals per hour and weekday, which gives us $\lambda(t)$ for $t \in \{1, 168\}$. During the data analysis in Section 5.1, we observed that for some specialties, there were very few or no patient arrivals on holidays. To ensure a more accurate estimation, we exclude these holidays when calculating the mean number of arrivals. For the length of stay, we use the empirical distribution, such that:

$$P(S > x) = \frac{\text{number of patients with length of stay } > x}{\text{total number of patients}}. \tag{7}$$

Now, based on this infinite-server queue we want to approximate the time-dependent loss system $M_t/G/c/c$. Massey and Whitt (1994) showed that the Modified-Offered Load (MOL) approximation performs well, making it a suitable approach for our model. In this method, the stationary loss model remains unchanged, but the offered load is replaced by the *modified* offered load $m(\tau)$, obtained from the infinite-server queue with a time-dependent arrival process. Using this approximation, the blocking probability at time τ can be calculated as follows:

$$P_{\rm c}(\tau) = \frac{\frac{(m(\tau))^c}{c!}}{\sum_{k=0}^{c} \frac{(m(\tau))^k}{k!}}, \tag{8}$$

Now we can calculate for each $\tau \in \{1, 168\}$ the blocking probability for any number of beds c. We want to have enough beds on the busiest hour of the week, so we iteratively increase the number of beds until the blocking probability of the busiest hour $\leq P_{\max}$. This enables us to compute bed requirements for day treatment departments using the Erlang loss model, while still accounting for the nonstationary arrivals.

6.3.3 Discrete Event Simulation

As an alternative to the Erlang loss model, we also use simulation for determining bed requirements. Previous research by Berge Holm et al. (2013) demonstrated that simulation models effectively capture the complexity of patient flow through hospital wards. They utilized a Discrete Event Simulation (DES) approach, which models a system as a sequence of discrete events occurring at specific points in time. Each event represents a change in the system's state, and between events, no changes are assumed to occur, allowing the simulation time to advance directly to the next event's occurrence.

The DES model in this research is built on several key assumptions:

- **Patient arrivals follow a Poisson process**: The arrival rate is defined for each hour, weekday, and week to capture the time-dependent arrival patterns typical of day treatment departments. Under the assumption of Poisson-distributed arrivals, we assume that the time between two patient arrivals is exponentially distributed with rate $\lambda(t)$ in hour t. However, since the rate $\lambda(t)$ varies over time, we need to take into account the time-varying nature of the process when simulating these times. This is explained later on in this section.

- **Length of stay follows an empirical distribution**: We base the length of stay on historical data from various departments, considering variations and outliers. By using this empirical distribution, we can more accurately represent the real dynamics of patient care within the hospital. This choice is also influenced by practical considerations, such as the difficulty of fitting alternative distributions.

- **Number of beds are assumed to be unlimited**: In the simulation, we assume an unlimited bed capacity, meaning no patient is blocked due to a lack of available beds. Once the simulation is complete, we can analyze simulated patient arrivals and departures for varying bed capacities (c) to assess the blocking probability and occupancy rate.

- **The simulation period spans one year**: We compute the results of the simulation over one year of simulated admissions. We include a 6-week warm-up period for clinical departments to ensure that the wards begin with a realistic patient load rather than starting empty.

- **Departments are simulated independently of each other**: We simulate arrivals and departures of patients per department. There are no interactions between departments, thereby simplifying the model by eliminating the need to account for patient transfers.

Figure 18 illustrates the various steps in the DES model, which follows a cycle consisting of six steps repeated at each iteration. Before entering this cycle, we first need to jump to an hour when patient arrivals are expected. Although the simulation begins at 00:00:00 on 01/01, day treatment departments for example are empty during the night, resulting in a zero arrival rate at that time. Therefore, the simulation jumps forward to the first hour where the arrival rate, $\lambda(t)$, is greater than 0, which is typically around 07:00:00. At this point, the simulation starts and we proceed to the first step of the cycle.

The first step involves sampling the time until the next arrival from an exponential distribution with rate $\lambda(t)$, where $\lambda(t)$ corresponds to the arrival rate for the weekday and week of 01/01 at the seventh hour. We then compute the next arrival time by adding the sampled inter-arrival time to the current time and we can go to step two. If the resulting arrival time falls within the next hour, it needs to be adjusted. If the arrival rate in the next hour is higher, we expect more patients during that hour, so the arrival time will be shifted slightly earlier. Conversely, if the arrival rate is lower, the arrival time will be adjusted to a later time. This adjustment is calculated as follows:

$$\text{New arrival time} = \text{start next hour} + \text{minutes in next hour} \cdot \frac{\lambda(\text{current hour})}{\lambda(\text{next hour})} \quad (9)$$

We illustrate Equation 9 with an example. Assume the simulation starts at 07:00, and we sample an inter-arrival time of 70 minutes. This results in the next arrival being scheduled for 07:00 + 70 minutes = 08:10. Since this falls into the next hour, a correction is required. The start of the next hour is 08:00, and 10 minutes are remaining in that hour. Suppose that $\lambda(07{:}00\ 01/01) = 2$ and $\lambda(08{:}00\ 01/01) = 4$. Because the arrival rate in the next hour is higher, the arrival time will be adjusted to an earlier time. Using Equation 9, we calculate the corrected arrival time as follows:

$$\text{New arrival time} = 08{:}00 + 10 \cdot \frac{2}{4} = 08{:}05$$

After determining the inter-arrival time in step one and potentially correcting the arrival time in step two, we proceed to step three, where we sample the length of stay for that arrival using the empirical distribution. To do this, we first sample a random number between 0 and 1 from a uniform distribution.

This value is then used as a quantile to obtain the length of stay from the empirical distribution.

In step four, we add the sampled length of stay to the arrival time to compute the departure time. At this point, one patient's simulation is complete, and we add both the arrival and departure times to the event list in step five. The final step of the cycle involves jumping to the time of the sampled arrival. In the previous example, the sampled arrival time was 08:05, so we jump from 07:00 to 08:05 and begin the process again from step one. This cycle is repeated until the simulation reaches the end of the year.

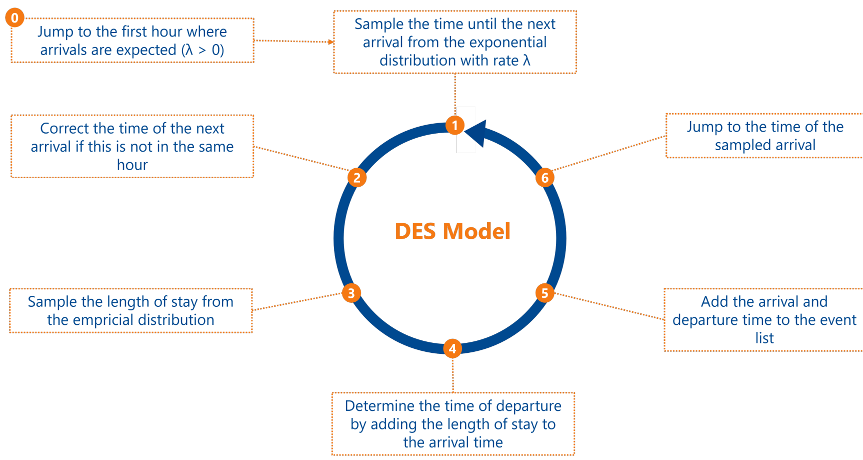


Figure 18: Discrete Event Simulation Model

The simulation output is an event list with the arrival and departure times of all simulated patients for each department, as shown in Table 7.

	Event List	
Ward	**Time of Arrival**	**Time of Departure**
Ward 1	08:05:00 01/01/25	12:10:23 01/01/25
Ward 1	08:23:37 01/01/25	14:16:48 01/01/25
Ward 1	08:45:56 01/01/25	13:48:56 01/01/25
Ward 1	...	...
Ward 1	15:02:18 31/12/25	16:58:33 31/12/25

Table 7: Example event list

This simulation output serves as the basis for determining the minimum number of beds c required to achieve a predefined blocking probability $P_{\max}$. Specifically, we aim to find the lowest bed capacity c for which the blocking probability $P_c \leq P_{\max}$. To achieve this, we iteratively increase the number of available beds, processing patient arrivals and departures chronologically. For each bed capacity c, we track which patients are admitted and which are blocked under the assumption that blocked patients leave the system. By monitoring the number of occupied beds at each event, we calculate the blocking probability as:

$$P_c = \frac{\text{Number of Blocked Patients}}{\text{Number of Blocked Patients} + \text{Number of Admitted Patients}} \tag{10}$$

The approach differs for clinical and day treatment departments due to their distinct demand patterns:

- For **clinical departments**, we account for seasonal variations by evaluating the blocking probability based on the busiest quarter of the year. Specifically, we ensure that the mean blocking probability over all days in the busiest quarter remains below the maximum desired blocking probability $P_{\max}$.
- For **day treatment departments**, patient demand follows a weekly and hourly pattern. Therefore, we compute the mean blocking probability per hour and weekday and ensure that the blocking probability of the busiest hour in the week does not exceed $P_{\max}$.

6.4 Cluster to ward assignment model

In this final section of the methodology, all components discussed in the previous sections come together to form the input of the cluster-to-ward assignment model. We use the following input:

- Available beds per ward
- Required beds per cluster
- Ward preferences per specialty
- Need to be close to the OR per specialty
- Distances between wards and between each ward and the OR

The goal is to determine which specialties should be clustered together and to which wards they should be assigned. We do this by formulating the problem as an Integer Linear Program (ILP). Table 8 gives an overview of the sets, decision variables, and parameters. The optimization problem is solved in Python using Pyomo and the HiGHS Solver, which is well-suited for large-scale linear programming problems. Although R is the standard programming language within the R&D Capacity Management team, its limitations in handling large-scale ILPs and the availability of fewer high-performance solvers made Python the more suitable choice for this application.

Since the input differs between clinical and day treatment departments, we solve the model twice: once for clinical departments and once for day treatment departments. We also divide the set of wards W based on the current clinical ward layout, ensuring that wards designated for example for day treatment are only available for that purpose in the model. While we acknowledge that this approach limits the flexibility to reassign wards or combine day treatment and clinical departments, addressing this issue falls outside the scope of the current research and will be discussed further in Section 10. By dividing the wards between the two types of admission departments, we improve the model's performance by restricting the possible cluster-to-ward assignments. This results in a significant reduction in the number of decision variables.

The following subsections will provide a more detailed explanation of the ILP model. Section 6.4.1 discusses the decision variables, Section 6.4.2 presents the objective function, and lastly, the constraints are presented in Section 6.4.3.

	Sets
S	Set of specialties
C	Set of possible clusters of specialties
W	Set of available wards
	Decision Variables
$X_{c,w} \in \{0,1\}$	Binary variable, equal to 1 if cluster $c \in C$ is assigned to ward $w \in W$, and 0 otherwise
$Y_c \in \{0,1\}$	Binary variable, equal to 1 if cluster $c \in C$ is assigned to any ward, and 0 otherwise
$C_{\max} \in \mathbb{Z}_{\geq 0}$	Continuous variable, representing the size of the largest assigned cluster
	Parameters
$h_{s,c}$	Binary variable, equal to 1 if specialty $s \in S$ is in cluster $c \in C$, and 0 otherwise
b_w	The number of available beds at ward $w \in W$
r_c	The number of required beds for cluster $c \in C$
$m_{s,w}$	Binary variable, equal to 1 if specialty $s \in S$ prefers to be assigned to ward $w \in W$, and 0 otherwise
t_s	Binary variable, equal to 1 if specialty $s \in S$ prefers to be assigned close to the OR, and 0 otherwise
d_w	Distance from ward $w \in W$ to the OR
$d_{\max}$	Maximum allowed distance from any ward to the OR
$p_{\max}$	Maximum allowed distance between any two wards within the same cluster
$p_{w,v}$	Binary variable, equal to 1 if the distance between ward $w \in W$ and $v \in W$ does not exceed $p_{\max}$

Table 8: Notation of the sets, decision variables, and parameters

6.4.1 Decision Variables

The goal is to determine which specialties should be assigned to which wards. In our modeling approach, a significant portion of this optimization is handled during the preprocessing phase, before solving the ILP. As explained in the previous sections, during this preprocessing, we identify the possible clusters of specialties and calculate the required number of beds for each cluster. In the ILP, we then determine which clusters are assigned to which wards, which, in turn, indicates the specialties allocated to each ward. We introduce the binary variable $X_{c,w}$, equal to 1 if cluster $c \in C$ is assigned to ward $w \in W$, and 0 otherwise.

A specialty can belong to multiple clusters. In our ILP model, the objective is to select a subset of these clusters, rather than all of them, since each specialty should be assigned to only one cluster. This implies that not every cluster will be assigned to a ward. To determine which clusters are assigned, and which ones not, we introduce another binary variable Y_c, equal to 1 if cluster $c \in C$ is assigned to any ward, and 0 otherwise. By combining these two variables, the ILP model can efficiently determine the optimal specialty assignments while respecting the constraints of the problem.

We introduce an additional decision variable, $C_{\max}$, which is a continuous variable representing the size of the largest assigned cluster. This variable is added to prevent the model from assigning all specialties into a single large cluster, even if the compatibility matrix permits such an arrangement. While the compatibility matrix defines feasible clusters, the introduction of $C_{\max}$ ensures that the clustering remains balanced and avoids the creation of excessively large clusters.

6.4.2 Objective Function

In order to ensure that the ILP finds the optimal solution, the following objective function is used in the model:

$$\min \sum_{c \in C} r_c Y_c + C_{\max} \tag{11}$$

This objective function is designed to minimize two important components: the total number of required beds across all clusters assigned to wards and the maximum cluster size. Each of these components plays a critical role in ensuring the efficient allocation of resources within the hospital.

The first term represents the total number of required beds, denoted by r_c, where beds are counted only if cluster c is assigned to a ward (as indicated by the binary variable Y_c). In other words, this term captures the total bed requirement for the proposed clinical ward layout. By combining specialties effectively, fluctuations in bed occupancy can be balanced, as peaks in demand for one specialty may be offset by lower occupancy in another. This stabilization can lead to a reduced overall bed requirement, potentially resulting in a more balanced workload for nurses or even a reduction in the total number of nurses needed.

Naturally, minimizing the number of beds would suggest grouping all specialties into a single cluster, as this could lead to the most efficient use of space. This approach is not practical in hospital settings, where maintaining specialized care and avoiding overcrowding is essential. While the compatibility matrix helps by restricting certain combinations of specialties, the introduction of the decision variable $C_{\max}$ provides an additional safeguard. The second term of the objective function therefore represents minimizing the maximum cluster size $C_{\max}$. It prevents the creation of excessively large clusters, ensuring that specialty care remains adequately specialized and practical for hospital operations, while still aiming to minimize bed usage.

Another approach we considered was using a weighted objective function, where the walking distances between wards and the OR, as well as the distances between wards within the same cluster, would be minimized as part of the objective rather than enforced through constraints. This would mean treating some constraints as soft constraints instead of hard constraints. However, we opted for hard constraints because they significantly improve computational efficiency. By explicitly restricting the feasible region, the model has fewer possibilities to explore, leading to faster optimization. Additionally, using soft constraints would require assigning appropriate weights to these distance penalties in the objective function, introducing the challenge of determining suitable weight values.

6.4.3 Constraints

In order to find a feasible solution, we need to set some constraints. This section will discuss each constraint.

To ensure a feasible assignment, each specialty $s \in S$ must be assigned to exactly one ward. Since we assign clusters to wards rather than specialties directly, we introduce a binary parameter $h_{s,c}$, which equals 1 if specialty s is part of cluster c. To illustrate this, consider an example where four clusters contain specialty CHI:

$$\{\text{CHI}\}, \{\text{CHI, GYN}\}, \{\text{CHI, MDL}\}, \{\text{CHI, GYN, MDL}\}$$

For all four clusters that include CHI, we have $h_{CHI,c} = 1$. However, assigning multiple clusters containing CHI to a ward would result in the specialty being placed in multiple locations in the hospital, which is not allowed. Instead, CHI must be assigned exactly once. To enforce this, the model ensures that $Y_c = 1$ or exactly one of these four clusters, while for the remaining three $Y_c = 0$. This requirement is formalized in Constraint 12:

$$\sum_{c \in C} h_{s,c} Y_c = 1 \qquad \forall s \in S \tag{12}$$

A ward can accommodate multiple specialties if they belong to the same cluster. However, a ward cannot be shared by multiple clusters. To enforce this, we ensure that no more than one cluster is assigned to each ward, denoted by Constraint 13:

$$\sum_{c \in C} X_{c,w} \leq 1 \qquad \forall w \in W \tag{13}$$

For each cluster $c \in C$, we compute the required number of beds r_c based on a predefined blocking probability. We introduce variable b_w as the number of available beds in ward $w \in W$. To ensure the blocking probability is met, the total number of available beds in the ward(s) to which cluster c is assigned must be greater than or equal to the required number of beds. This condition is enforced by Constraint 14. The term Y_c on the right-hand side acts as an activation condition. If cluster c is not assigned to any ward ($Y_c = 0$), the right-hand side becomes zero, effectively deactivating the constraint for that cluster. This means the constraint only applies when $Y_c = 1$, ensuring that it is relevant only for clusters that are actively assigned to a ward.

$$\sum_{w \in W} X_{c,w} b_w \geq r_c Y_c \qquad \forall c \in C \tag{14}$$

The decision variable Y_c indicates whether cluster c is assigned to any ward, and the decision variable $X_{c,w}$ indicates whether cluster c is assigned to ward w. Constraints 15 and 16 ensure consistency between Y_c and $X_{c,w}$. The first constraint ensures that if $X_{c,w} > 0$ for any w, then Y_c must equal 1. The second constraint ensures that if cluster c is not assigned to any ward ($X_{c,w} = 0$ for all $w \in W$), then Y_c must be equal to 0. This guarantees that $Y_c = 1$ if and only if cluster c is assigned to at least one ward.

$$\sum_{w \in W} X_{c,w} \leq |W| Y_c \qquad \forall c \in C \tag{15}$$

$$Y_c \leq \sum_{w \in W} X_{c,w} \qquad \forall c \in C \tag{16}$$

A cluster is defined by a set of specialties that can be assigned to a ward. However, to ensure that clusters remain within a manageable size and prevent the formation of overly large clusters, we introduce a constraint that limits the maximum size of a cluster. By minimizing $C_{\max}$ in the objective function and bounding it through this constraint, we effectively limit the size of the largest cluster, balancing the reduction of required beds with the need for practical and specialized care in hospital settings. The constraint is formulated as follows:

$$C_{\max} \geq \sum_{s \in S} h_{s,c} Y_c \qquad \forall c \in C \tag{17}$$

The last three constraints address hospital-specific settings and incorporate user-defined input to tailor the model to practical requirements. These constraints ensure that assignments respect logistical and spatial considerations within the hospital environment.

Constraint 18 enforces that if a specialty s must be assigned to a specific ward w (indicated by $m_{s,w} = 1$), then the cluster c containing specialty s must be assigned to that ward. This constraint ensures that all mandatory assignments are respected by the model.

$$\sum_{c \in C} h_{s,c} X_{c,w} \geq m_{s,w} \qquad \forall s \in S, \forall w \in W \tag{18}$$

Certain specialties require close proximity to the operating room (OR) due to the nature of their procedures or the need for quick access in emergencies. To model this, we introduce the parameter t_s, which is set to 1 for specialties that require close proximity to the OR and 0 otherwise. For these specialties, we ensure that the distance between the assigned ward w and the OR (denoted by d_w) does not exceed a predefined maximum $d_{\max}$. Constraint 19 guarantees that wards assigned to these specialties meet the necessary proximity requirements.

$$d_w t_s h_{s,c} X_{c,w} \leq d_{\max} \qquad \forall s \in S, w \in W, c \in C \tag{19}$$

Finally, Constraint 20 ensures that a single cluster $c \in C$ cannot be assigned to two wards $w \in W$ and $v \in V$ if the distance between these two wards exceeds a predefined maximum allowable distance $p_{\max}$. The parameter $p_{w,v}$ is a binary variable, equal to 1 if the distance between wards w and v does not exceed $p_{\max}$, and 0 otherwise. The left-hand side of the constraint $X_{c,w} + X_{c,v} - 1$ equals 1 only if the cluster c is assigned to both wards w and v. In this case, the constraint requires $p_{w,v}$ to be 1, ensuring that the distance between these two wards satisfies the maximum distance requirement. If $p_{w,v} = 0$ meaning the distance between w and v exceeds $p_{\max}$, the constraint prevents the simultaneous assignment of cluster c to both wards, maintaining the spatial feasibility of ward assignments.

$$X_{c,w} + X_{c,v} - 1 \leq p_{w,v} \qquad \forall c \in C, w \in V, v \in W \tag{20}$$

With these constraints in place, we ensure that each specialty is assigned to an appropriate ward while adhering to hospital-specific requirements. Appendix C shows the complete ILP formulation.

7 Results Development Data

In this section, we present the results of applying the bed requirements model and the cluster-to-ward assignment model. We base all results in these sections on ChipSoft's development dataset, the same as we performed our data analysis on in Section 5. In the Section 8, we will conduct a case study to further illustrate the results. It is important to note that in all results bed requirements are based on a blocking probability of 5%.

We start by analyzing bed requirements for clinical departments in Section 7.1 and day treatment departments in Section 7.2. Then, we will examine the computational performance in Section 7.3 and the results of the cluster-to-ward assignment model in Section 7.4.

7.1 Bed Requirements for Clinical Departments

In Section 6.3, we defined two methods for calculating bed requirements for clinical departments: the Erlang loss model and the Discrete Event Simulation (DES) model. We compute bed requirements for each cluster, but given the large number of possible combinations, evaluating each in detail in this section is impractical. We therefore analyze bed requirements first for each specialty individually and then select a few clusters to analyze.

Table 9 presents the required number of beds along with the corresponding occupancy percentages for both the Erlang loss and DES models. For six out of fifteen specialties, both models yield the same number of beds. For the remaining specialties, the differences are generally small, with a maximum discrepancy of two beds. While expert opinion suggests that differences become practically significant at four beds or more, a discrepancy of two beds may still be relevant. Notably, the Erlang loss model tends to estimate higher bed requirements for most specialties. This may be due to its assumption of a consistently high arrival rate in a stationary system. This could result in a slightly busier scenario compared to the simulation model. Furthermore, the Erlang loss model uses the mean length of stay to estimate bed requirements, whereas the DES model relies on the empirical distribution. As highlighted in Section 5.2, some specialties exhibit skewed distributions with outliers, where a few patients have exceptionally long stays. Because the Erlang model is based on the mean length of stay, it is more sensitive to these outliers, which can lead to higher required bed capacity for such specialties.

Table 9 also presents the occupancy rate, a key metric that represents the proportion of available beds occupied by patients. For each specialty, the daily occupancy rate is first determined using historical data by dividing the maximum number of occupied beds on that day by the number of available beds. The final value reported in the table is obtained by averaging these daily occupancy rates over the entire observation period.

A higher occupancy rate suggests more efficient bed utilization but may also indicate potential capacity constraints, while a lower occupancy rate may imply underutilization of resources. We can see that in general, specialties with a lower bed requirement (fewer than ten beds) tend to have lower occupancy percentages, typically around or below 50%. In contrast, for specialties with a higher bed requirement, occupancy rates range between approximately 65% and 88%. The maximum number of occupied beds for each day in 2019 can be seen in Appendix D together with the required beds found by DES and Erlang models for each specialty.

	Erlang loss Model		*DES Model*	
Specialty	**Beds**	**Occupancy %**	**Beds**	**Occupancy %**
ANE	2	15.5%	2	15.5%
CAR	46	83.7 %	44	87.5%
CHI	35	79.4 %	35	79.4%
GER	20	65.0 %	18	72.2%
GYN	5	62.8%	6	52.4%
INT	49	83.7%	47	87.3%
KAA	5	39.9%	5	39.9%
KNO	5	46.6%	5	46.6%
LON	39	83.4%	40	81.3%
MDL	24	79.3%	22	86.5%
NCH	14	64.9%	13	69.9%
NEU	29	83.2%	29	83.2%
ORT	32	76.1%	31	78.6%
PLA	7	51.8%	7	51.8%
URO	17	78.5%	18	74.1%

Table 9: Bed requirements and corresponding occupancy percentage per specialty for Erlang loss and DES model within clinical departments

We select a subset of specialties—ANE, CHI, GYN, and ORT—to form clusters and analyze how the models perform when multiple specialties share a ward. Since these specialties are fully compatible, this results in 15 possible clusters. Table 10 presents the estimated bed requirements for each cluster, comparing results from both the Erlang loss and DES models. The "Separate" column sums the bed requirements for each specialty placed individually, while the "Combined" column shows the requirements when the specialties are placed together in one ward. Since clusters 1-4 only contain one specialty, the combined bed requirements are not provided.

Comparing the results of the Erlang loss and DES models in Table 10, we observe that the estimated number of required beds is generally very similar, with a maximum discrepancy of two beds. Focusing on clusters 5–15 (which include only the combined clusters), the estimates are identical for three clusters, differ by one bed for five clusters, and differ by two beds for the remaining three clusters. Notably, there is no consistent trend regarding which model predicts a higher or lower bed requirement; in some cases, the Erlang loss model estimates a higher number of beds, while in others, the DES model provides the higher estimate.

Cluster	**Specialties in Cluster**				**Erlang loss Model (Beds)**		**DES Model (Beds)**	
	ANE	CHI	GYN	ORT	Separate	Combined	Separate	Combined
1	●				2	-	2	-
2		●			35	-	35	-
3			●		5	-	6	-
4				●	32	-	31	-
5	●	●			37	35	37	36
6	●		●		7	6	7	7
7	●			●	34	33	33	31
8		●	●		40	37	41	37
9		●		●	67	61	66	61
10			●	●	37	34	37	35
11	●	●	●		42	37	43	37
12	●		●	●	39	34	39	36
13	●	●		●	69	62	68	60
14		●	●	●	72	63	72	62
15	●	●	●	●	74	63	74	64

Table 10: Bed requirements per cluster for the Erlang loss and DES models, showing both the requirements of placing the specialties separate and combined

Furthermore, we analyze the impact of placing specialties in separate wards versus combining them within a single ward for each cluster. The difference in required beds between these two approaches ranges from 0 to 11 beds. If all four specialties were assigned to separate wards, a total of 74 beds would be required. However, when all four specialties are combined into a single ward, that cluster requires only 63 beds according to the Erlang loss model and 64 beds according to the DES model. This pooling effect can therefore result in a reduction of up to 11 beds. We can see that in general pooling effects become more pronounced as both the cluster size and the sizes of the specialties within the cluster increase.

Lastly, we analyze the performance of an existing cluster of four specialties—KAA, KNO, PLA, and URO—that share a ward in the development data in Figure 19. The ward has 27 available beds. The Erlang loss model estimates the required number of beds at 25, while DES estimates 26, both of which are very close to the available capacity. We observe that at some moments in time the occupancy exceeded the number of available beds, indicating capacity shortages during peak moments.

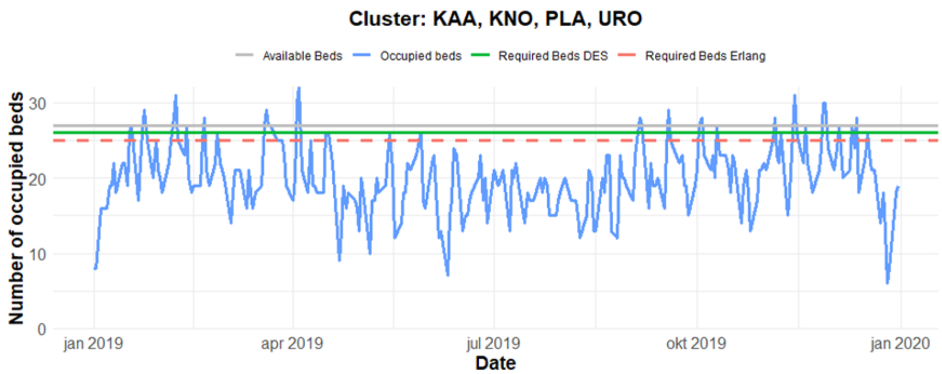


Figure 19: Occupied, required and available beds for an existing cluster in ChipSoft's development data: KAA, KNO, PLA, URO

7.2 Bed Requirements for Day Treatment Departments

In Section 6.3, we introduced two approaches for determining bed requirements in day treatment departments: the time-dependent Erlang loss model and the Discrete Event Simulation (DES) model. This section presents a comparative analysis of the results obtained from both models. Given the large number of possible specialty clusters, we first examine the bed requirements for each specialty individually before selecting a specific cluster for a more detailed analysis.

Table 11 presents the required number of beds along with the corresponding occupancy percentages for both the time-dependent Erlang loss and DES models. The difference in bed estimates between the time-dependent Erlang loss model and DES is at most 2 beds. For 5 out of 17 specialties, both models yield the same estimate. In 10 cases, the difference is just 1 bed, while for 2 specialties, the estimates differ by 2 beds. Notably, neither model consistently produces higher or lower estimates than the other. We computed the occupancy rates similar to clinical departments, which range from 5.0% to 95.4%. The low occupancy rates are for the very small specialties. Appendix E presents plots for all specialties with the maximum occupation per day and the bed estimates of both models for each day in 2019.

	Erlang loss Model		*DES Model*	
Specialty	**Beds**	**Occupancy %**	**Beds**	**Occupancy %**
ANE	2	7.9%	1	15.7%
CAR	11	50.2%	12	46.0%
CHI	10	53.6%	12	44.6%
GYN	5	33.5%	4	41.9%
INT	26	91.7%	25	95.4%
KAA	3	5.0%	1	14.9%
KNO	6	36.7%	5	44.0%
LON	11	55.6%	12	51.0%
MDL	17	68.0%	16	72.3%
NCH	2	8.2%	1	16.5%
NEU	4	37.1%	4	37.1%
OOG	11	49.3%	11	49.3%
ORT	9	45.5%	10	41.0%
PLA	6	25.5%	6	25.5%
REU	5	48.0%	5	48.0%
TAN	4	10.6%	4	10.6%
URO	6	35.2%	7	30.2%

Table 11: Bed requirements and corresponding occupancy percentage per specialty for Erlang loss and DES model for day treatment

In practice, these specialties within day treatment departments mostly do not have their own dedicated ward, but are combined with other specialties. A common department within day treatment is often one ward with all surgical specialties. We therefore analyze a cluster of 8 different surgical specialties: CHI, GYN, KNO, NCH, OOG, ORT, PLA, URO. The results of the bed requirements are presented in Table 12 and show that the DES model estimates an additional 4 beds and a resulting difference of around 10% in occupancy percentage.

	Erlang loss Model		*DES Model*	
Cluster	**Beds**	**Occupancy %**	**Beds**	**Occupancy %**
Surgical Cluster	26	76.6%	30	66.4%

Table 12: Bed requirements and occupancy percentage for the surgical cluster

Figure 20 illustrates the bed requirements for both models, along with the maximum daily occupancy for the surgical clusters. This figure shows that 4 additional beds estimated by the DES model allow it to capture more peak occupancy moments compared to the time-dependent Erlang loss model. The influence of holidays is also evident, as bed occupancy drops to (almost) zero on these days. Additionally, a slight decline in occupancy can be observed around August, possibly due to the summer holiday.

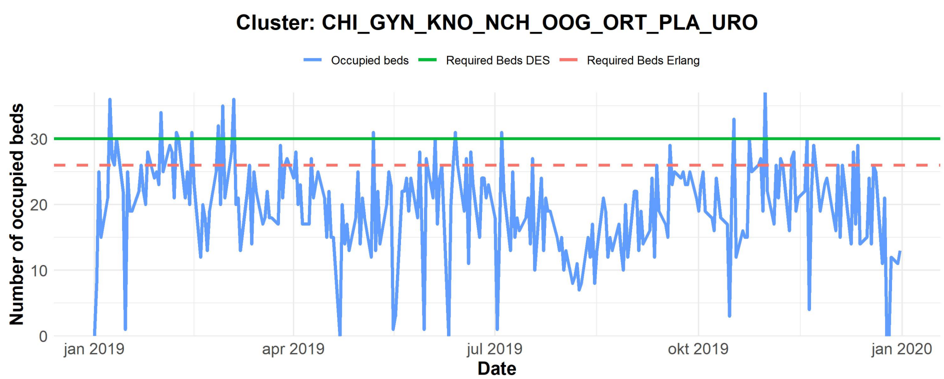


Figure 20: Bed estimation and maximum daily occupation for surgical day treatment cluster

7.3 Computational Performance

It is important that the bed requirements model can solve the problem in a reasonable amount of time, as excessively long runtimes would make it impractical for real-world use. We compare the computation time of all proposed models in terms of run time of one cluster, the results can be seen in Table 13 for each model and type of department.

Department	**Model**	**Runtime per Cluster**
Clinical	Erlang loss DES	0.05 sec 125 sec
Day Treatment	Time-dependent Erlang loss DES	0.67 sec 160 sec

Table 13: Average runtime of a single cluster for Erlang loss and DES model

The results show that both the regular Erlang loss model and the time-dependent Erlang loss model are much faster than the DES model. Even with 1000 clusters, the Erlang loss model completes its computation in under a minute and the time-dependent Erlang loss model in around 11 minutes. Computing bed requirements for 1000 clusters with DES results in a runtime of more than 34 hours for clinical departments and 44 hours for day treatment departments. While strategic-level models are not run frequently, excessively long runtimes reduce usability, making it impractical for the model to take so many hours to complete.

7.4 Cluster to Ward Assignments

In this section, we analyze the results of the ILP model presented in Section 6.4, where clusters of specialties are assigned to wards. Since we have no user input for the ChipSoft development data, the results are generated using hypothetical user input. This section will therefore primarily serve to demonstrate the model's ability to generate feasible solutions. Section 8 will provide a more comprehensive evaluation of the model's performance using real user input.

For the clinical departments, we use the compatibility matrix from Section 6.1 (Table 4) and for day treatment departments, we assume that all surgical specialties are mutually compatible, as are all non-surgical specialties. This assumption leads to 595 possible clusters within clinical departments and 702 possible clusters within day treatment departments. We assume that all surgical specialties must be located near the operating room, while non-surgical specialties do not have this constraint. Additionally, we assume that the neurology and cardiology clinical department are required to remain in their currently assigned ward. As detailed in Section 5.4, there are a total of 20 available wards, with 493 beds distributed across these wards. Using this data, along with the computed bed requirements for each cluster, we input the information into the ILP model.

The complexity of the problem is reflected by the number of decision variables. We have 8331 decision variables for clinical departments and 5617 for day treatment departments. The results of the ILP are presented in terms of the objective value, which represents the required bed estimates plus the size of the largest cluster, and the runtime of the ILP solver in Table 14.

Department	**Model**	**Objective**	**Runtime**
Clinical	Erlang loss	306	133.4 sec
	DES	297	49.2 sec
Day Treatment	Time-dependent Erlang loss	83	39.5 sec
	DES	79	30.6 sec

Table 14: Objective value and runtime of cluster-to-ward assignment model

The objective values within each department type are relatively similar, with slightly lower values observed when the DES model is used for bed requirement estimation. While a lower objective value may indicate an improved solution in the context of a minimization problem, it does not necessarily correspond to a better overall outcome. If the DES model would estimate lower bed requirements, the resulting objective value is inherently reduced. Consequently, a more detailed analysis of the resulting clinical ward layout is required to fully understand these differences. Regarding computational efficiency, three solutions were obtained within one minute, while one required slightly more than two minutes. We will discuss the result in more detail for clinical departments and day treatment departments seperately.

Clinical Departments

Figures 21 and 22 illustrate the assignment of clusters to wards based on Erlang and DES bed estimates respectively. The horizontal axis represents the clusters of specialties. The light blue bar indicates the available beds in the wards to which the cluster is assigned, while the dark blue bar represents the required number of beds for the cluster.

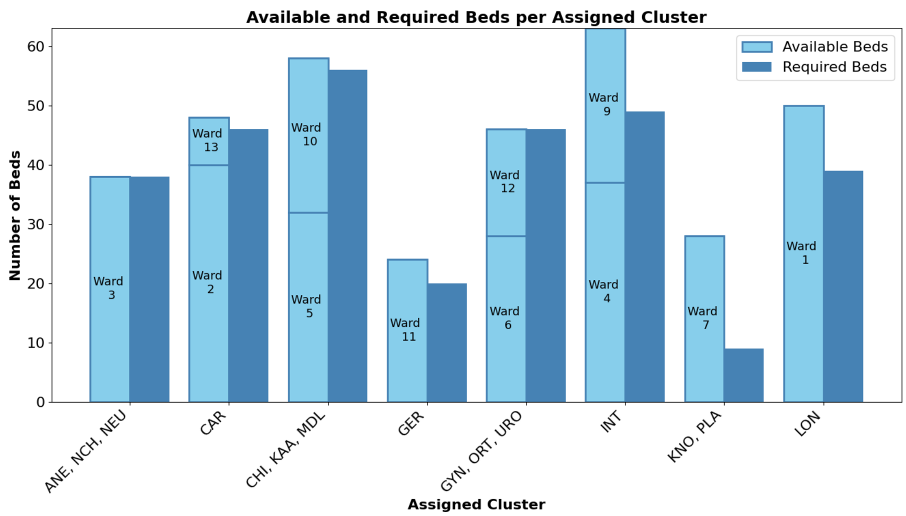


Figure 21: Clinical Ward Layout based on Erlang bed estimates

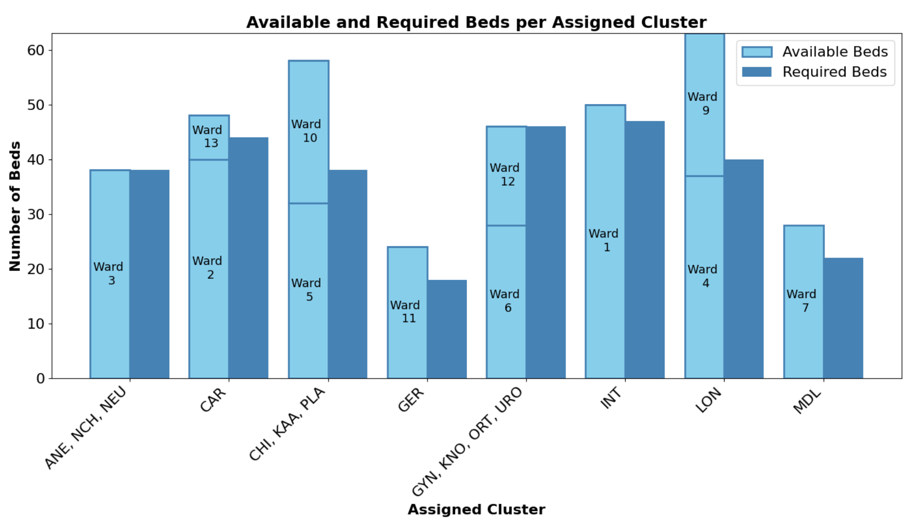


Figure 22: Clinical Ward Layout based on DES bed estimates

While the clinical ward layout is not identical, both layouts form 8 clusters, of which 5 are identical: {ANE, NCH, NEU}, {CAR}, {GER}, {INT}, {LON}. Only the wards to which the clusters {INT} and {LON} are assigned are switched. The ward ID's correspond to how the wards are defined in Figure 16 in Section 5.4, where ward 1 corresponds to the wards with the most beds. The required beds per cluster can be seen in Table 15 for Erlang bed estimates and Table 16 for DES bed estimates. Here we can indeed see that for the similar clusters DES sometimes estimates less beds.

Cluster	Beds
ANE, NCH, NEU	38
CAR	46
CHI, KAA, MDL	56
GER	20
GYN, ORT, URO	46
INT	49
KNO, PLA	9
LON	39
Total	**303**

Table 15: Required beds per cluster (Erlang bed estimates)

Cluster	Beds
ANE, NCH, NEU	38
CAR	44
CHI, KAA, PLA	38
GER	18
GYN, KNO, ORT, URO	46
INT	47
LON	40
MDL	22
Total	**293**

Table 16: Required beds per cluster (DES bed estimates)

Day Treatment Departments

Figures 23 and 24 depict the assignment of day treatment clusters to wards based on Erlang and DES bed estimates, respectively. The assigned clusters are nearly identical, with only URO and ORT switched, and differences in the wards to which they are assigned. Both solutions result in two surgical clusters and two non-surgical clusters. The required number of beds per cluster is provided in Table 17 for Erlang bed estimates and Table 18 for DES bed estimates.

Cluster	Beds
ANE, NEU, TAN	5
CAR, INT, LON, MDL, REU	43
CHI, GYN, KAA, NCH, ORT	16
KNO, OOG, PLA, URO	14
Total	**78**

Table 17: Required beds per cluster (Erlang bed estimates)

Cluster	Beds
ANE, NEU, TAN	3
CAR, INT, LON, MDL, REU	38
CHI, GYN, KAA, NCH, URO	15
KNO, OOG, ORT, PLA	18
Total	**74**

Table 18: Required beds per cluster (DES bed estimates)

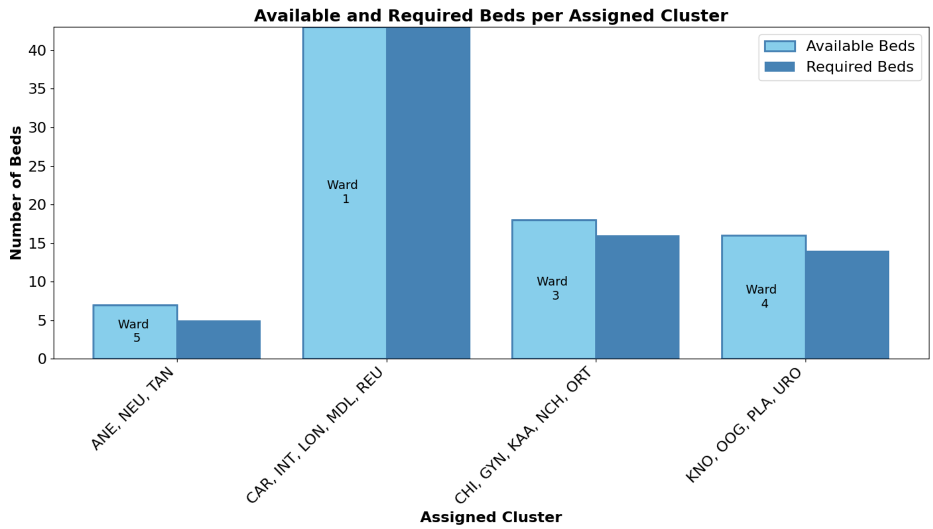


Figure 23: Day Treatment Ward Layout based on Erlang bed estimates

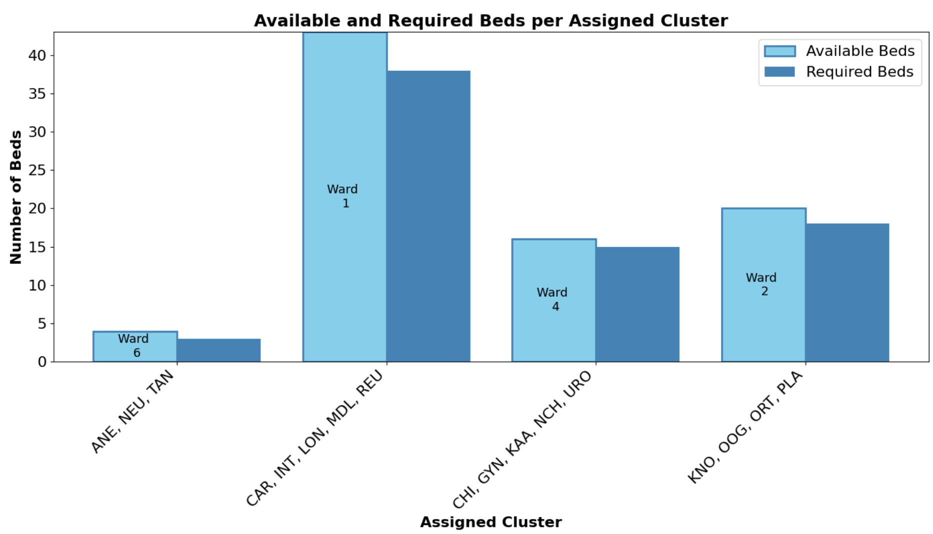


Figure 24: Day Treatment Ward Layout based on DES bed estimates

8 Case Study

In this chapter, we conduct a case study in collaboration with a medium-sized regional hospital. Unlike in Section 7, where we relied on general assumptions for the user input, this case study allows us to incorporate real user input. For this study, we spoke with the manager of the hospital's capacity management department and several capacity advisors. Key topics included strategies for optimizing bed occupancy, the assignment of specialties across wards, and constraints related to medical and logistical requirements. Additionally, we examined the hospital's current tools and methodologies for determining bed requirements. Beyond sharing insights into their existing practices, the hospital provided us with the necessary user input and granted permission to run our model on their data. This section is structured as follows. Section 8.1 describes the application of the methodology. Section 8.2 presents the results of the case study and lastly, Section 8.3 discusses the assessment of the case study hospital on the results.

8.1 Application of the Methodology

We follow the methodology described in Section 6, which is divided into four steps: (1) processing data from HiX, (2) processing user input, (3) calculating bed requirements, and (4) assigning clusters to wards.

1. Processing Data from HiX

We were given access to the hospital's database containing data up to the end of November 2023. Although this is not the most recent dataset, hospital staff confirmed that no significant changes have occurred, making it representative for modeling purposes. Since we require one year of data, we selected the period from November 2022 to November 2023. Using SQL queries, we extracted the admissions and bed capacity datasets from HiX and imported them into RStudio for model implementation.

To clean the admissions dataset, we removed departments that were excluded from the analysis based on discussions with the hospital. Similar to ChipSoft's development dataset, we excluded the Intensive Care Unit, Coronary Care Unit, Obstetric department, Pediatric department, and Neonatal department. Additionally, this hospital has an admission lounge, a day treatment ward designated for scopy, and a mental health care center (GGZ), which we also excluded. All these departments have distinct operational dynamics and staffing requirements that set them apart from regular clinical and day treatment wards.

Although traumatology is typically considered a subspecialty within general surgery, the hospital treats it as a separate specialty. As a result, traumatology patients are not necessarily assigned to the same ward as other general surgery patients. To distinguish these patients, we identify traumatology admissions based on their specific diagnosis codes. These codes are defined by the Nederlandse Zorg Autoriteit (NZa), the regulatory body overseeing healthcare markets in the Netherlands. The NZa establishes standardized diagnosis codes that classify various medical conditions, including those related to traumatology. All admissions categorized under the specialty of general surgery and with a diagnosis code related to traumatology are classified as TRAU.

Oncology is another example of a subspecialty that spans several departments but is treated as a distinct group within the hospital. The hospital explains that this group specifically includes patients admitted for cytostatic treatment, which refers to therapies that inhibit the growth and division of cancer cells, commonly used in chemotherapy. They categorize these admissions under the admission type "Cytostica", enabling us to assign the specialty ONC to all admissions with this type.

After excluding the specified wards and defining the two extra specialties, the dataset now contains a total of 16 specialties. Table 19 provides an overview of the specialties assigned to clinical wards and day treatment wards. Each specialty is listed in the first column, while the second and third columns indicate whether the specialty operates within clinical departments, day treatment departments, or both. A filled circle (●) signifies that the specialty is present in the corresponding department type.

Specialty	**Clinic**	**Day Treatment**
ANE		●
CAR	●	●
CHI	●	●
GYN	●	●
INT	●	●
KAA		●
KNO	●	●
LON	●	●
MDL	●	●
NEU	●	●
ONC		●
OOG		●
ORT	●	●
REU		●
TRAU	●	
URO	●	●

Table 19: Specialties operating in clinical and day treatment departments

We also exclude the specified departments from the bed capacity dataset, resulting in 11 wards: 3 designated for day treatment and 8 for clinical departments. In total, these wards offer 235 available beds. Figure 25 illustrates the current clinical ward layout, displaying the locations of the wards, the number of available beds in each, and the specialties assigned to them.

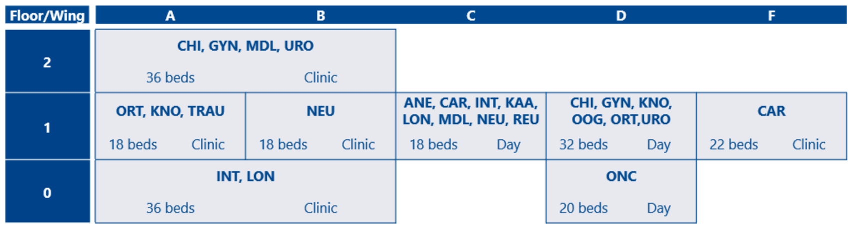


Figure 25: Current layout of clinical and day treatment wards

2. Processing User Input

The hospital provided us with all components of the required user input.

- **Compatibility Matrix**: The compatibility matrix for clinical departments is presented in Appendix F. For day treatment, all specialties, except Oncology, are generally permitted to be assigned to the same ward. However, a distinction is currently made between surgical and non-surgical specialties.

- **Ward Preferences**: Neurology must be assigned to ward B1 due to the presence of the stroke unit. Additionally, Cardiology needs to be placed in ward F1, as it must be located next to the Cardiac Care Unit (CCU).

- **Close to OR**: In principle, all surgical specialties are required to be located near the operating room (OR). However, the hospital indicated that since it is not a very large hospital, all wards are considered sufficiently close, and no additional restrictions are imposed.

- **Ward & OR Locations**: The locations of the wards are shown in Figure 25. If a cluster requires multiple wards, these must be located on the same floor in adjacent wings. Equation 1 is used to compute the distance, with a maximum allowed distance set to 1. As for the operating room (OR), its location is not a concern, as all wards are considered close enough.

3. Calculating Bed Requirements

The provided compatibility matrices are used to form clusters, for which bed requirements are computed. A blocking probability of 5% is applied. Section 6.3 introduced both the (time-dependent) Erlang and DES model for calculating bed requirements. For clinical departments, the compatibility matrix limits the number of possible clusters to 41. Given this relatively small number, we can feasibly run bed estimates using Discrete Event Simulation (DES) within a reasonable timeframe. Using Table 13 from Section 7.3, we can calculate that this process would take less than 1.5 hours. Despite the longer runtime compared to the Erlang model, we prefer using DES for this case study. This choice aligns with our goal of ensuring transparency in our methodology and making the results more interpretable for the hospital. The hospital emphasized the importance of understanding why a model predicts a certain number of beds or a specific ward layout. A simulation model is more intuitive and easier to explain than a mathematical formula like the Erlang loss model. DES allows us to track exactly when and how many patients are admitted and which of these patients had to be rejected. This can be presented to the hospital to enhance confidence in the model's predictions.

For day treatment departments, if we assume full compatibility among all specialties in day treatment departments, the number of possible combinations becomes 16,384. With an average runtime of 160 seconds per combination using DES, this would result in an excessively long run time. In contrast, using the time-dependent Erlang loss model would take approximately 6 hours, which is considered manageable. Consequently, when investigating the impact of full compatibility, bed requirements can only be estimated using the time-dependent Erlang loss model. Since the hospital communicated with us that in principle all specialties are allowed to be placed together, we do want to explore this option and compare it to their current day ward layout. This means that we analyze two scenarios:

1. Assume three fixed clusters: surgical, non-surgical and oncology
2. Assume full compatibility between all specialties, except oncology

Although we could use DES for the first scenario, we apply the time-dependent Erlang loss model in both cases to ensure that any differences in layout are not influenced by variations in bed estimates between the two models.

4. Assigning Clusters to Wards

With all input correctly in place, we can run the ILP model as defined in Section 6.4 to assign clusters to wards. We will do this once for clinical departments and once for day treatment departments.

8.2 Case Study Results

This section presents the results of applying the proposed methodology to the case study hospital. First, we analyze the outcomes for clinical departments, followed by an examination of the results for day treatment.

Clinical departments

Figure 26 presents the optimized layout of clinical wards, as determined by the ILP model using bed estimates from the DES model.

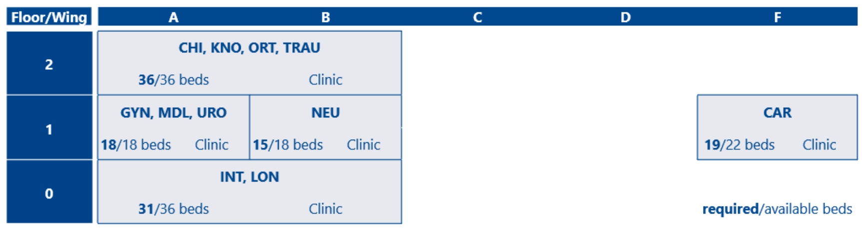


Figure 26: Optimized layout of clinical wards found by the model

The results show that two clusters have been adjusted. Wards A2 and B2 now belong to the cluster with CHI, KNO, ORT, and TRAU, whereas previously, KNO, ORT, and TRAU formed their own cluster and were assigned to ward A1. In the optimized solution, A1 is now grouped with specialties GYN, MDL, and URO. The remainder of the ward layout remains unchanged.

Table 20 presents the blocking probability and occupancy rate for each ward in the optimized clinical ward layout. The blocking probability is calculated over the busiest quarter. The occupancy rate is calculated by dividing the maximum number of occupied beds per day by the total number of available beds. The final occupancy rate reported in the table represents the average of these daily occupancy rates over the entire observation period. The observed occupancy rates range from 60.4% to 79.0%.

Ward	**Required Beds**	$P_{\mathbf{block}}$	**Occupancy Rate**
AB0	31	4.7%	79.0%
A1	18	3.7%	70.2%
B1	15	5.2%	68.7%
F1	19	0%	60.4%
AB2	36	3.3%	73.8%
Total	119	3.4%	70.4%

Table 20: Blocking Probability and Occupancy Rates for the optimized clinical ward layout

To investigate whether the results in Table 20 have actually improved, we need to calculate these results for the current clinical ward layout as well. Our analysis will now focus exclusively on the modified sections of the layout, specifically wards A1, A2, and B2, along with their assigned specialty clusters. Table 21 shows a detailed analysis for the current and optimized clinical ward layout. In the table, ward AB2 denotes the combination of wards A2 and B2. In both the current and optimized layouts, there are 18 beds in A1 and 36 in AB2.

Layout	**Ward**	P_{block}	**# Blocked**	**# Admitted**	**Occupancy %**
Current	A1	10.3%	175	1914	75.8%
	AB2	1.6%	36	3203	69.4%
Optimized	A1	3.7%	52	1771	70.2%
	AB2	3.3%	82	3423	73.8%

Table 21: Comparison of current and optimized clinical ward layout

Table 21 shows that ward A1 experiences a substantial improvement, with the blocking probability decreasing from 10.3% to 3.7%, resulting in a reduction of 123 blocked patients (from 175 to 52). Additionally, the occupancy rate decreases from 75.8% to 70.2%. Ward AB2 sees a slight increase in blocking probability from 1.6% to 3.3%, corresponding to an increase of 46 blocked patients (from 36 to 82). However, this remains below the acceptable threshold of 5%, indicating that the increase is manageable. Meanwhile, the occupancy rate rises from 69.4% to 73.8%. The difference between the occupancy rates on the two wards got smaller, suggesting a better distribution of bed demand. We can state that the optimized clinical ward layout causes an overall reduction of 77 blocked patients and a more balanced occupancy rate.

Current Clinical Ward Layout

Figure 27: Bed occupancy and blocked patients for current clinical ward layout

Figure 27 illustrates the bed occupancy trends throughout the year for the current clinical ward layout. The blue line represents the mean number of occupied beds per day, while the shaded light blue area indicates the daily variability in bed occupancy. The green line denotes the total number of available beds, establishing an upper limit that occupancy cannot exceed. The pink line represents the number of patients unable to be admitted due to full capacity. Ideally, this value should remain as low as possible. While the number of blocked patients is minimal for ward AB2, ward A1 experiences a significant number of blocked patients.

Figure 28 presents similar bed occupancy plots for the two wards, but now under the optimized clinical ward layout. A notable reduction in the number of blocked patients is observed for ward A1. While there is a slight increase in the number of blocked patients at ward AB2, this increase remains relatively minor. We can again observe that patient demand is now more equally spread over these two wards.

Optimized Clinical Ward Layout

Figure 28: Bed occupancy and blocked patients for current clinical ward layout

Day treatment departments

Figure 29 presents the current day ward layout for the three fixed clusters: the non-surgical cluster is assigned to ward C1, the surgical cluster to D1, and oncology to D0. Notably, the number of available beds significantly exceeds the required number of beds.

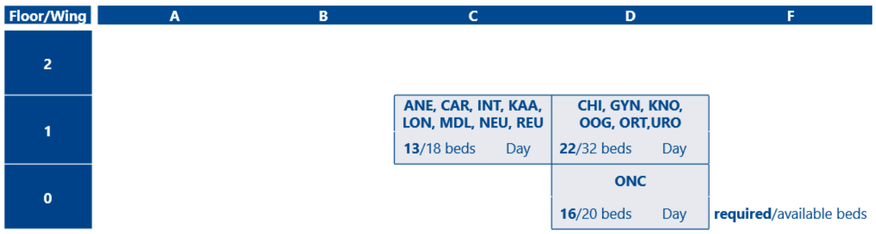


Figure 29: Bed occupancy and blocked patients with the current day ward layout

Figure 30 illustrates the day ward layout for the full compatibility of specialties. The composition of specialties has changed into two quite equally distributed clusters. The required number of beds increases with 4 on ward C1 and decreases with 4 on ward D1, resulting in an equal number of beds needed in total.

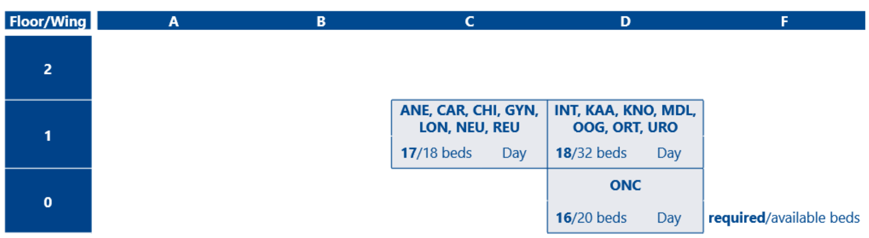


Figure 30: Bed occupancy and blocked patients with the optimized day ward layout

Table 22 presents the blocking probability, number of blocked patients, number of admitted patients, and occupancy rates for the two scenarios. The blocking probability is based on the busiest hour of the week to ensure enough beds on peak moments of the week. This is similar to how the bed requirements model compute it for day treatment departments. For comparison, the number of available beds is set equal to the required beds, meaning any additional physical beds in the wards are not utilized. Once again, we compare only the wards that differ between the two layouts, in this case C1 and D1. While the total number of beds used remains the same in both the current and optimized scenarios, we observe a decrease in the blocking probability on ward C1. Despite a slight increase in the blocking probability on ward D1, the average blocking probability for the two wards decreases from 5.15% to 4.05%, indicating a modest improvement.

Layout	Ward	P_{block}	# Blocked	# Admitted	Occupancy %
Current	C1	6.3%	72	4039	69.0%
	D1	4.0%	24	5290	62.7%
Optimized	C1	3.1%	17	4164	62.1%
	D1	5.0%	51	5193	66.7%

Table 22: Comparison of current and optimized day ward layout

For further analysis of the results, Figure 31 illustrates the bed occupancy trends throughout the year for the current day ward layout. The shaded blue area indicates the daily variability in bed occupancy. The green line denotes the total number of available beds and the pink line represents the number of patients unable to be admitted due to full capacity. We left out the mean occupancy line because of the empty hours that bring down the mean. Figure 32 presents similar bed occupancy plots for the two wards, but now under the optimized day ward layout.

Current Day Ward Layout

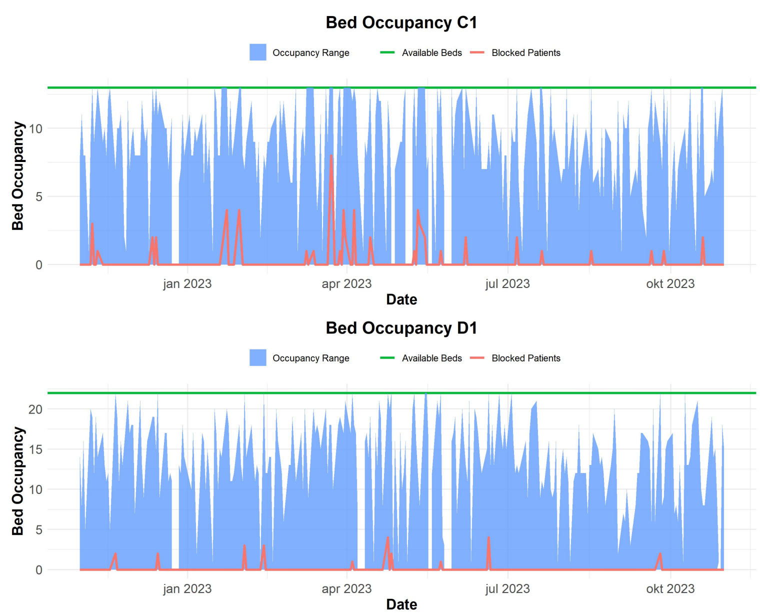


Figure 31: Bed occupancy and blocked patients for current day ward layout

Optimized Day Ward Layout

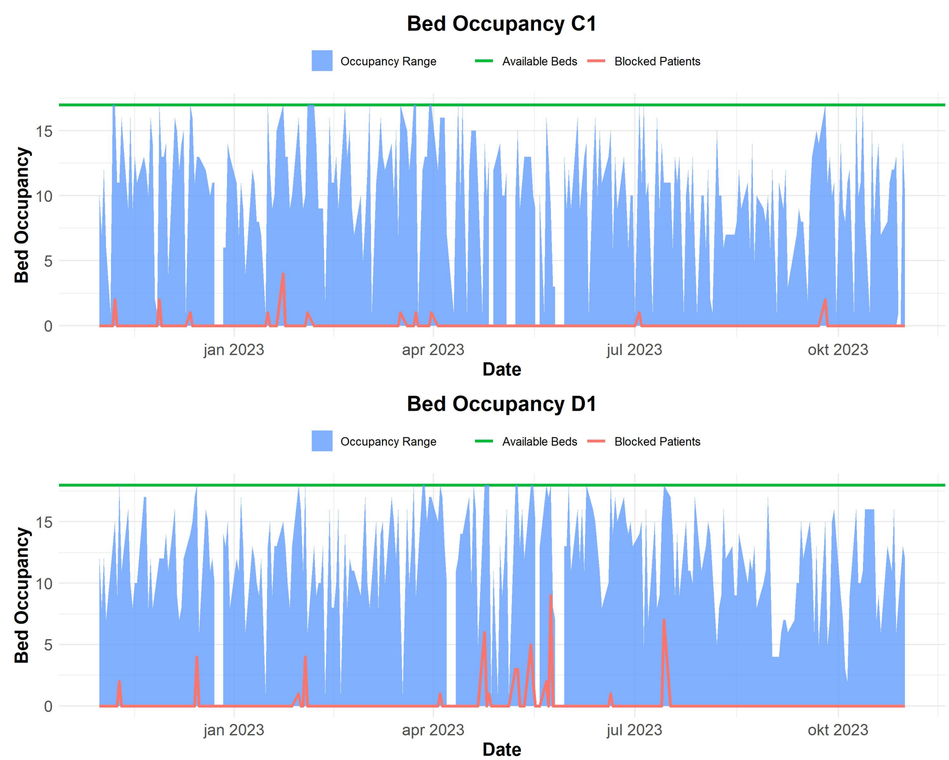


Figure 32: Bed occupancy and blocked patients for current day ward layout

The plots indicate that the number of blocked patients is lower in the optimized layout. However, the total number of blocked patients remains relatively low for both layouts. This is primarily because the bed requirements are calculated based on the busiest hour of the week, which represents peak demand. On an operational level, certain days or during specific time slots, beds may be closed for usage because they are not actually required at those times, even though they were accounted for in the peak demand calculation.

8.3 Case Hospital's Assessment of the Findings

We discussed the application of the model and the results with the hospital. They recognized the high blocking probability of ward A1 in the current clinical ward layout, noting that they have been seeking ways to address the congestion in this ward. We proposed the optimized layout to the hospital, where now orthopedics and general surgery share a ward. They pointed out that, traditionally, orthopedics and general surgery are not placed together in the same ward because infected wounds of general surgery patients could pose health risks to patients recovering from orthopedic procedures, such as prosthesis placements. However, other hospitals have also recently combined orthopedics and general surgery within the same ward and we discussed the possibility of placing at-risk patients in separate rooms within the same ward. The case hospital mentioned that they are definitely open to discuss this option since the results demonstrate a promising reduction in blocking probability.

Regarding day treatment, the hospital observed that the prediction of required beds showed an underestimation. This discrepancy arises because the model does not differentiate between bed types. For instance, the eye surgery specialty uses special chairs instead of regular beds. However, the model currently treats these chairs as regular beds, allowing patients from any specialty to occupy them. This oversight leads to the assumption that any patient can occupy any available bed, disregarding the specific requirements of certain specialties and thus leading to an underestimation of the required beds. To improve the model and provide a more realistic ward layout for day treatments, it is crucial to incorporate bed types into the calculations. This is an important consideration for future research.

9 Conclusion

This research aimed to develop a mathematical model for optimizing the allocation of specialties within clinical and day treatment departments while considering bed requirements, medical and patient-related constraints, and spatial limitations.

We developed an Integer Linear Program (ILP) model to assign clusters of specialties to available wards. A crucial modeling decision was to first identify feasible specialty clusters and calculate their corresponding bed requirements before addressing the cluster-to-ward assignment problem. This strategy substantially reduced the number of decision variables, allowing the problem to be solved efficiently without the need for heuristics. To facilitate this process, we incorporated user input, which included a compatibility matrix specifying which specialties can be placed together. For each feasible cluster, the bed requirements were then computed.

This study employed two distinct models to determine the required number of beds: the Erlang loss model and Discrete Event Simulation (DES). The Erlang loss model is widely used in the literature for modeling inpatient capacity; however, it is not applicable to day treatment settings due to their pronounced time-dependent characteristics. Given the absence of established models for day treatment in previous literature, a time-dependent Erlang loss model was introduced to capture the time-dependent nature of day treatment departments.

While DES is significantly slower computationally, it provides a much more detailed and intuitive representation of patient flow. Unlike a mathematical formula like Erlang, DES allows for tracking the number of simulated patients, rejections due to bed shortages, and real-time bed occupancy. Explainability is a crucial factor in model adoption within hospitals, as decision-makers must understand and trust the model's predictions. However, when the number of compatible specialties grows, the number of possible specialty combinations increases exponentially, making DES computationally infeasible.

Our results demonstrated that both models yielded similar outcomes, with only a difference of only a few beds. In both cases, we observed that the combination of specialties resulted in a reduction in the number of beds needed, attributed to the advantages of bed pooling. By pooling beds across a larger patient population, variability in patient arrivals and lengths of stay is mitigated, leading to more efficient bed utilization and a reduction in the total number of beds required to maintain the same blocking probability.

Given the importance of computational efficiency, the (time-dependent) Erlang loss model is preferred when computational time is a constraint. However, for clinical departments—where the number of compatible specialty combinations is often manageable—DES remains the preferred method due to its greater transparency and interpretability for hospitals. For implementation, this trade-off must be made based on the wishes of hospitals.

In addition to bed requirements constraints, we incorporated user input to define preferred ward assignments for specific specialties and proximity constraints for surgical specialties that require wards near the operating room. Furthermore, we considered the physical locations of the wards to determine which wards could belong to the same cluster, thereby prohibiting long walking distances for patients and staff.

To identify an optimal ward layout, the objective function was formulated to minimize the required number of beds while also controlling for the maximum cluster size to prevent the formation of excessively large specialty groups. The ILP succeeds in finding optimal solutions based on this objective function within a few seconds without having to rely on heuristics.

The case study demonstrated that the model identified an alternative layout compared to the current arrangement. By mixing the specialties of two wards, the blocking probability at one clinical ward was reduced from 10.3% to 3.3%, while maintaining the blocking probability of the other wards below the prespecified threshold of 5%. Additionally, it ensured that patient demand was more evenly distributed across the wards, leading to more balanced bed utilization rates. For the day treatment wards, we assumed full compatibility between specialties, except for oncology, and were able to find a new ward layout with lower blocking probabilities while using the same number of beds. However, the ignorance of bed types led to an underestimation of the required number of beds.

Hospitals stand to benefit significantly from this model, as it enables the identification of optimized ward layouts that enhance resource utilization. The process of manually determining the most efficient specialty-to-ward assignments is highly complex and infeasible due to the vast number of possible scenarios. A mathematical model, in contrast, can systematically compare numerous scenarios, ensuring more informed and effective decision-making.

This model focuses specifically on strategic capacity management, but strategic capacity optimization positively influences integral capacity management across the entire hospital, as effective planning lays the foundation for improved capacity management at both tactical and operational levels. The model demonstrated that even minor adjustments in specialty assignments can lead to a substantial reduction in blocking probability and a more consistent utilization rate. This, in turn, reduces pressure on nursing staff. Fewer required beds often translates to a reduced need for nursing personnel, while more balanced utilization rates distribute workloads more evenly, further mitigating stress on nurses. This optimization not only improves operational efficiency but also promotes a more sustainable work environment.

10 Discussion

This section outlines the limitations of the study and suggests directions for future research. A key limitation is that the model does not account for overlap between day treatment and clinical departments. In practice, some hospitals may integrate these units, but this possibility is not incorporated into the current framework. This could limit the direct applicability of the proposed ward layout in hospitals where such integration occurs. Additionally, the exclusion of certain wards presents another constraint, as the proposed ward layout does not include all inpatient departments.

Furthermore, the model does not explicitly account for different types of beds, such as treatment chairs or specialized equipment designed for specific patient groups. This simplification may affect the accuracy of the results, as some specialties depend heavily on alternative bed types. For example, oncology day treatment units often use treatment chairs for chemotherapy patients, which differ significantly from standard inpatient beds in terms of space requirements.

Additionally, when a user specifies very few restrictions in the compatibility matrix, the solution space can become excessively large, leading to extreme runtimes when using DES for bed estimations. While DES offers a more detailed representation of patient flow, this finding implies that hospitals seeking highly flexible ward assignments may be limited to using the Erlang loss model to ensure timely decision-making. However, this issue could be mitigated with more efficient coding practices and the use of faster computers, which would reduce the computational burden.

Moreover, if the required number of beds exceeds the available capacity, the model is unable to find a feasible solution for the specified blocking probability, as this is currently enforced as a hard constraint. This means that now the user has to increase the specified blocking probability until the required number of beds decreased enough for the model to find a feasible solution. However, since this requires the entire model to run again this is not user friendly.

Lastly, there is no available data on patients who were blocked and subsequently transferred to another hospital. As a result, if the number of such patients is significant, the arrival rate may be underestimated, potentially leading to an underestimation of the required number of beds.

For future research, several avenues could be explored to address the limitations discussed above. One potential direction is to develop a more flexible model that accounts for the integration of day treatment and clinical departments, accommodating the diverse ways in which hospitals organize their units. Additionally, future studies could focus on improving the accuracy of bed estimations by explicitly considering different types of beds. Both these directions ensure better aligning with real-world hospital practices. Furthermore, we could replace the hard constraint for the required number of beds in the cluster-to-ward assignment model with a soft constraint. This soft constraint would impose a penalty on clinical ward layouts where the bed capacity is insufficient, thus preventing the model from failing to find a solution while still discouraging infeasible layouts.

Another important aspect that could be incorporated into the model is the consideration of nursing staff requirements. The case study hospital has already indicated that reducing the required number of beds by, for example, only one bed may not lead to significant improvements if the number of nurses cannot be adjusted accordingly. Future research could explore how the model could account for nursing staffing levels. This would provide a more comprehensive understanding of how changes in bed allocation affect overall hospital resource planning, particularly in terms of staffing needs.

References

Arnolds, I. V., & Nickel, S. (2013). Multi-period layout planning for hospital wards. *Socio-Economic Planning Sciences*, *47*(3), 220–237. https://doi.org/10.1016/j.seps.2013.02.001

Bekker, R., & de Bruin, A. M. (2010). Time-dependent analysis for refused admissions in clinical wards. *Annals of Operations Research*, *178*, 45–65.

Bekker, R., Koole, G., & Roubos, D. (2017). Flexible bed allocations for hospital wards. *Health Care Management Science*, *20*, 453–466. https://doi.org/10.1007/s10729-016-9364-4

Berge Holm, L., Lurås, H., & Dahl, F. A. (2013). Improving hospital bed utilisation through simulation and optimisation: With application to a 40% increase in patient volume in a Norwegian general hospital. *International Journal of Medical Informatics*, *82*(2), 80–89. https://doi.org/10.1016/j.ijmedinf.2012.05.006

Burdett, R. L., Corry, P., Cook, D., & Birgan, S. (2024). Optimisation approaches for organising hospital infrastructure function. *International Journal of Production Research*, 1–23. https://doi.org/10.1080/00207543.2024.2347563

de Bruin, A. M., Bekker, R., van Zanten, L., et al. (2010). Dimensioning hospital wards using the Erlang loss model. *Annals of Operations Research*, *178*, 23–43. https://doi.org/10.1007/s10479-009-0647-8

Devapriya, P., Strömblad, C. T. B., Bailey, M. D., et al. (2015). Stratbam: A discrete-event simulation model to support strategic hospital bed capacity decisions. *Journal of Medical Systems*, *39*(130). https://doi.org/10.1007/s10916-015-0325-0

Erlang, A. K. (1948). The theory of probabilities and telephone conversations. In E. Brockmeyer, H. L. Halstrøm, & A. Jensen (Eds.), *The life and works of A.K. Erlang* (pp. 131–137). Transactions of the Danish Academy of Technical Sciences, 2.

Green, L. (2002). How many hospital beds? *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, *39*(4), 400–412. https://doi.org/10.5034/inquiryjrnl_39.4.400

Green, L. (2006). Queueing analysis in healthcare. In *Patient flow: Reducing delay in healthcare delivery* (pp. 281–307). Springer.

Green, L., & Nguyen, V. (2001). Strategies for cutting hospital beds: The impact on patient service. *Health Services Research*, *36*, 421–442.

Hans, E., van, M., & Hulshof, P. (2012). A framework for healthcare planning and control. In *Operations Research in Health Care* (pp. 12–20). Springer. https://doi.org/10.1007/978-1-4614-1734-7_12

Hu, X., Barnes, S., & Golden, B. (2018). Applying queueing theory to the study of emergency department operations: A survey and a discussion of comparable simulation studies. *International transactions in operational research*, *25*(1), 7–49.

Hübner, A., Kuhn, H., & Walther, M. (2018). Combining clinical departments and wards in maximum-care hospitals. *OR Spectrum*, *40*, 679–709. https://doi.org/10.1007/s00291-018-0522-6

Izady, N., Arabzadeh, B., Sands, N., & Adams, J. (2024). Reconfiguration of inpatient services to reduce bed pressure in hospitals. *European Journal of Operational Research*, *316*(2), 680–693. https://doi.org/10.1016/j.ejor.2024.02.008

Koole, G. (2021). *Optimization of business processes*. MG Books.

Li, Y., Liao, P., Song, Y., & Chi, H. (2023). A systematic decision-support approach for healthcare facility layout design integrating resource flow and space adjacency optimization with simulation-based performance evaluation. *Journal of Building Engineering*, *77*, 107465. https://doi.org/10.1016/j.jobe.2023.107465

Liu, T. M. (2012). *A generic bed planning model* [Doctoral dissertation, University of Toronto (Canada)].

Massey, W. A., & Whitt, W. (1994). An analysis of the modified offered load approximation for the Erlang loss model. *Annals of Applied Probability*, *4*, 1145–1160.

Mitchell, P. H. (2008). Defining patient safety and quality care. *Patient safety and quality: An evidence-based handbook for nurses*.

Moengin, P., Septiani, W., & Herviana, S. (2014). A discrete-event simulation methodology to optimize the number of beds in hospital. *Proceedings of the World Congress on Engineering and Computer Science*, *2*.

Ravaghi, H., Alidoost, S., Mannion, R., et al. (2020). Models and methods for determining the optimal number of beds in hospitals and regions: A systematic scoping review. *BMC Health Services Research*, *20*(1), 186. https://doi.org/10.1186/s12913-020-5023-z

Stylianou, N., Fackrell, R., & Vasilakis, C. (2017). Are medical outliers associated with worse patient outcomes? A retrospective study within a regional NHS hospital using routine data. *BMJ Open*, *7*(5), e015676. https://doi.org/10.1136/bmjopen-2016-015676

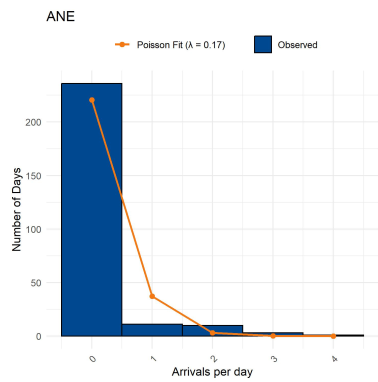
van Essen, J. T., van Houdenhoven, M., & Hurink, J. L. (2015). Clustering clinical departments for wards to achieve a prespecified blocking probability. *OR Spectrum*, *37*, 243–271. https://doi.org/10.1007/s00291-014-0368-5

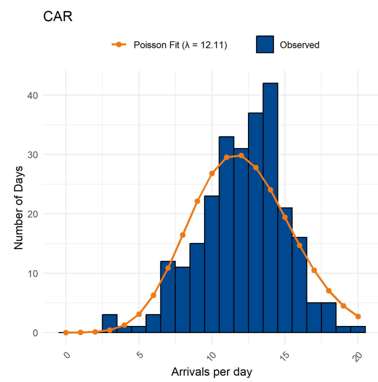
Veneklaas, W., Leeftink, A. G., van Boekel, P. H. C. M., & Hans, E. W. (2021). On the design, implementation, and feasibility of hospital admission services: The admission lounge case. *Omega*, *100*, 102308. https://doi.org/10.1016/j.omega.2020.102308

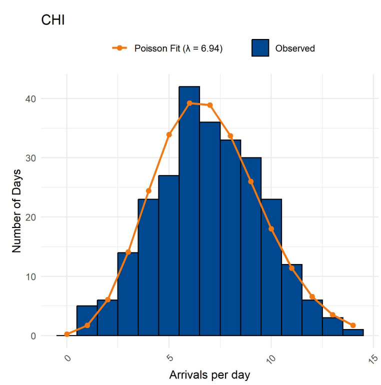
Walther, M. (2020). *Bed planning: Advanced applications of operational research in large hospitals:[cumulative dissertation]* [Doctoral dissertation, Katholische Universität Eichstätt-Ingolstadt].

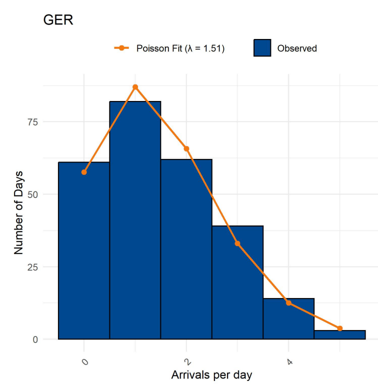
A Histogram of clinical arrivals per specialty

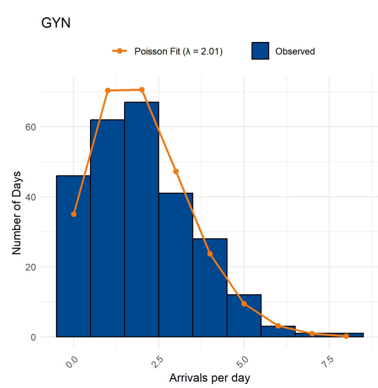
We plot for each specialty a histogram with the daily number of arrivals on the x-axis and the frequency on the y-axis. The orange line is the Poisson fit.

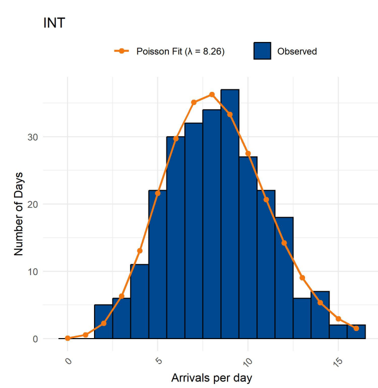


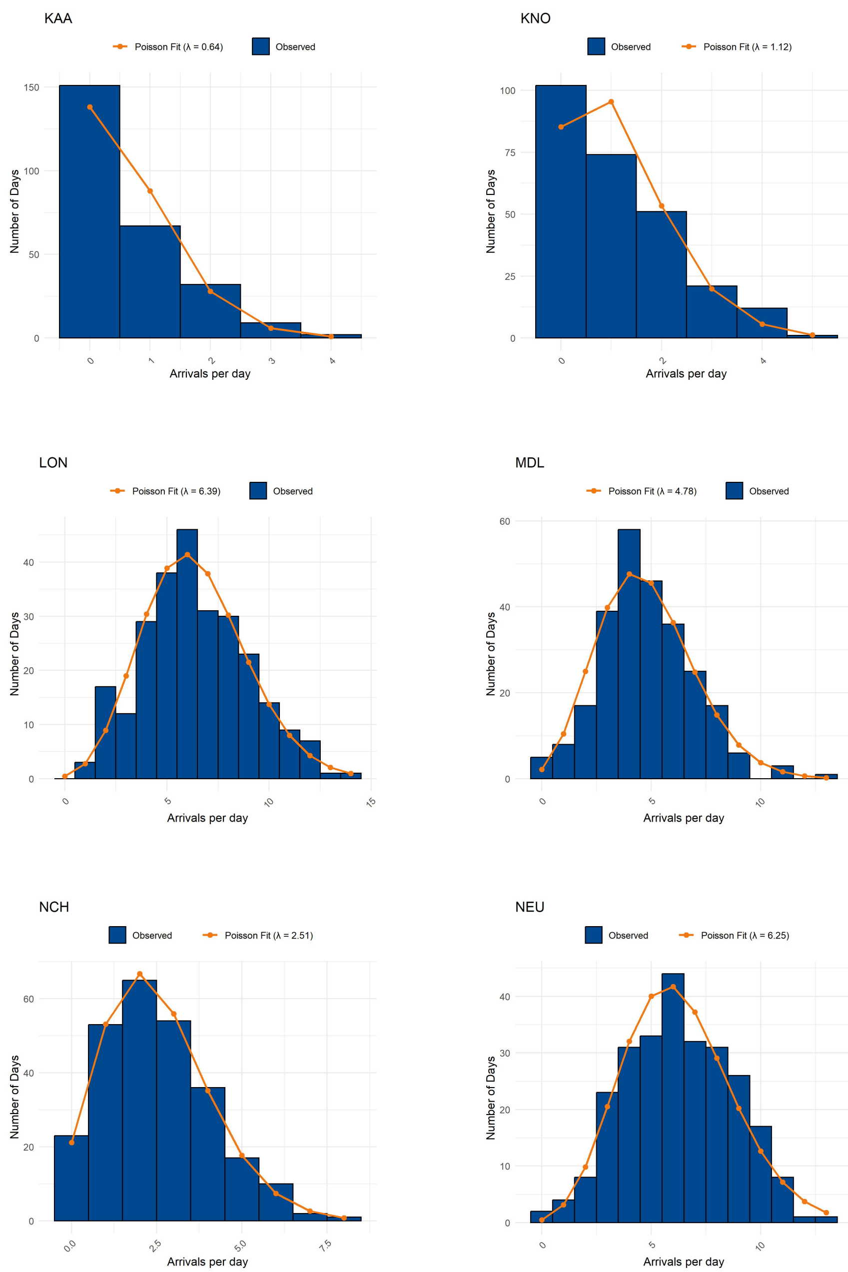


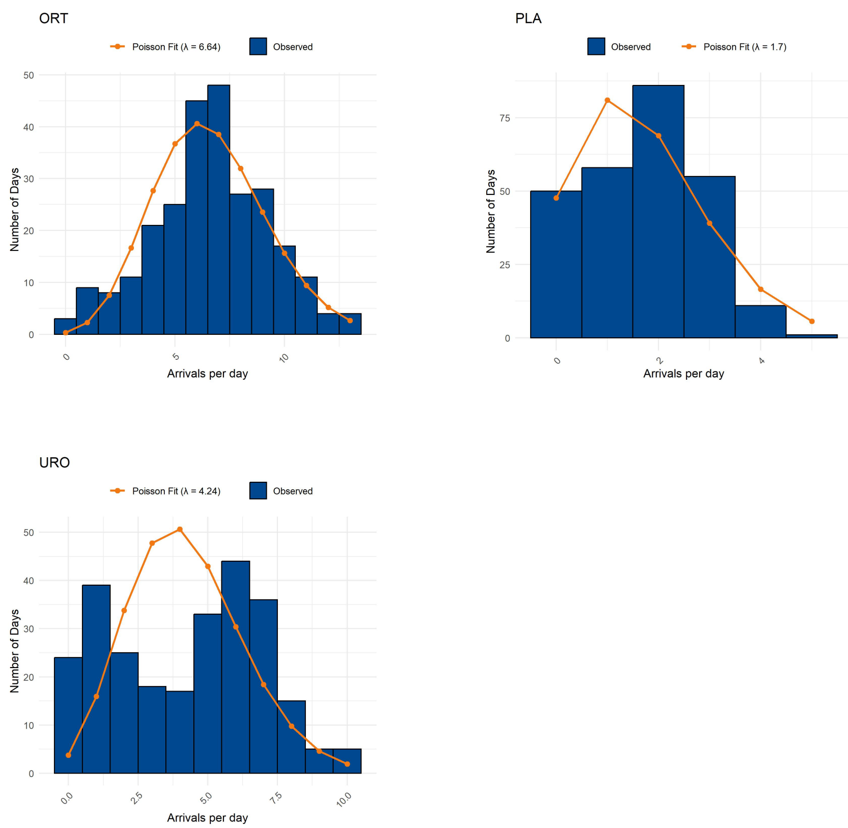




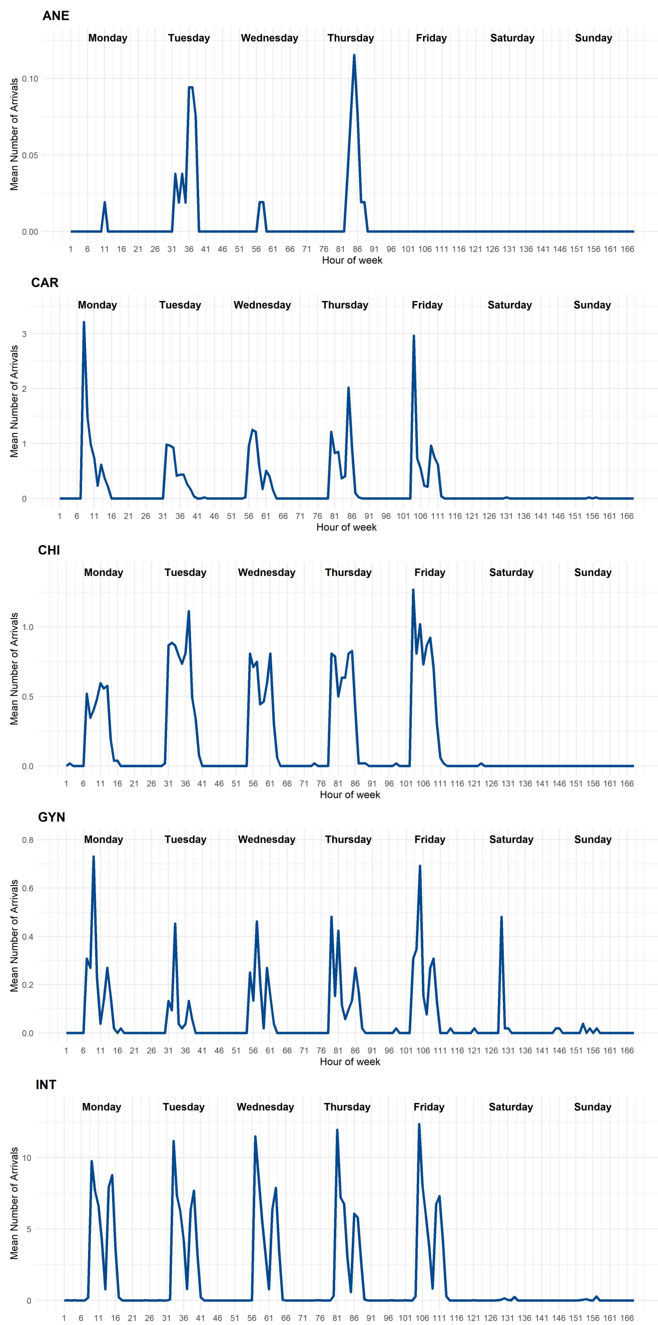


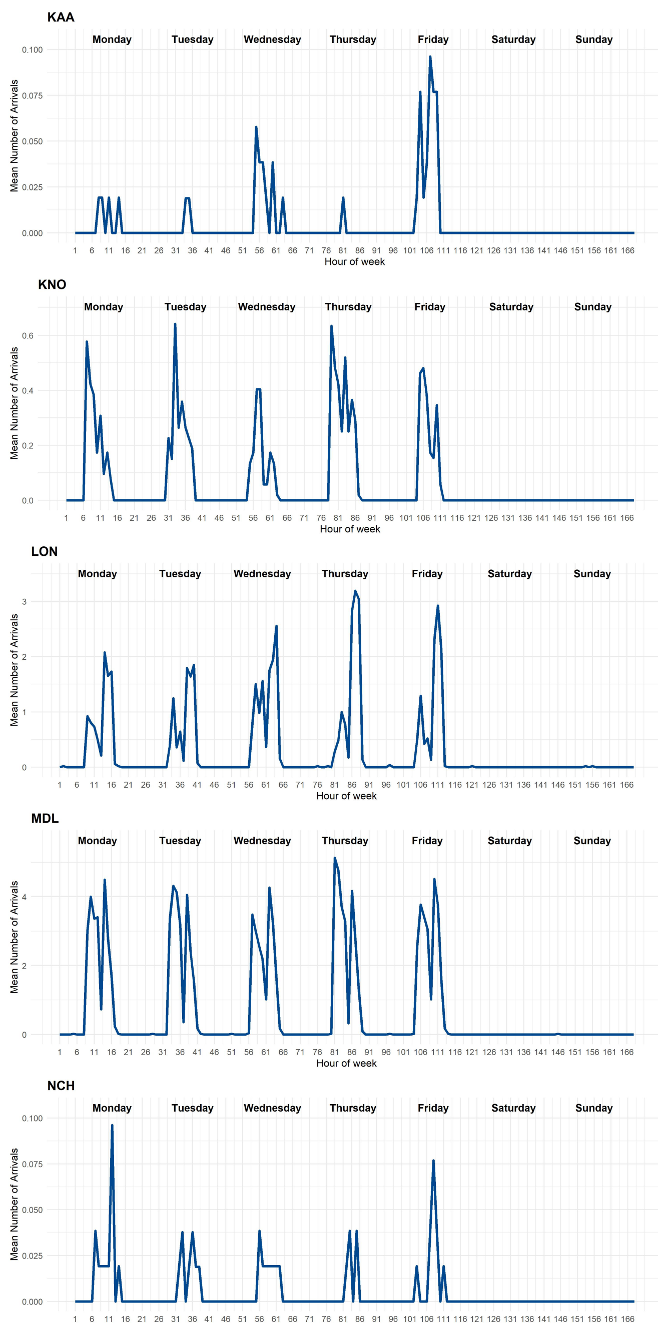


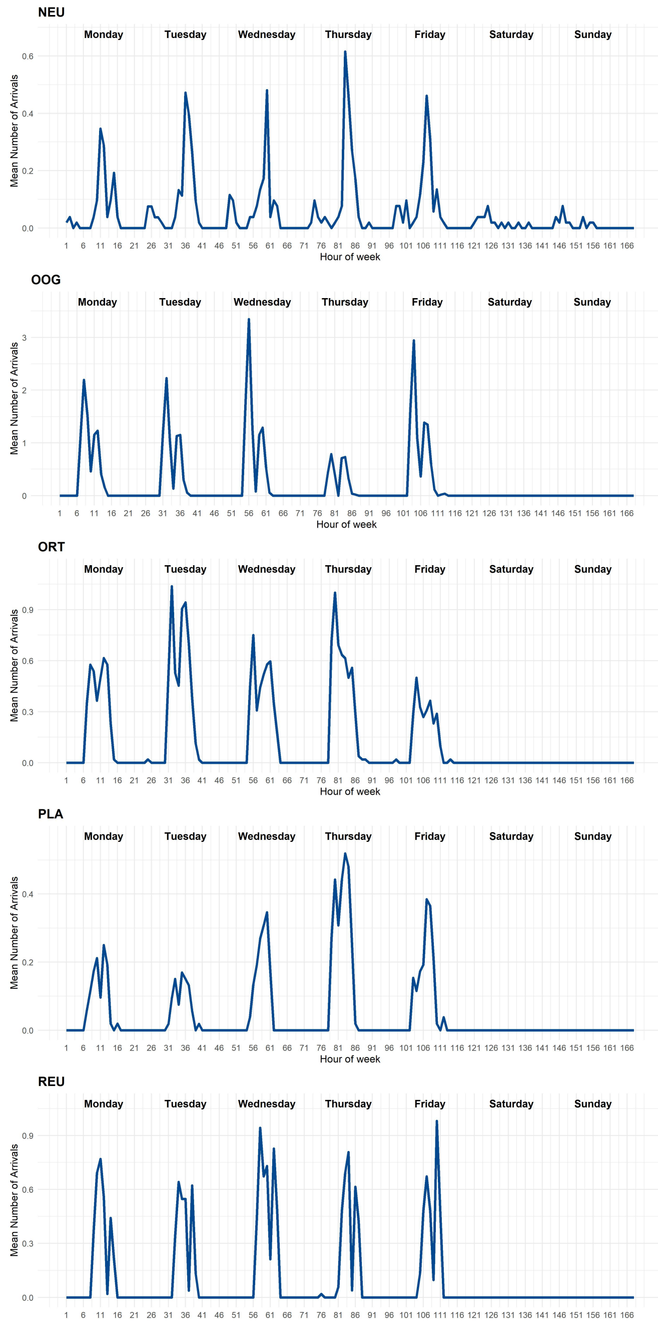

KAA
Poisson Fit (λ = 0.64)
Observed
Number of Days
Arrivals per day
KNO
Observed
Poisson Fit (λ = 1.12)
Number of Days
Arrivals per day
LON
Poisson Fit (λ = 6.39)
Observed
Number of Days
Arrivals per day
MDL
Poisson Fit (λ = 4.78)
Observed
Number of Days
Arrivals per day
NCH
Observed
Poisson Fit (λ = 2.51)
Number of Days
Arrivals per day
NEU
Observed
Poisson Fit (λ = 6.25)
Number of Days
Arrivals per day

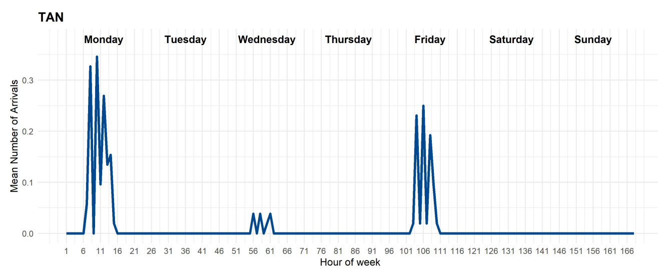

ORT
Poisson Fit (λ = 6.64)
Observed
Number of Days
0
10
20
30
40
50
0
5
10
Arrivals per day
PLA
Observed
Poisson Fit (λ = 1.7)
Number of Days
0
25
50
75
0
2
4
Arrivals per day
URO
Poisson Fit (λ = 4.24)
Observed
Number of Days
0
10
20
30
40
50
0.0
2.5
5.0
7.5
10.0
Arrivals per day

B Weekly distribution of day treatment arrivals per specialty

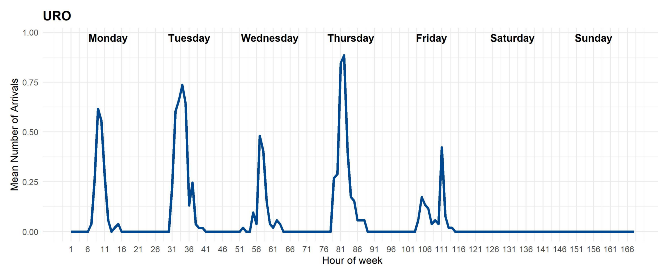



KAA
KNO
LON
MDL
NCH
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Mean Number of Arrivals
Hour of week


NEU
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Mean Number of Arrivals
Hour of week
OOG
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Mean Number of Arrivals
Hour of week
ORT
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Mean Number of Arrivals
Hour of week
PLA
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Mean Number of Arrivals
Hour of week
REU
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Mean Number of Arrivals
Hour of week



TAN
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Mean Number of Arrivals
Hour of week

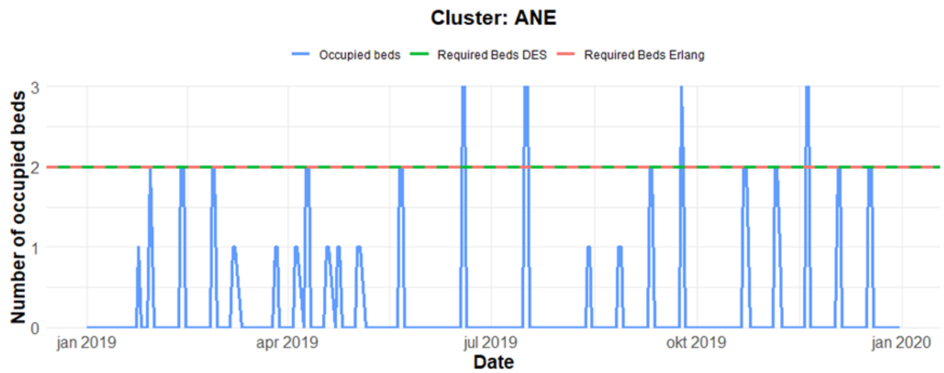


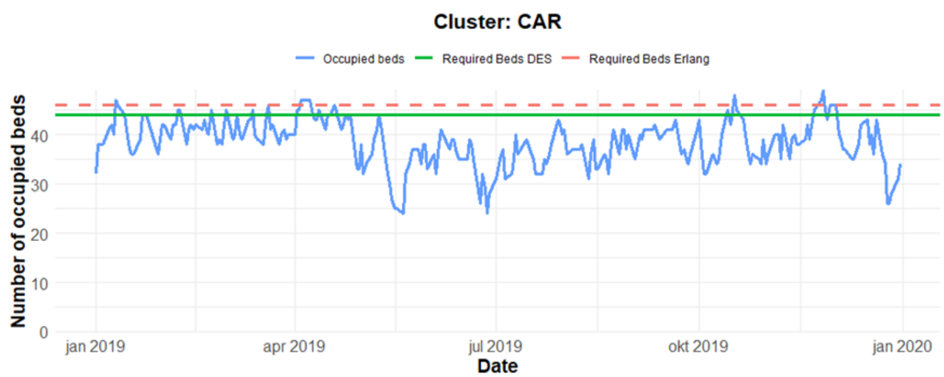
URO
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Mean Number of Arrivals
Hour of week

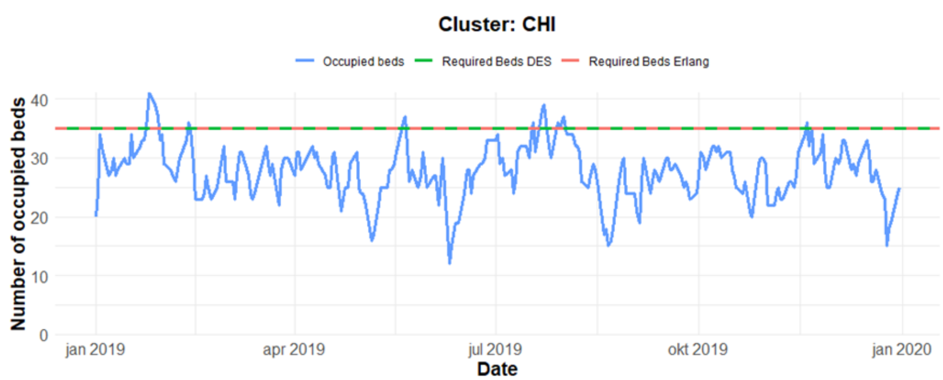
C ILP Cluster-to-Ward Assignment

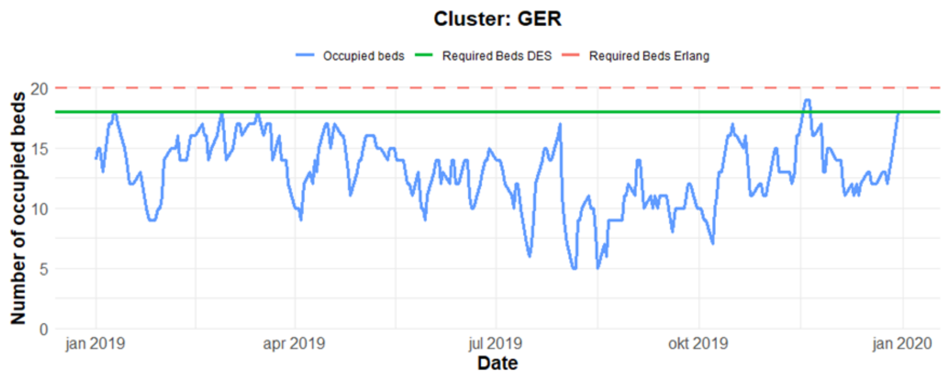
$$
\begin{array}{llll}
\min & \sum_{c \in C} r_c Y_c + C_{\max} & & \\
\text{s.t.} & \sum_{c \in C} h_{s,c} Y_c = 1, & \forall s \in S & \text{(Specialty assigned exactly once)} \\
& \sum_{c \in C} X_{c,w} \leq 1, & \forall w \in W & \text{(At most one cluster per ward)} \\
& \sum_{w \in W} X_{c,w} b_w \geq r_c Y_c, & \forall c \in C & \text{(Sufficient bed capacity per cluster)} \\
& \sum_{w \in W} X_{c,w} \leq |W| Y_c, & \forall c \in C & \text{(Linking } Y_c \text{ and } X_{c,w}\text{)} \\
& Y_c \leq \sum_{w \in W} X_{c,w}, & \forall c \in C & \text{(Linking } Y_c \text{ and } X_{c,w}\text{)} \\
& C_{\max} \geq \sum_{s \in S} h_{s,c} Y_c, & \forall c \in C & \text{(Bound maximum cluster size)} \\
& \sum_{c \in C} h_{s,c} X_{c,w} \geq m_{s,w}, & \forall s \in S, \forall w \in W & \text{(Mandatory specialty assignments)} \\
& d_w t_s h_{s,c} X_{c,w} \leq d_{\max}, & \forall s \in S, w \in W, c \in C & \text{(Distance to OR limit)} \\
& X_{c,w} + X_{c,v} - 1 \leq p_{w,v}, & \forall c \in C, w \in V, v \in W & \text{(Proximity constraint for wards)} \\
& X_{c,w} \in \{0,1\}, & \forall c \in C, w \in W & \\
& Y_c \in \{0,1\}, & \forall c \in C & \\
& C_{\max} \in \mathbb{Z}_{\geq 0} & &
\end{array}
$$

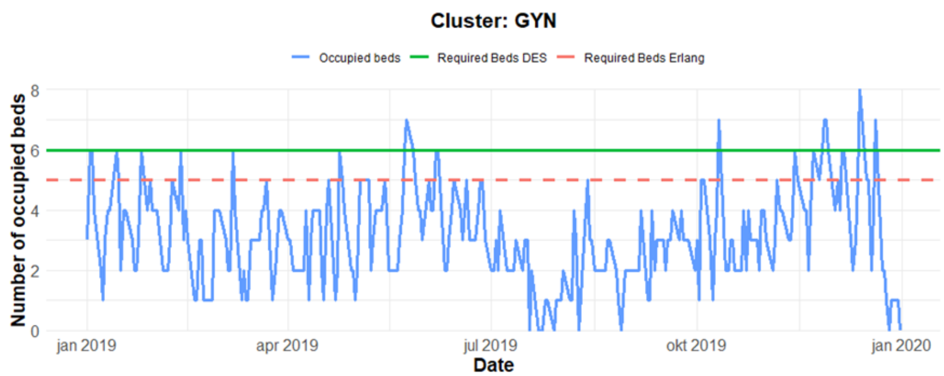
D Clinical Bed Requirements Results

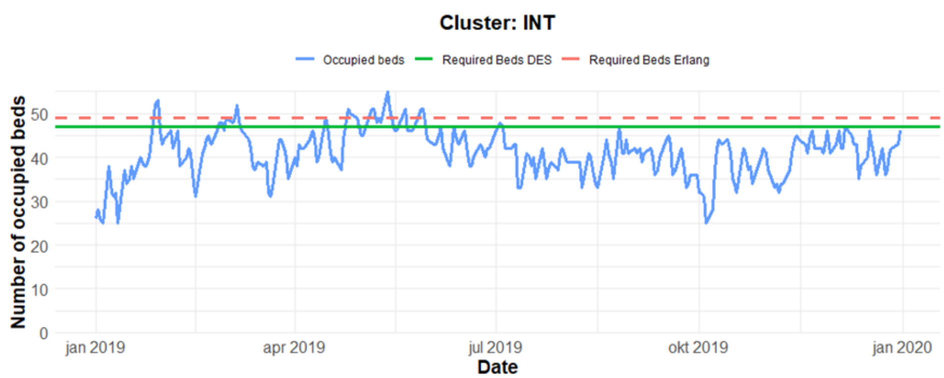


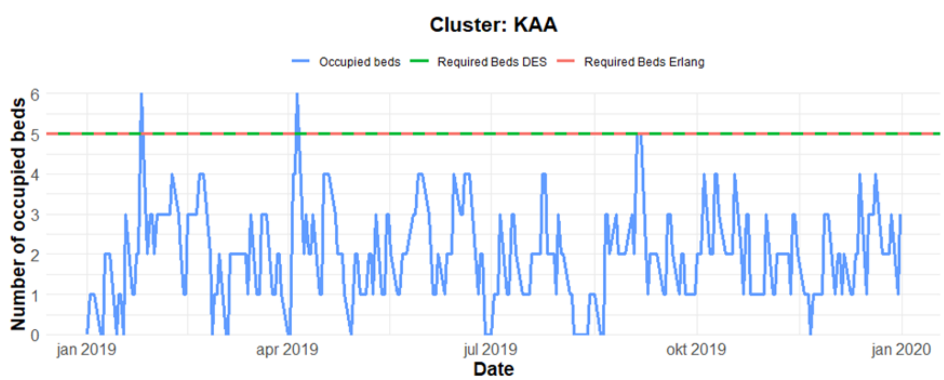


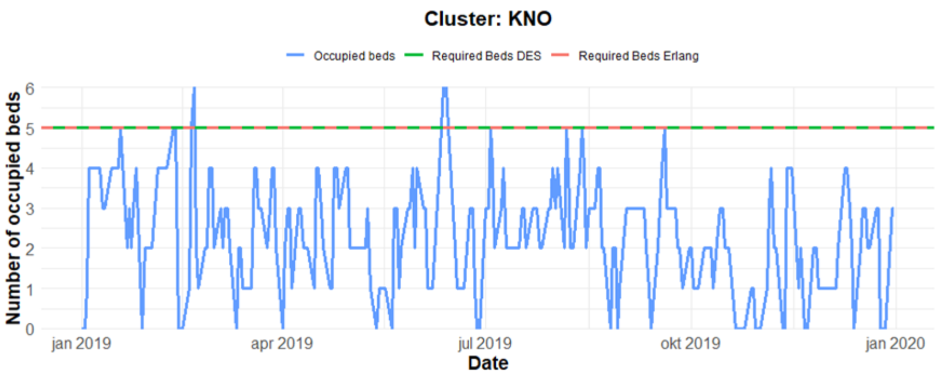


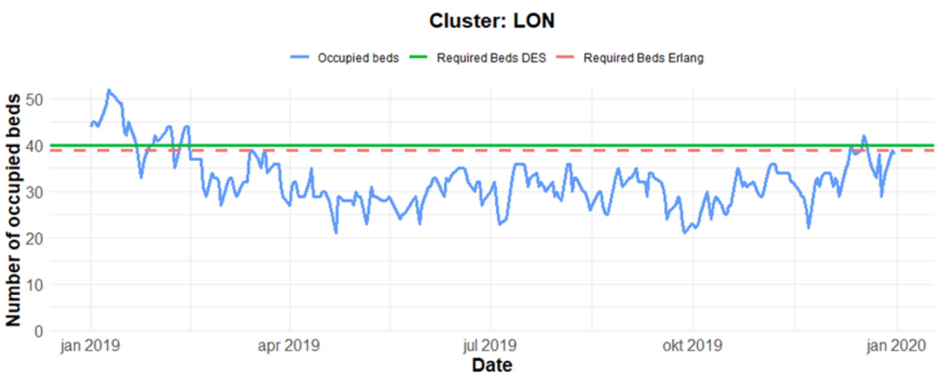

Cluster: GER
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
Date
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020

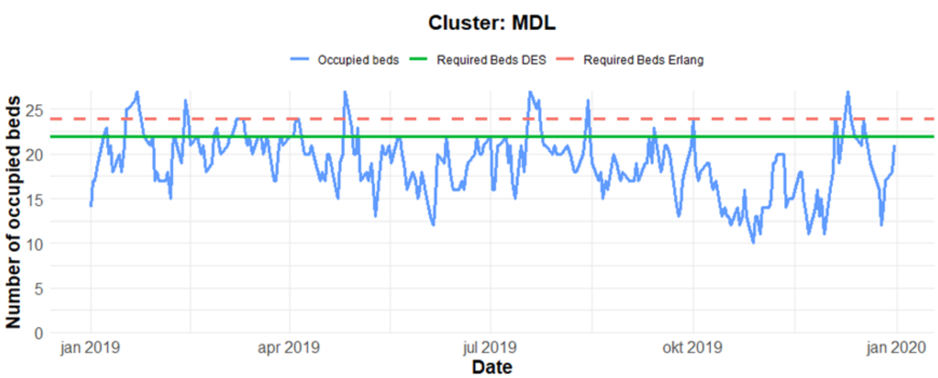

Cluster: GYN
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
Date
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020

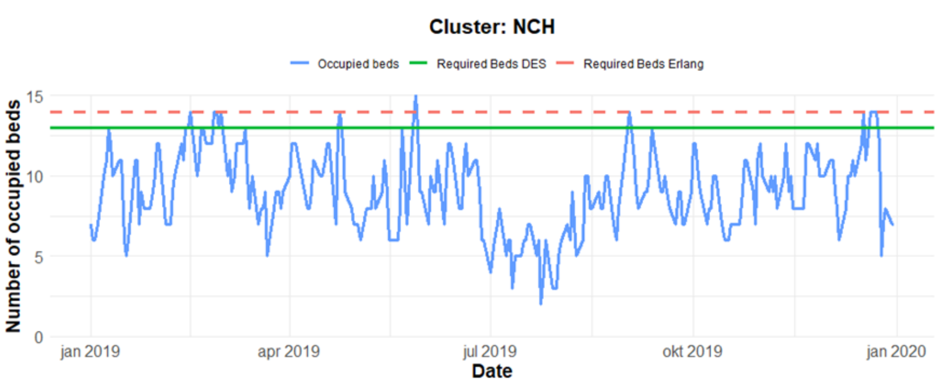

Cluster: INT
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
Date
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020

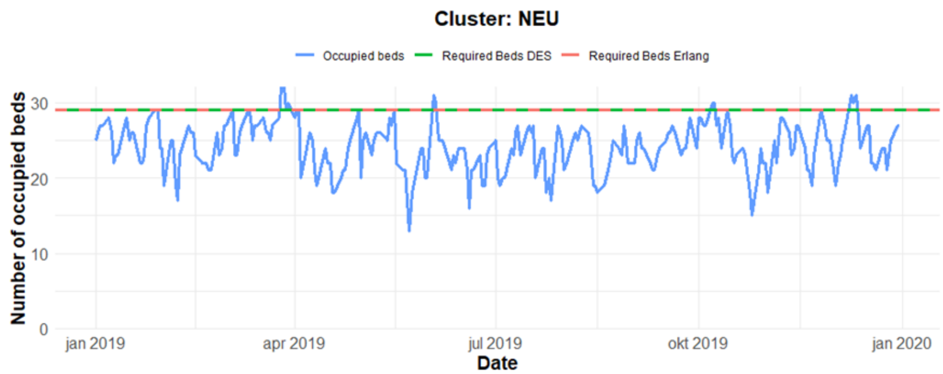

Cluster: KAA
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
Date
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020

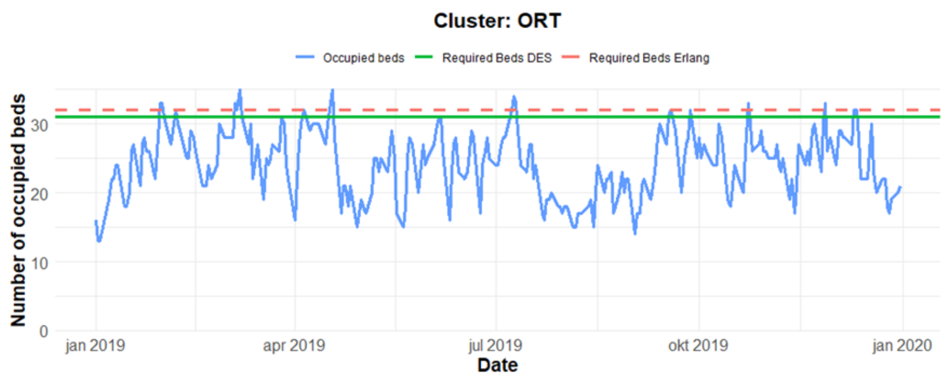

Cluster: KNO
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
1
2
3
4
5
6
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

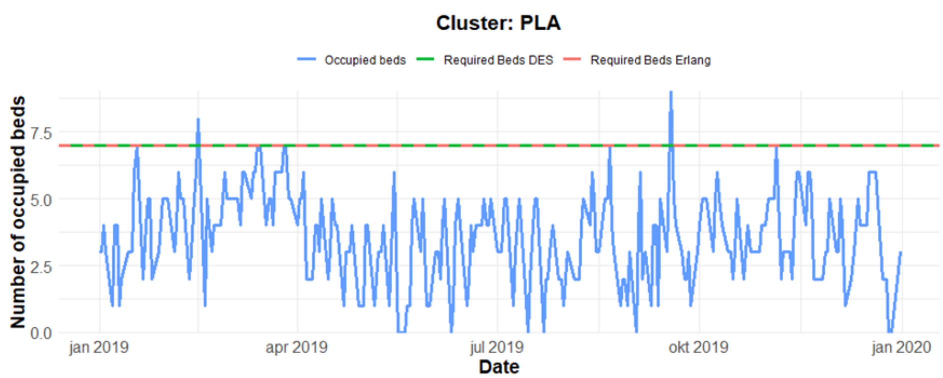

Cluster: LON
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
10
20
30
40
50
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

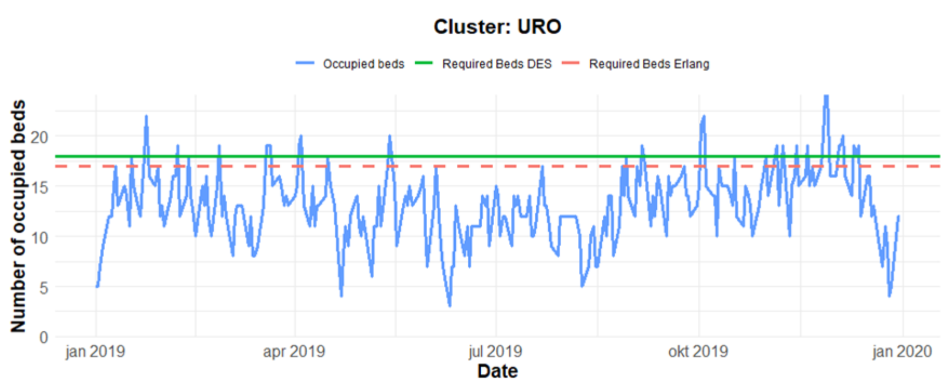

Cluster: MDL
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
5
10
15
20
25
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date


Cluster: NCH
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
5
10
15
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

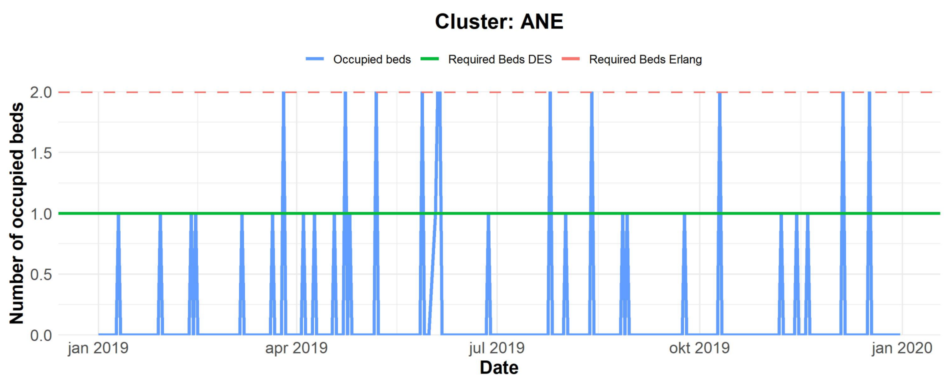

Cluster: NEU
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
10
20
30
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

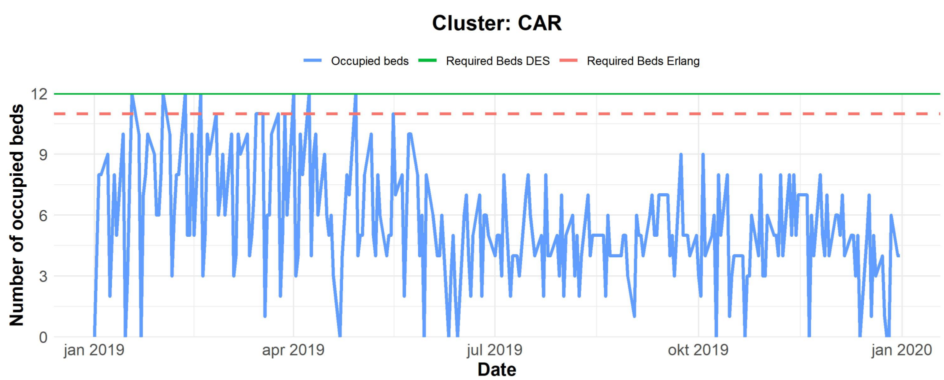

Cluster: ORT
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
10
20
30
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

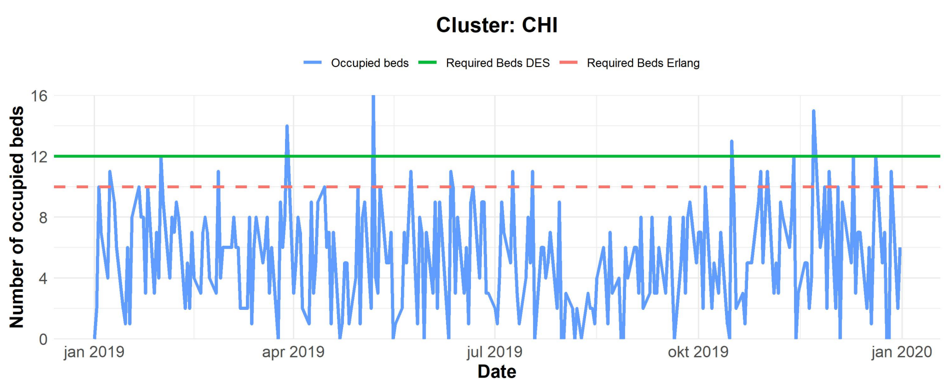

Cluster: PLA
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0.0
2.5
5.0
7.5
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

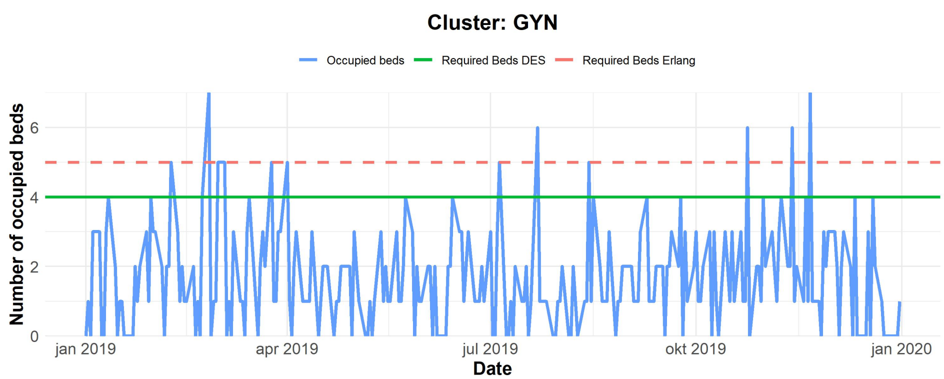

Cluster: URO
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
5
10
15
20
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

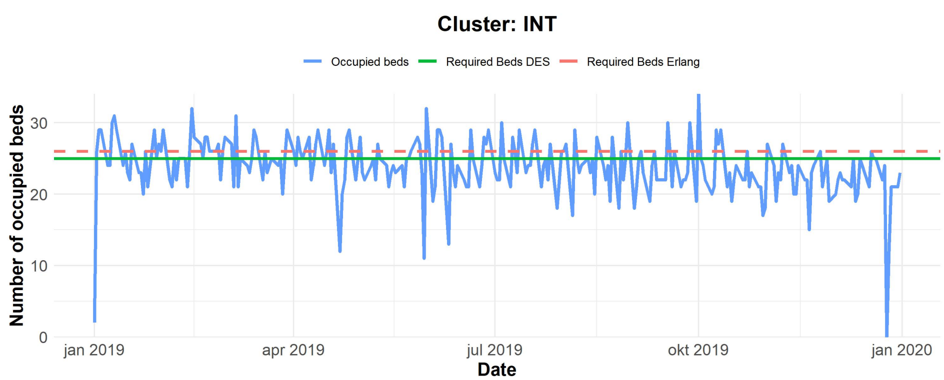
E Day Treatment Bed Requirements Results

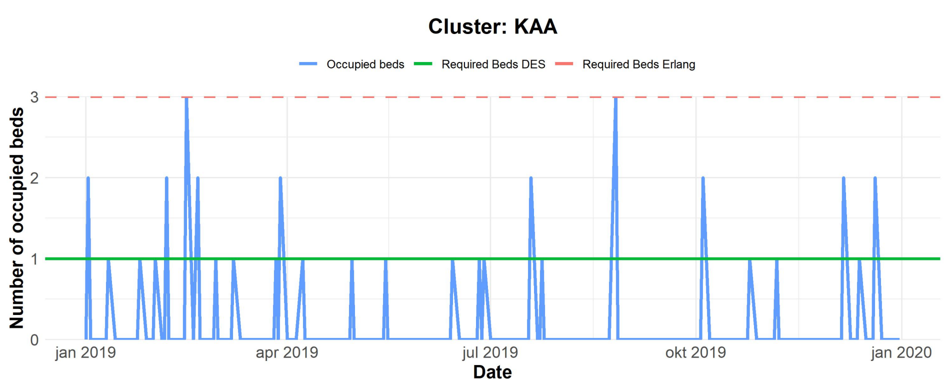


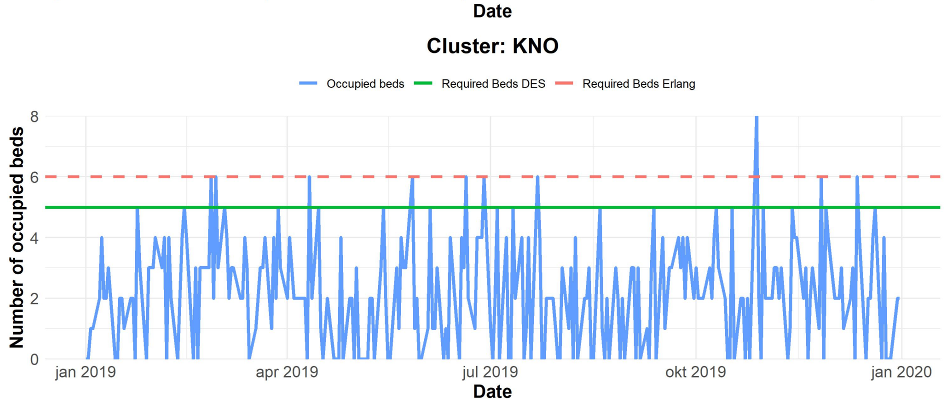


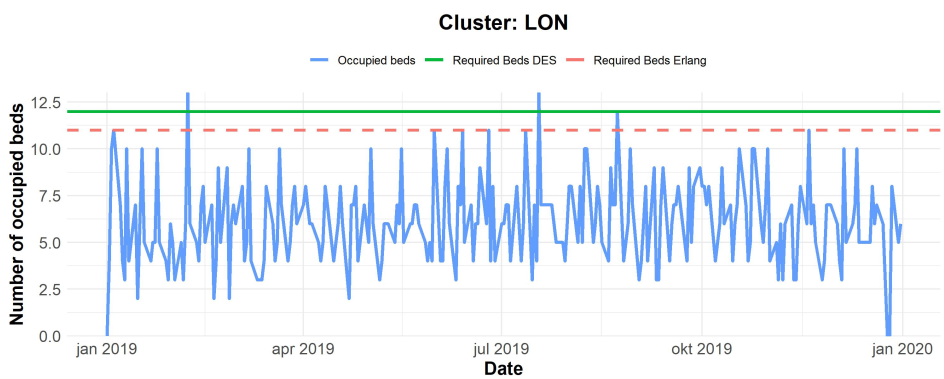


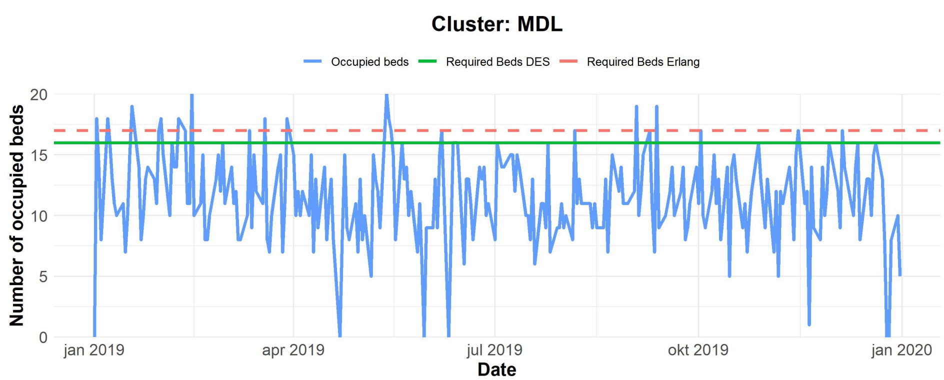

Cluster: GYN
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
2
4
6
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

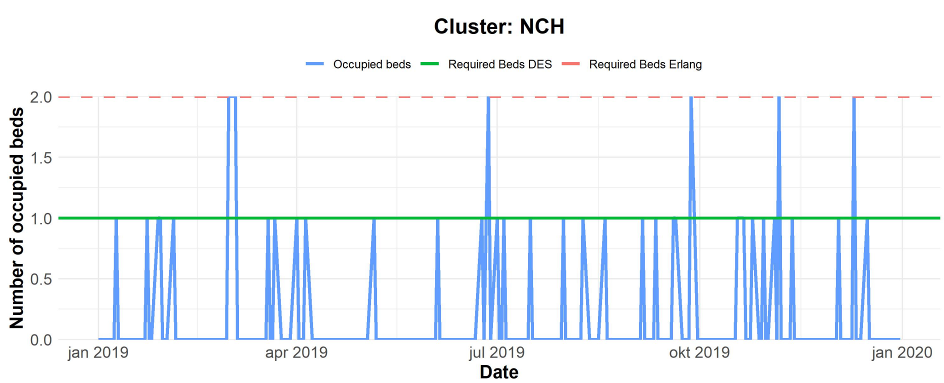

Cluster: INT
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
10
20
30
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

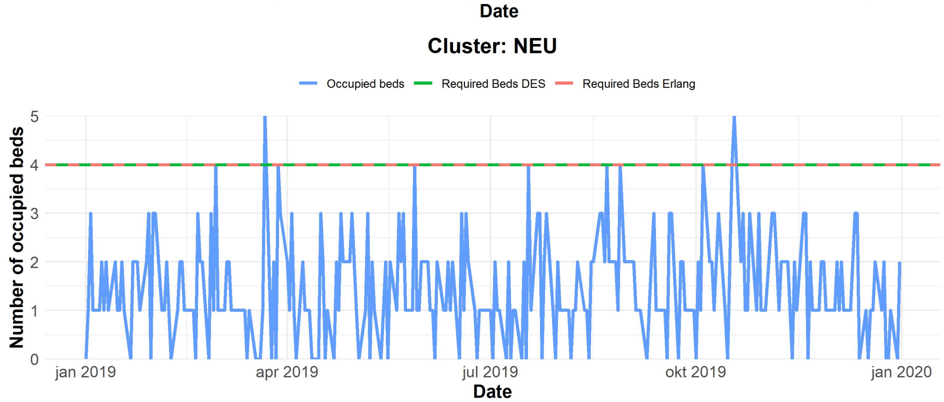

Cluster: KAA
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
1
2
3
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

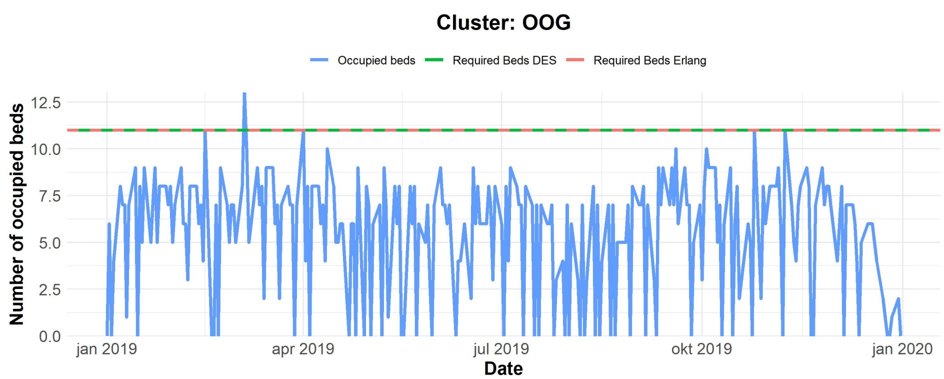

Cluster: KNO
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
2
4
6
8
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

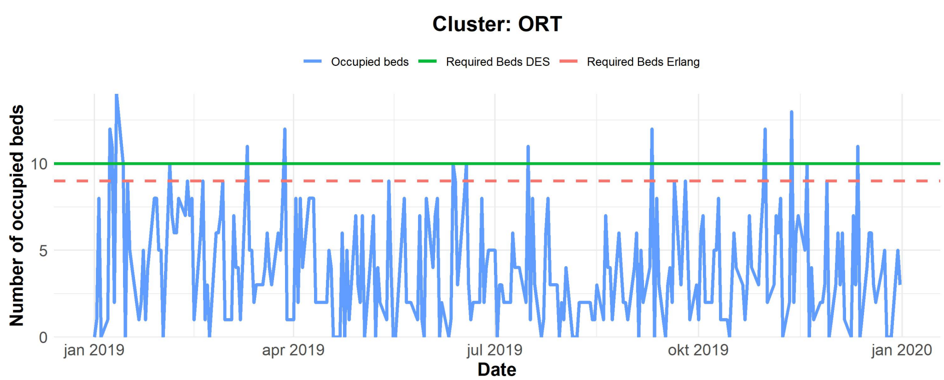

Cluster: LON
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0.0
2.5
5.0
7.5
10.0
12.5
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

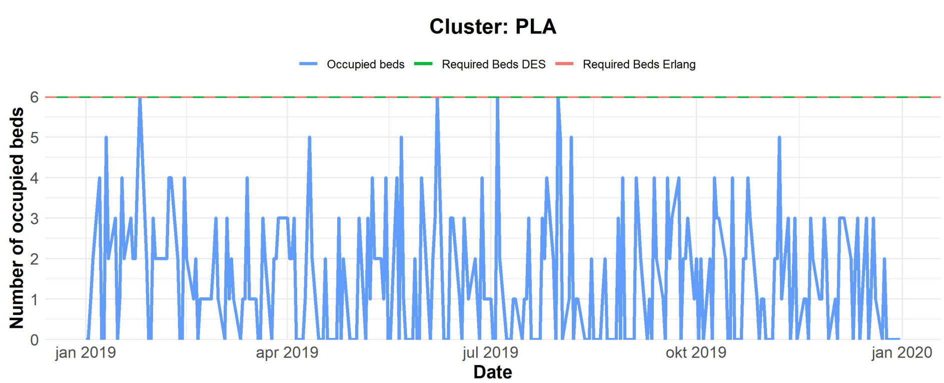

Cluster: MDL
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
5
10
15
20
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

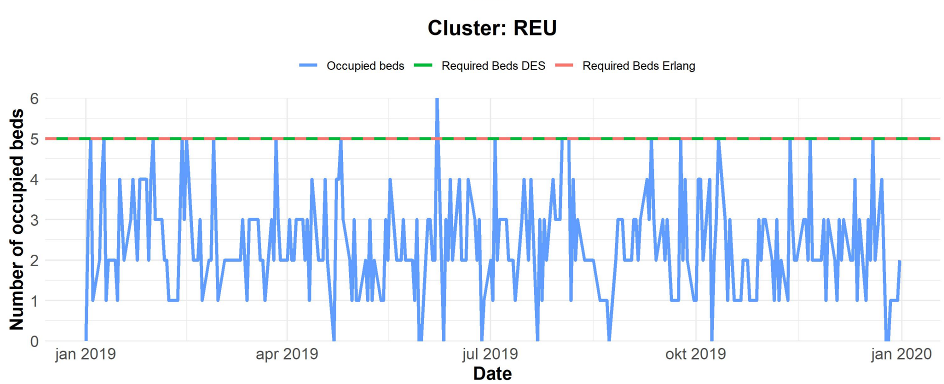

Cluster: NCH
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0.0
0.5
1.0
1.5
2.0
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

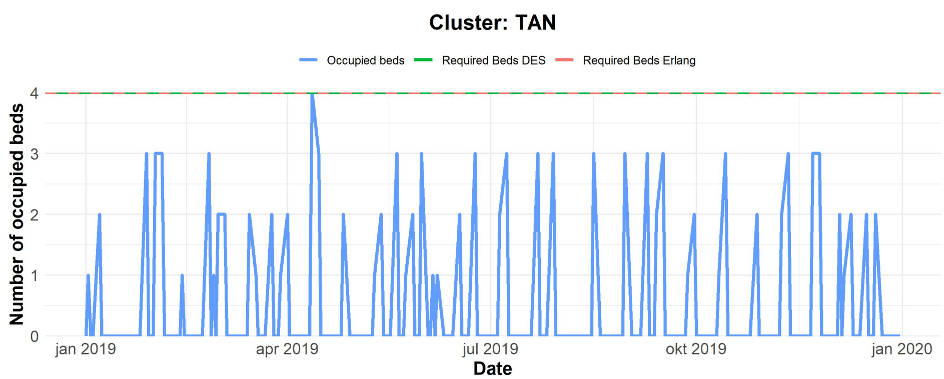

Cluster: NEU
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
1
2
3
4
5
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date


Cluster: OOG
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
12.5
10.0
7.5
5.0
2.5
0.0
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

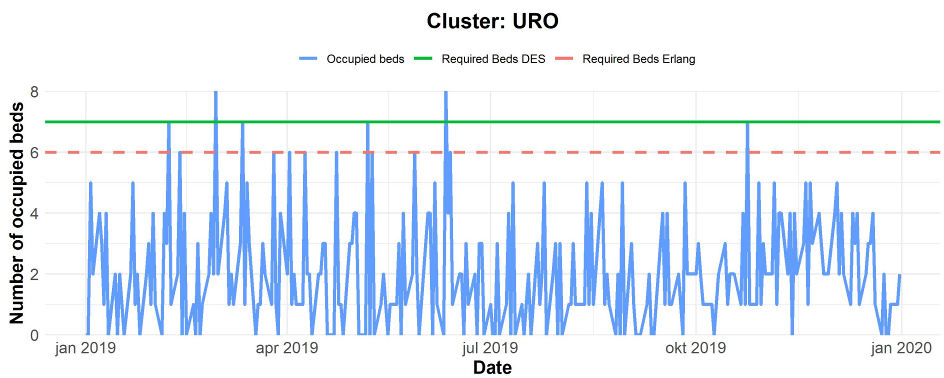

Cluster: ORT
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
10
5
0
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date


Cluster: PLA
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
6
5
4
3
2
1
0
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date


Cluster: REU
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
6
5
4
3
2
1
0
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date



Cluster: TAN
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
1
2
3
4
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date



Cluster: URO
Occupied beds
Required Beds DES
Required Beds Erlang
Number of occupied beds
0
2
4
6
8
jan 2019
apr 2019
jul 2019
okt 2019
jan 2020
Date

F Compatibility Matrix Case Study

	CAR	CHI	GYN	INT	KNO	LON	MDL	NEU	ORT	TRAU	URO
CAR	**1**	0	0	0	0	0	0	0	0	0	0
CHI		**1**	1	0	1	0	1	0	1	1	1
GYN			**1**	0	1	0	1	0	0	0	1
INT				**1**	0	1	1	0	0	0	0
KNO					**1**	0	0	0	1	1	1
LON						**1**	0	0	0	0	0
MDL							**1**	0	0	0	1
NEU								**1**	0	0	0
ORT									**1**	1	0
TRAU										**1**	0
URO											**1**

Table 23: Compatibility matrix for selected clinical specialties