MASTER ASSIGNMENT

Nurse Shift Scheduling Emergency Department St. Olav’s Hospital

12/06/2025

ANNIEK PELLEBOER

S2397919

FACULTY OF BEHAVIORAL, MANAGEMENT & SOCIAL SCIENCES (BMS)

UNIVERSITY OF TWENTE.

ST. OLAVS HOSPITAL
UNIVERSITETSSYKEHUSET I TRONDHEIM

Research information

Master thesis:

Nurse shift scheduling for the emergency department of St. Olav's Hospital

Author:

A.J. Pelleboer

Study program:

Industrial Engineering and Management

University of Twente

Drienerlolaan 5, 7522NB Enschede

Company:

St. Olav's Hospital

Princess Kristinas gate 2-4, Øya, Trondheim

Supervisors

Company supervisors:	Dr. C.H. Andersson
	Dr. T.R. Bovim
	Dr. A.N. Gullhav
University supervisors:	Dr. ir. A.G. Leeftink
	Prof. dr. ir. E.W. Hans

Preface

Dear reader,

Before you lies the report, I have written for my master's thesis Industrial Engineering and Management at the University of Twente. The research has been done at St. Olav's Hospital, a university hospital located in Trondheim, Norway. I am grateful for all help I received and the opportunities that were handed to me.

St. Olav's Hospital handed me the opportunity to perform my research and gain some experience in their culture. I thank everyone who took the time to help me and supported me in finding my way around. In particular, I thank Thomas Bovim, Henrik Andersson, and Anders Gullhav for being my supervisors, for putting in hours to help me and my research to go forward, and for the pleasant collaboration.

I also received help from my home university. I thank Gréanne Leeftink for being my first supervisor. Besides receiving a lot of useful feedback, I found it helpful to discuss challenges I was experiencing and determine the next steps. I would also like to thank Erwin Hans for being my second supervisor and therefore putting in time and effort to give me feedback on my work.

Lastly, I just want to say a big thank you to everyone who helped make it possible for me to do my thesis in Trondheim. It was such a great experience, and I've learned a lot and made some amazing memories along the way.

Have fun reading this thesis!

Anniek Pelleboer

May 2025

Management summary

Shift scheduling plays a crucial role in healthcare operations, balancing patient-care needs against labour costs, staff well-being, and regulatory requirements. Across hospitals, planners must work with variable patient inflows, fixed shift lengths, mandated rest periods, and contract rules, while also respecting nurses' work-life balance and preferences. In Emergency Departments (EDs) in particular, hourly demand can highly differ, making static 8-hour rosters prone to coverage gaps during peaks and unnecessary overstaffing in quieter times.

At St. Olav's University Hospital, scheduling challenges are affected by wage premiums, weekend-work rules, the mix of contract percentages. Currently, the ED works with a shift schedule which does not follow the demand fluctuations very well. This makes it difficult to meet nursing demand with the available nurses and negatively impacts the staff experience. Besides, the ED has been struggling with nurse shortages, which again highlights the importance of efficient use of the available manpower. The solution approach we took in this thesis is optimize the shift schedule for the nurses at ED-nurses, by balancing staffing to cover demand, while minimizing total staffing costs.

From historic patient data we derive hourly nursing-demand percentiles and adopted a 90 % service-level target. A two-stage mixed-integer programming model was then used to experiment with the settings. Different instances for shift lengths, start times, weekend-shift rules, and contract mixes were solved. For each of these instances we compared several measures such as Full Time Equivalents (FTEs) needed, idle time, unused contract hours, and total costs. Sometimes we looked at more case specific measures when we found this of interest for the specific matter. We also compared how the different settings performed on a new dataset to test the robustness of the found shift schedule.

- Allowing mixes of 8 hour and 12-hour shifts combined with seven start times saves 7.31% of the costs, while for eight start times this is 7.65%.
- Paying extra to create extra weekend capacity can be well worth the investment. One extra weekend shift saves 1.32% of the total costs.
- We provided data which allowed the ED to make educated decisions on whether to allow longer shifts only in the weekend or throughout the whole week. Looking at the settings which allows 8- and 12-hour shift with seven start times the cost improvement goes from 7.31% to 6.83% when allowing long shifts only in the weekend. For the settings with eight start times the cost improvement goes from 7.65% to 7.06%.
- Choosing to allow more start times and shift lengths, either only in the weekend or throughout the whole week, results in a needed FTE decrease of 7.6%.
- A lower average contract percentage helps fill weekend gaps, but this effect fades when the settings for the schedule get more flexible. The data shows that with independent of which settings are chosen, the average contract percentage should not be above 90%.

Based on these experiments and their outcomes we made some recommendations:

- Base rostered capacity on the 90% demand percentile. This ensures that service-level targets are met in most hours while avoiding the excessive overstaffing that comes from planning to the absolute peak.
- Complement standard 8-hour shifts with 12-hour shifts for the weekends.
- Increase the number of shifts start options to the eight start times.

- Begin systematically collecting additional patient characteristics and treatment-duration data so it is possible to improve the patient classification system and make the demand modelling more realistic.
- Look into whether nurses are willing to work extra weekends in return for an extra premium.
- Develop the tool further, so its use becomes more beneficial.
- Use an average contract percentage of 90% or lower.

This thesis adds to both the theory and practice of nurse shift scheduling in high-variability clinical settings through three contributions.

1) We introduce a two-stage framework that translates historical arrival data into percentile-based staffing targets and then applies mixed-integer programming to assign shifts. By combining demand forecasting with optimization, our approach extends the shift scheduling models and offers a template adaptable to other situations.
2) We demonstrate that certain combinations of shift lengths, specifically mixes of 8- and 12-hour blocks with staggered start times, outperform traditional fixed-length shift schedules. Experiments show these flexible configurations yield the service-level coverage with lower total nurse-hours than standard 8-hour schedules, underlining the efficiency gains achievable by tailoring shift settings to demand patterns.
3) We deliver a functional prototype scheduling tool, which can be built further on when this research is continued. Currently, it can already be used to compare what different settings do with the efficiency of the shift schedule.

Contents

Tables

Figures

1. Introduction

This chapter introduces St. Olav's Hospital, specifically in the Emergency Department (ED), as they are interested in improving the efficiency of their personnel scheduling process. We start with introducing St. Olav's in Section 1.1. Section 1.2 continues with the introduction of the ED, since this will be the focus of our research. Section 1.3 discusses the experienced problem and 1.4 the current shift scheduling approach. Section 1.5 describes the problem approach. Section 1.6 concludes with the problem scope.

1.1 About St. Olav's

St. Olav's Hospital is a university hospital located in Trondheim, Norway. With its 11,000 employees it is one of Norway's largest healthcare institutions. Their goal is to offer excellent treatment to the population of Central Norway (St. Olav's Hospital, 2024).

St. Olav's consists of more than twenty clinics, each responsible for their own specializations. These clinics consist of several departments, e.g. department of nurses or department of doctors. The departments in turn consist of several sections, each having their own manager. Examples for these sections are wards or outpatient clinics. Figure 1 shows the structure of the hospital.

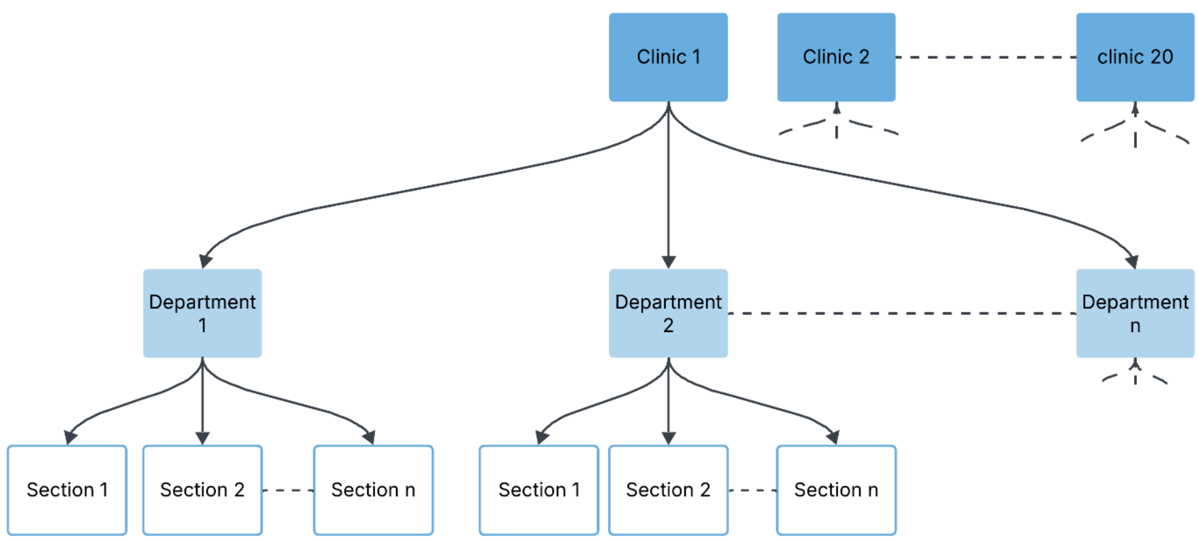

Figure 1: Structure St. Olav's Hospital

Compared to other hospitals, St. Olav's has a special layout. It is often the case that the operating rooms are the "heart" of the hospital, around which the rest is built. In St. Olav's the operating rooms are instead divided over various buildings, which belong to different clinics.

St. Olav's is working on an initiative in personnel scheduling. This is a project which can affect the decision-making processes on each of the levels in the whole organization. It consists of multiple smaller initiatives, one of which is nurse shift scheduling.

1.2 About the emergency department

This research is done in collaboration with the ED, which forms a clinic together with the Heart and Long Centre. Together they share one building in which the ED can be found on the ground floor and the basement. The ED treats about 27,000 patients per year. Over the last years they have experienced a large increase in number of patients per year. Patients that arrive by themselves enter at the reception on the ground floor. This is where a first check on their situation is done, and eligible patients are sent down to the ED. Here, they stay in the waiting room until there is capacity to help them. Some patients arrive by ambulance or helicopter, and are directly brought into the ED. Figure 2 shows a layout of the ED.

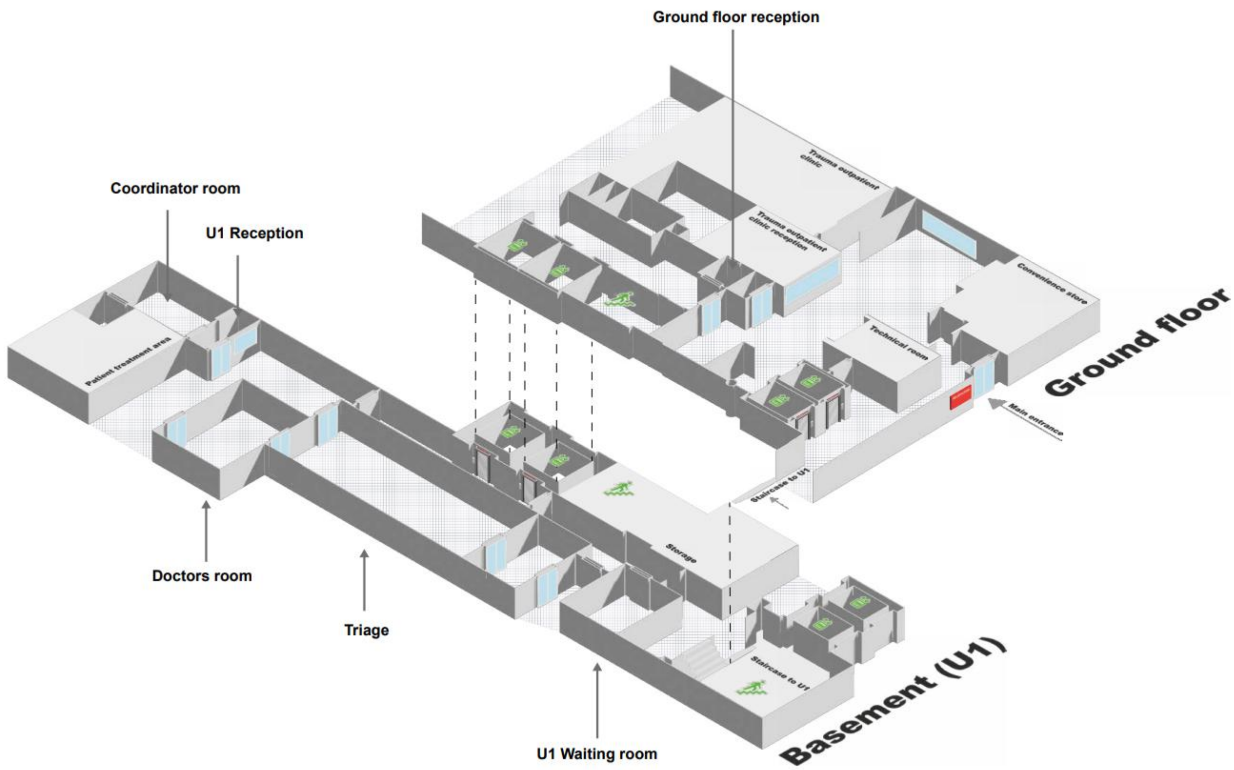

Figure 2: Layout ED

There are three emergency rooms (ERs), which can be used for surgery if it is really urgent. Preferably these rooms are not used for surgeries. There is one large triage room, which can hold nine patients at a time. There are nine smaller rooms which are used for assessment and treatment.

1.3 Background of the problem

The planning team at the ED faces several challenges. They report that there are too few nurses available to fully cover all shifts. Additionally, when unexpected demand arises, it is often difficult to find nurses willing to work extra hours. Many nurses indicate that their current workload is already high, leaving them too exhausted to take on additional shifts during emergencies. As a result, patient care sometimes has to be delivered with insufficient staff, which can negatively impact the quality of care.

Nurses have also mentioned that some shifts are very quiet, and they do not have much to do, while others are very busy and experienced as stressful. This indicates that the current shifts do not fit the demand curve well which results in less efficient use of the available nurses. Therefore, shift scheduling is a topic of interest for this department.

1.4 Current nurse shift scheduling

Currently, the ED schedules their shifts via a set of clear rules. They work with five shifts and five start times:

- 07:30 - 15:15
- 12:00 - 20:00
- 15:00 - 22:30
- 20:00 - 04:00
- 22:15 - 07:45

When scheduling the nurses, the first thing which is considered is the budget. The number of nurses hired must be in line with this. Next, the people responsible for planning ask each nurse their preference with regards to working weekends, night shifts, and other types of shifts. After getting all the preferences from the nurses, they start with planning the weekend shifts, as these are most difficult to fill. Then they continue with the night shifts. Next, they fill the rest according to shift patterns. Then the schedule is sent to the nurses, which now get the chance to give feedback. Based on the feedback the schedule is changed and is sent to the union, which needs to approve it before the official schedule is determined.

At each hour of the day there is a minimum number of nurses which need to be present. These numbers were determined based on experience. If the planning team at the ED wants to alter these minimums, they need to do this in discussion with the nurse union. When there is more staff available than the minimum nurses needed, they get scheduled during convenient hours to do work that is not related to patient care. When it is not possible to schedule as many nurses as is required by the minimum, it is possible to leave “gaps”, but these need to be very well explained to the union, otherwise they will not approve of the schedules.

While scheduling, there are certain rules from either the government or the union, that need to be followed:

- The maximum working hours per week is 48, which may be averaged over the planning period of 24 or 27 weeks, depending on whether we are talking about the first or the last half of the year.
- The rest period between shifts should normally be 11 hours but may be reduced to 9 hours.
- The breaks are included in the working hours
- Employees work every third weekend and no more than 9 weekend over the 27-week period.

The work for nurses does not only consist of task related to care for patients. There are also many indirect care tasks. Nurses at the ED explained that these tasks are more flexible in the sense of when to perform them than patient related tasks. Some of these tasks can be performed at any time, but some cannot. Some tasks can be done in any time span, but some need a minimum time block to be performed. For example, completing a course. Working on this for 10 minutes and then switch to something else, would not make sense. On the other hand, restocking some inventory can be done for a short while before switching to patient care.

Earlier research in the department tried mapping the number of nurses present onto the number of patients in the ED. Figure 3 shows the results. From this figure we can see that the flow of the number of nurses present is not in line with how the number of patients change. This shows that there is room for improvement in using the available nursing capacity.

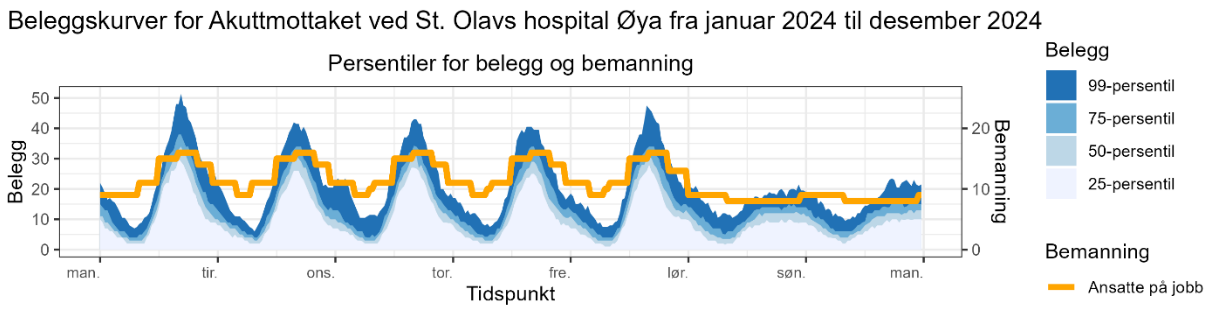

Figure 3: Number of patients and number of nurses during the week

1.5 Research design

The aim of this research is to find out how to develop a tool that makes the shift scheduling process efficient and finds an optimal shift schedule, and how the use of this tool can be implemented. Therefore, the main research question is:

"How can we develop a tool that can make the shift scheduling process more efficient and that can optimize the shift schedules in the emergency department at St. Olav's Hospital?"

We have formulated sub questions below, which help us to complete the main research goal.

1. How is the shift scheduling currently done in the emergency department at St. Olav's hospital and what is the current performance?

This question is meant to give a better understanding of the background of the problem and what information is already available and is discussed on Section 1.4. We use (informal) interviews to answer this question. Questions were asked to people involved in the nurse shift scheduling of the ED and the people working at St. Olav's involved in the project on optimizing the personnel scheduling.

2. Which model can best optimize the nurse shift schedule in the ED at St. Olav's hospital?
    a. What is shift scheduling?
    b. What are the different ways to define workload for nurses?
    c. What mathematical formulations are used to find optimal shifts for nurses?

The questions in this section are meant to help construct a solution to the experienced problem and reflect on how well it works for the emergency department. It is related to already existing knowledge which is then combined with the knowledge about the emergency department to formulate a solution approach. This question is answered by using a literature search and personal interpretation and is discussed in Chapter 2, 3, and 4.

3. How well does the model perform for the emergency department at St. Olav's hospital?

The goal of this question is to reflect on how well the constructed tool works. This is done by performing experiments with the model which is explained in Chapter 5.

4. How can the use of the tool be incorporated in the current way of working?

Due to the restriction of time, we cannot implement the solution we find. We do want to give advice on implementing the use of the tool. With the help of literature and by observing the organization we answer the fourth question in Chapter 6.

1.6 Research scope

The goal is to deliver a prototype of a tool which can schedule the nurses shifts optimally. This tool can also be used to experiment with different rules regarding the shift scheduling. This research focuses on the situation at the ED. There are several important factors which need to be defined for this tool to work:

1. The nurse workload
2. The mathematical formulation
3. Functionalities for the tool
4. A solution approach

The focus of this research is on defining these four factors and on getting a prototype of the tool. Besides, we evaluate the tool, and we give advice on implementation.

2. Literature review

The goal of this chapter is to explore the literature and use this to construct a solution for the ED. The topic of personnel scheduling is introduced in Section 2.1. Then we dive into nurse shift scheduling in Section 2.2. We conclude this chapter by summarizing in Section 2.3.

2.1About personnel scheduling

Personnel scheduling is an important topic for hospitals. In literature, personnel scheduling is getting much attention as a big percentage of the total direct costs comes from personnel (Akkermans et al., 2021; Ernst et al., 2004). Therefore, it is of interest to these organisations to handle the personnel scheduling process as efficiently as possible. This thesis relates to a subtask of the personnel scheduling domain, nurse shift scheduling. Specifically, in the emergency department where we deal with high demand variability. In this section we explain how nurse shift scheduling in the emergency department is positioned within the personnel scheduling topic.

Hulshof et al. (2012) defines a taxonomy for healthcare decision making existing of 6 types of care services and 3 levels of hierarchy. The decision-making regarding personnel scheduling may differ between these types of care services. Below the hierarchical levels are described and examples of decision making related to personnel scheduling are given.

- Strategic planning: addresses structural decision making. An example of decisions made on this level is how much staff needs to be hired, and which staff with what background is hired (Hulshof et al., 2012; Ozcan, 2009).
- Tactical planning: translates strategic planning decisions to guidelines that facilitate operational planning decisions. This level also includes scheduling, the process of deciding what people should be available at what moment (Hulshof et al., 2012; Ozcan, 2009).
- Operational planning: involves the short-term decisions making related to the execution of the health care delivery process. This level can be divided into online operational and offline operational planning.
    - Offline operational planning: reflects the in advance planning of operations, e.g. assigning nurses to shifts.
    - Online operational planning: reflects the control mechanisms that deal with monitoring the process and reacting to unplanned events, e.g. dealing with changes in staff needed because of emergency arrivals (Hulshof et al., 2012; Ozcan, 2009).

Ernst et al. (2004) define personnel scheduling as the process of constructing work timetables for staff so that an organization can satisfy the demand for its goods or services. Ernst et al. (2004) mention that personnel scheduling consists of several modules. Depending on the situation, some modules can be combined or performed at the same time. The following modules are considered:

- Demand modelling: To determine how many staff are needed at different times over some planning period, or rostering horizon. This is the process of translating some predicted pattern into associated duties, which can be in turn translated to demand for staff.
- Days off scheduling: this module focuses on determining how rest days should be distributed among workdays for various job roles.

- Shift scheduling: Selecting what shifts are to be worked and how many employees are needed on each shift.
- Line of work construction: In this step the work schedules or roster lines for each staff member are determined.
- Task assignment: Assigns tasks to be carried out during each shift.
- Staff assignment: Assigns individual staff to the lined of work.

This research is performed for the emergency department of St. Olav's Hospital. This department has to deal with high variability in patient arrivals, which in turn determines the demand for nurses. Therefore, the demand modelling is an important topic in this research. To define personnel scheduling, we combine the steps of the different papers. We split the scheduling tasks from Hulshof et al. (2012) into two tasks which we call shift scheduling and personnel allocation. Both of these tasks have subtasks as shown in Figure 4, which are based on the approach used by Ernst et al. (2004).

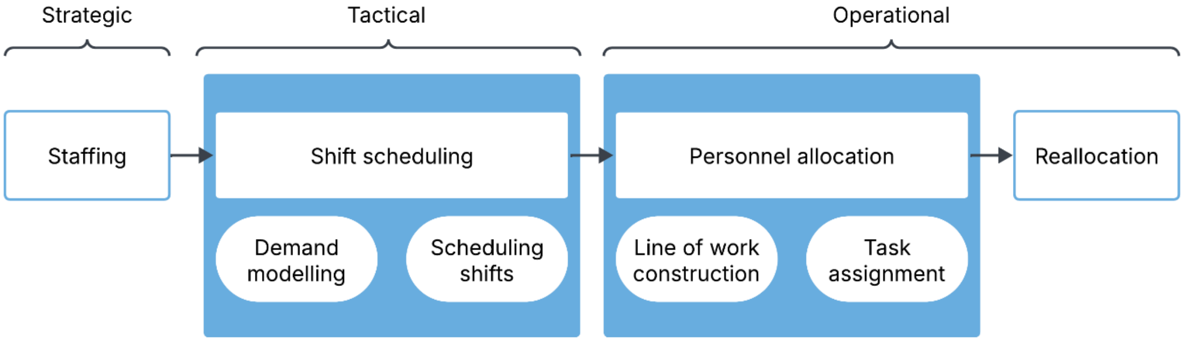

Figure 4: Shift scheduling process

2.2 Nurse shift scheduling

In this section we investigate the literature on the topic of nurse shift scheduling. We start by defining nurse shift scheduling. Then we continue by discussing two parts of nurse shift scheduling, starting with demand modelling. We analyze different ways of defining the workload for nurses and what data is needed to make use of these workload approaches. We conclude with the optimization model. We described the optimization model based on different characteristics and how these characteristics could be included in the model.

2.2.1 Nurse shift scheduling defined

Shift scheduling deals with the problem of selecting what shifts are to be worked and how many employees should be assigned to each shift to meet patient demand (Ernst et al., 2004). There are various decisions possible related to the shifts. Examples are when a shift should start, how long the shift continues. The objective is often to find shifts schedules that minimize the number of staff hours required to cover the desired staffing levels (Cardoen et al., 2010). Shift scheduling is important as it affects job satisfaction and productivity (Hulshof et al., 2012).

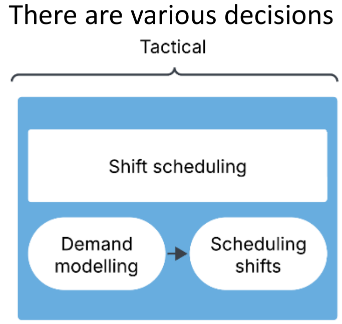

Figure 5: Shift scheduling process

As mentioned in Section 2.1, we consider shift scheduling to consist of both modelling the demand and scheduling suitable shifts (Ernst et al., 2004; Griffiths et al., 2020; Hulshof et al., 2012; Ozcan, 2009). The used definition of shift scheduling is visualized in Figure 5. The goal is to find a schedule which results in good use of the available resources.

When translating the demand into shifts, various constraints can be taken into account. An example is the shift length or the nurse-to-patient ratio (Daniel Wright et al., 2006).

Ernst et al. (2004) explains that there are three broad categories on which staff demand can be based:

- Task based demand: demand is obtained from lists of individual tasks to be performed. When rostering to task-based demand, the main task is to select a good set of feasible shifts to cover all tasks.
- Flexible demand: in this case the likelihood of future incidents is less well known and must be modelled using forecasting techniques. When rostering to flexible demand one should consider the timing of work and meal breaks within the limits allowed by the workplace regulations and company requirements.
- Shift-based demand: demand is obtained directly from a specification of the number of staff that ae required to be on duty during different shifts.

There are two types of shift scheduling:

- Cyclical scheduling: the schedule covers a designated number of weeks, the cycle length, and then repeats itself (Maier-Rothe & Wolfe, 1973; Ozcan, 2009). All employees of the same class perform exactly the same work, but with a different moment for the first shift (Ernst et al., 2004). This roster type is most applicable for situations with repeating demand.
- Discretionary scheduling: consists of two categories. The first is staggered scheduling. Here the hours of work per week do not change, but the employees are allowed to choose when to start their workday. The second is flexible scheduling. The employees select the schedule pattern that best meets their needs (Ozcan, 2009).

In nurse scheduling, the goal is to provide appropriate staff levels in the different medical wards in a hospital (Ernst et al., 2004). A more specific topic is nurse shift scheduling, which belongs to the tactical decision-making level, and is about determining how many nurses need to be present at what time and what shifts can cover this demand.

As mentioned in the previous section there are different approaches to shift scheduling. We discussed demand categories and types of shift scheduling. In nurse shift scheduling the demand is often based on shift-based demand (Ernst et al., 2004). Hospitals typically work with measures like nurse-to-patient ratios, which translate into a needed staff level. In nurse shift scheduling, cyclical scheduling is often used to meet staffing requirements in a way that is consistent with personnel policies and employee preferences (Maier-Rothe & Wolfe, 1973).

Applying our previously explained definition for shift scheduling to nurse shift scheduling results in two parts. First, to determine the nurse workload throughout the planning horizon, from which the number of nurses to be present at each moment can be determined. Next, this information is used as input to find the optimal shift schedule. This is visualized in Figure 6.

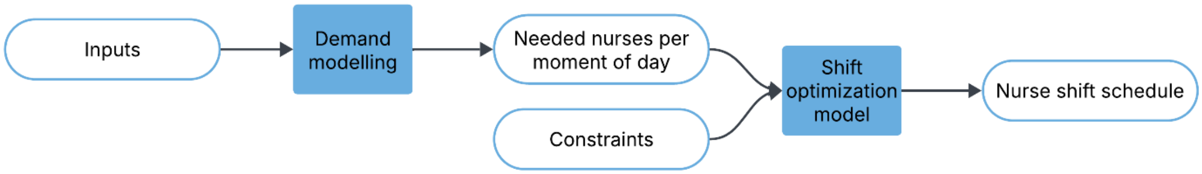

Figure 6: Information flow nurse shift scheduling

2.2.2 Demand modelling

Demand modelling is about defining some expected demand for care and translating this into a demand for nurse hours, as known as the workload for nurses. In this section we discuss the different definitions of nurse workload, and we look into some assumptions related to demand modelling.

Workload definition

Literature describes several ways of defining workload for nurses, but it is good to keep in mind that there is overlap between the approaches (Griffiths et al., 2020). To keep overview, we have divided the approaches in five different categories. We first discuss the census, the number of patients currently under care. Second, we look at the nurse-to-patient category. Third, we talk about expected work based on patient classification systems, which is divided into prototype and factor analysis systems. Lastly, we discuss how regression can help with defining nurse workload. It is structured in a way that the approaches get more and more complex and built further on some of the previously mentioned approaches. Each of the approaches is structured as follows: we start with a basic explanation of what it is, followed by what is found in the literature on it, then we discuss the benefits and downsides of using this approach, and we conclude with what data is needed when using this approach.

Census

Census is the number of patients occupying a bed at a given time, which relates to shift-based demand.

In many organizations the census is determined per day and doesn't differentiate between different parts of the day (Beswick et al., 2010). It is however also possible to look at intra-day census, meaning we look at the number of patients in the ward at several moments during the day. Beswick et al. (2010) explains that the patient volumes counted at midnight will differ significantly from those counted at three different times throughout the day. This can be related to nurse workload in a way that each patient represents an average amount of nursing hours needed.

Relating the census to the workload is a relatively easy approach as we only multiply the number of used beds with the average number of nursing hours needed. The downside of using the census to determine the nursing workload is that it does not differentiate between patient types.

To predict the census, we would need to know the probability distribution of certain number of a specific type of patients arriving at a certain moment. Besides, we would need to know the probability that a patient who is still in the ward on day n, is to be discharged that day. It would be a decision to make for how many separate points in the day we want to predict the census separately.

Nurse-to-patient ratio

Related to census, we have the nurse-to-patient ratio. The nurse-to-patient ratio builds further on the census. Nurse-to-patient ratios imply acceptable levels of patient care and nurse workload (Coffman et al., 2002).

This approach assumes that patient care is inversely related to the amount of work that nurses are given, therefore it addresses both nurse workload and patient care (Daniel Wright et al., 2006). Nurse-to-patient ratio relates to shift-based demand as this method assigns nurses based on a preferred of fixed ratio. Warner (2006) describes a more specific measure, Nursing Hours per Patient Day (NH/PD). If this value is known the Nursing Hours per Patient Shift can be calculated. For both of these values mentioned above, the management can communicate which minimum is wanted for what percentage of the time. This combined with knowing the census, shift decisions can be made.

The benefit of using nurse-to-patient ratio for determining workload is relatively easy. Once the ratio is determined and the census is known, it is easy to determine how many nurses are needed. The downside is that this method does not look at how to find the census. Furthermore, it does not differentiate between different types of patients and their corresponding care needs (Hurst, 2003).

To use this ratio in shift scheduling, it would make sense to get an agreed minimum and preferable value for the ratio and what an acceptable chance is to be below this ratio. The next step is to know the census. With these two values we can calculate how many nurses are needed at each moment to live up to the ratio.

Patient classification system

The goal of a Patient Classification System (PCS) is to determine the intensity of nursing care for a patient, including both direct and indirect nursing requirements (Williams & Crouch, 2006). This is done by identifying and classifying patients into care groups or categories, which each have their own required nursing effort. Certain characteristics can result in a very different number of nursing hours needed. We differentiate between two PCSs.

Prototype system

The first is called the prototype system. This system uses only a few tasks that have been shown to be predictors of the amount of care provided (Williams & Crouch, 2006). Patients are categorized into groups based on whether they demonstrate one or more of these critical indicators. Ozcan (2009) explains it as patients are classified according to the type of care needed. This type of classification is described as subjective. This approach relates to shift-based demand as patients are assigned to care categories and each of these categories have their own care requirements.

Experts have argued that patients with different problems in different diagnostic categories should generally receive differing amounts of nursing care (Sovie & Smith 1986). Warner (2006) explains that we could refine the NH/PD by dividing patients into patient classes. Each class has a corresponding number of nursing hours required during each part of the day. For example, class A needs X nursing hours during the early shift, Y during the late shift, and Z during the night shift.

Griffiths et al. (2020) also explains a way of using a prototype system to assign workload to certain patients. It is based on pre-existing categorisation of patients, in which each category has a corresponding amount of nursing hours. An example is relating the categories to the patient diagnosis. This is also used by Park & Murray (2006). They developed a nursing diagnosis-based classification system showing that the nursing diagnosis could independently predict patient outcomes by linking nursing diagnosis with length of stay, hospital charges, discharge disposition and mortality.

The benefit of prototype systems is that they are relatively easy to setup and use, but the downside is that they are highly subjective (Ozcan, 2009).

Factor analysis system

The second is called factor type system, which classifies by summing the relative values assigned to individual tasks or indicators of patient needs (Ozcan, 2009). This system identified a comprehensive list of tasks or procedures performed, with a numerical value given to each task based on the time taken to perform them (Williams & Crouch, 2006). These values are summed, and the category is determined by the number of points. This type of classification is objective. Factor-analysis systems related to task-based demand as it determined the total amount of work by summing time needed for separate tasks.

In the Oulu Patient Classification (Aschan et al., 2009) patients are assigned to one of four classifications, representing different amounts of care required, based upon a weighted rating of care needs across six dimensions: planning and co-ordination of nursing care, breathing, blood circulation and symptoms of disease, nutrition and medication, personal hygiene and secretion, activity, sleep and rest, and teaching guidance in care and follow up care, emotional support. Griffiths et al. (2020) also describes a factor analysis system. Here the categorisation of patients is done in a “bespoke” way, instead of them being pre-determined. An example is basing the categorization on level of acuity. When grouping patients based on acuity, we look at how severe the medical condition of a patient is. (Ozcan, 2009) describes an acuity-based system called NPAQ, which assesses various factors related to patient care requirements based on a questionnaire approach. Another approach mentioned by both Griffiths et al. (2020) and Ozcan (2009), considers all patient-related variables in the determination of the amount of care each patient should receive daily.

Factor analysis techniques provide a highly developed set of workload data which can be translated into nursing workload. However, developing a factor analysis method is both time consuming and difficult (Ozcan, 2009).

Using patient classification systems to determine nurse workload requires to connect certain nursing hours to certain characteristics or patient categories. For each characteristic or patient group, we need to determine a number of nursing hours needed during each point in time. Doing this for all expected patients, yields a prediction of nursing hours needed at each point in time. Combining this knowledge with the number of patients of each group we expect results in an estimation for needed nursing hours.

Regression on factors

Regression is related to flexible demand, as it uses historical data to forecast both patient arrivals and expected tasks. In some respects, regression-based models simply represent a particular approach to allocating time across a number of factors within an indicator-based system, rather than directly observing or estimating time linked to specific activities or patient groups.

Hurst (2003) identified regression-based approaches, which model the relationship between patient-, ward- and hospital-related variables, and establishment in adequately staffed wards. To obtain the recommended establishment for a particular ward, coefficients derived from the regression models are used to estimate the required staffing. Aschan et al. (2009) has used regression to determine the relationship between nursing time and type of nursing diagnoses. This was done with a regression equation of type $y = \alpha + \beta \times x$. The coefficient determination, R-squared would express the extent to which the variation in values of nursing time was accounted for by the type of nursing diagnoses that were identified. The final formula accounted for the following:

- Per-patient activity time constant specific to the discipline for a particular shift
- Per-patient activity time constant for each significant nursing diagnosis identified for the discipline
- Per-patient activity time constant for all the indirect patient care activities

Hurst (2003) explains that regressions are useful for situations where predictions are possible, such as the number of planned admissions. Furthermore, it is quite a cost-efficient method. The downside Hurst (2003) mentions is the need for appropriate data. If this data is not collected yet, this might be a struggle to set up

To use a regression to connect certain characteristics to nurse workload we need to have historical data which contains both the nursing hours as well as whether the characteristics of interest were present or not.

2.2.3 Shift optimization model

This section discusses the literature related to mathematical formulations for optimal nurse shift scheduling. We do this by going over certain characteristics of interest to the model.

Shift overlap

Traditional nurse scheduling often divides the day into three fixed 8-hour shifts, which works well when demand is evenly distributed. However, when demand fluctuates, alternative scheduling structures, such as overlapping shifts, can improve responsiveness during peak periods (Van Den Bergh et al., 2013). In some models, shift overlap emerges indirectly because of optimizing coverage. For example, in the work of Aickelin & Dowsland (2004), overlap occurs when multiple nurses are assigned to the same time period to meet demand, though overlap itself is not explicitly modelled. Aickelin et al. (2009) observe similar behavior, where overlapping shifts arise to reduce shortages or preference costs. In contrast, Awadallah et al. (2011) incorporate overlapping shifts directly into the shift design, such that early, day, and late shifts are intentionally structured to overlap during busy times. This built-in overlap improves coverage and can contribute to smoother transitions between shifts, even if handovers are not explicitly modelled.

Shift start times

Allowing flexibility in shift start times can significantly improve schedule adaptability to fluctuating demand (Awadallah et al., 2011). Several models adopt a predefined set of fixed start times to guide scheduling, offering structure but limiting flexibility (de Grano et al., 2009; Smet et al., 2014). Some models expand on this by including multiple start-time options within each shift type, enabling greater adaptability without full optimization (Di Gaspero et al., 2007; Topaloglu & Ozkarahan, 2004). More dynamic approaches treat shift start times as part of the optimization process. For example, Gutjahr & Rauner (2007) match precise hospital-defined demand intervals with nurse availability, and Özean (2005) uses a memetic algorithm to select start times from a set of allowed values. These models allow the schedule to respond more directly to demand peaks and staffing preferences. Overall, the treatment of shift start times ranges from static input to flexible decision variables, reflecting the balance between complexity and real-world usability.

Shift length

Flexible shift lengths can improve a schedule's ability to respond to fluctuating demand, making it an important factor in nurse rostering (Van Den Bergh et al., 2013). Several models use a fixed set of predefined shift durations, selecting from available options like 4-, 8-, or 12-hour shifts (de Grano et al., 2009; Gutjahr & Rauner, 2007; Özean, 2005; Smet et al., 2014). These models incorporate shift length into the scheduling process without optimizing it directly. In contrast, Topaloglu & Ozkarahan (2004) allow nurses to express preferences for shift duration, which are factored into the goal programming model to balance staff satisfaction with demand. The most flexible treatment is found in Di Gaspero et al. (2007), where shift length is actively optimized alongside start times to generate the most efficient coverage.

Coverage constraint

The coverage constraint is popular in the area of personnel scheduling (Van Den Bergh et al., 2013). As staff shortages are experienced everywhere it is necessary to work with short coverage constraints. Smet et al. (2014) distinguish between determined, range, and variable coverage types, incorporating them with weighted penalties to encourage load balancing and minimize understaffing. In contrast, Azaiez & Al Sharif (2005) and Topaloglu & Ozkarahan (2004) treat coverage as a goal rather than a hard requirement, allowing limited violations through weighted deviations in the objective function. Several models enforce coverage as a hard constraint, requiring exact fulfilment of staffing needs (Bellanti et al., 2004; Li et al., 2012; Yilmaz, 2012). Meanwhile, others such as Bilgin et al. (2012) and Parr & Thompson (2007) apply soft constraints, penalizing deviations while allowing flexibility in meeting coverage. Burke et al. (2006) refine this further by enforcing coverage over short time intervals, rather than full shifts, improving alignment with dynamic patient demand. Di Gaspero et al. (2007) approach coverage by generating minimal sets of shifts that meet hourly requirements, prioritizing efficiency. Purnomo & Bard (2007) ensure feasibility by incorporating exact or range-based coverage within the master problem of a branch-and-price framework. Finally, Sadjadi et al. (2011) incorporate uncertainty through a stochastic model that uses probabilistic thresholds to assess whether coverage is met. These varying formulations reflect the trade-off between strict feasibility, operational realism, and flexibility in handling staff shortages.

2.3Summary and conclusion

The goal of this chapter was to use literature to answer research question 2: "Which model can best optimize the nurse shift schedule in the ED at St. Olav's hospital?"

We began by defining nurse shift scheduling, which consists of two main components: demand modelling and shift scheduling. Literature identifies three approaches to defining staff demand: task-based, flexible, and shift-based (Ernst et al., 2004), and two types of scheduling: cyclical and discretionary (Ozcan, 2009).

Next, we explored demand modelling in more detail. Literature presents four main ways to define nurse workload: census, nurse-to-patient ratio, patient classification systems, and regression models. These methods can be combined and vary in complexity and data requirements. To translate hospital data into demand for care, we discussed several modelling assumptions. We found that both age and triage level influence nursing hours: older patients and those with more urgent triage levels require more care (Clopton & Hyrkäs, 2020; Iordache et al., 2020). We use this literature to define how the age and triage affect the care intensity, which is further explained in Section 4.1.

We also looked at how a patient's stay can be divided into phases, and how long each phase lasts. Wundavalli et al. (2019) measured 6-minute triage, with both discharge and admission taking 15 minutes. Duma & Aringhieri (2020) similarly report 5 minutes for triage, a variable period for tests and care, 10 minutes for revaluation, and 1 minute for discharge. Ruffing et al. (2014) found that, across six treatment categories the frequency-weighted average treatment time was 17.6 minutes. Based on this literature we decide to use triage (6 minutes), treatment (18 minutes), revaluation and discharge (13 minutes) as phases and discuss with the experts whether they think this is a good measure for the ED.

Literature on shift-optimization shows that shift overlap is treated both as coverage-driven assignments (Aickelin & Dowsland, 2004) and as a design element for peak-period responsiveness (Awadallah et al., 2011). Shift start times range from fixed slots (de Grano et al., 2009; Smet et al., 2014) to decision-variable approaches that align with demand intervals (Gutjahr & Rauner, 2007). Shift lengths are commonly selected from predefined options and weighted by nurse preferences (Topaloglu & Ozkarahan, 2004), while the most flexible frameworks optimize both length and start time for maximal coverage efficiency (Di Gaspero et al., 2007). Coverage constraints can be hard requirements guaranteeing staffing levels (Bellanti et al., 2004; Li et al., 2012), soft formulations permitting controlled understaffing (Bilgin et al., 2012), and stochastic thresholds that models demand uncertainty (Sadjadi et al., 2011). We will be using a set of allowed shift lengths and start times as input for the model, as this way it is easier to include ED preferences. Which ones to use will be decided by the optimization. We decided to allow shift overlap to naturally rise from the allowed shift lengths and start times. We will treat coverage as a hard constraint by setting a minimum service level. This level can be lower than the 100th percentile, depending on the user. We do use a penalty of the risk of understaffing when planning below the 100th percentile.

The next chapter presents the problem description, followed by a mathematical formulation for the model which can be used to solve the problem.

3. Problem description and model formulation

This section outlines the proposed approach to shift scheduling. We start with the problem description. Then we explain our decomposed approach, consisting of two steps. First, we model the demand for nursing care. Second, we optimize the shift schedule using a mathematical modelling approach.

3.1 Problem description

Emergency departments face highly variable patient arrivals throughout the week, creating a continuous tension between labour costs and adequate nursing coverage. To address this, we propose a two-stage planning model that first translates historical patient data into hourly care demand at a specified service-level percentile and then transforms this into a detailed nurse schedule.

The model finds the number of nurses needed for each hour of the day. The start times of shifts depends on the allowed start hours provided as input. The same goes for the allowed shift lengths. Across the week, each nurse is limited to one shift per day, no more than five shifts total, and total scheduled shift hours cannot exceed the sum of contracted capacity.

Saturdays and Sundays form the weekend, subject to special weekend-work rules. Nurses that work on Saturdays, also work on Sundays. Nurses that only work 8-hour shifts can work at most one in three weekends. Nurses that work shifts longer than 8-hours can work at most one in four weekends. It is possible to incentivize nurse to voluntarily work more weekends against a premium rate. How much extra shifts are available against which premium can be given as input.

There are different types of tasks for the nurses. The care tasks are specifically related to the patients and the care they need. Then there are also free flexible tasks which can be performed at any time and structured flexible tasks which can be only performed between certain hours and have a minimum length. The hours of flexible tasks depend on the total care hours.

Which service level to use is given as input by the user. The chosen service level forms a minimum for the amount of care which should be planned for in the shift schedule. The wage differs per time of day and day of week. There is also a different wage when using mertid and overtime. The difference between mertid and overtime is that mertid talks about the difference between the hired contract percentage and a full-time contract. Meaning a nurse who has an 80% contract will first work 20% mertid before the extra work counts as overtime. Mertid rates also depends on time of day and day of week, while overtime has a flat rate. If contracted nursing hours go unused, they still need to be paid.

There is a maximum of the total number of FTEs, and it should be possible to enforce a minimum average contract percentage among all nurses. It should also be possible to work with a limit on the number of nurses hired of certain contract percentages. Direct-care assignments must meet or exceed the demand level set by the chosen service level, with any shortfall covered by overtime or mertid but never beyond their allowed ratios.

Under these constraints, the model decides how many shifts of each allowable start time and duration to schedule over the week, how many nurses to employ at each contract percentage, and how to assign scheduled nurses each hour into direct care, free-flexible, or structured-flexible tasks. It also determines whether to use premium-paid extra weekend shifts and how many.

3.2Demand modelling

The goal of this step is to define the nurse workload. The product of this step is to the demand for care in a format that can be used in the optimization model. The desired format is the number of nurses needed for each hour to perform all tasks related to caregiving. How to do this highly depends on the available data, which is very case-specific. The approach explained in this section is in line with the literature in Section 2.2.2 under Demand modelling methods.

The idea is that we look at patient data from historical weeks, translate this into nurse demand, and then determine the demand percentiles for each hour during the week. A similar approach has been used to derive hourly nurse staffing needs in an emergency department setting (Gräff et al., 2016). The first step is to divide the patient's stay can be divided into phases. The phases should represent the activity each patient has to go through when being in the ED. To determine the nurse workload, it is important to know how long each of these phases take. The next step is to know how much care is needed for each of these phases. Therefore, we determine the nurse-to-patient ratios for each of the phases. These ratios vary from patient to patient. Therefore, to better estimate the nurse workload, we differentiate between patient groups. We use a patient classification system to make patient groups which have their own care intensity. We combine regression analysis with a patient classification system to find the impact of each characteristic. Patient types are defined based on characteristics which significantly impact nursing hours. Based on these characteristics the patient groups are defined.

Using the steps discussed before it is possible to transform the historical patient data into an estimate for nurse workload per hour. Outliers can be removed using the Interquartile Range (IQR) method, which identifies values that fall significantly outside the middle 50% of the data and apply Winsorizing to limit their effect (Dash et al., 2023). Based on the datapoints we calculate hourly percentiles, which serve as input for the shift optimization model. For example, this means that if we look at the 90th percentile that we would have enough nurses for 90% of the historical hours. At the end of the demand modelling, we find the number of nurses needed for each hour of the week. These will be decimal numbers which need to be rounded to integers. The rule which is used for rounding depends on the preferences of the user.

3.3Mathematical formulation

In this section we describe the mathematical model we developed to analyse the options regarding shift schedules. The goal is to provide a clear formulation of the simplified system in the emergency department of St. Olav's.

Sets

The used sets are described in Table 1.

Table 1: Sets

SET	DEFINITION
$\boldsymbol{t} \in \boldsymbol{T}$	time periods
$\boldsymbol{t^s} \in \boldsymbol{T^s} \subseteq \boldsymbol{T}$	possible shift start times
$\boldsymbol{l} \in \boldsymbol{L}$	allowed shift length
$\boldsymbol{d} \in \boldsymbol{D}$	days in a week
$\boldsymbol{k} \in \boldsymbol{K}$	contract percentages
$\boldsymbol{S_d}$	$\{(\mathrm{t}, \mathrm{l})\}$: shift starting time t with length l that starts on day d

$\boldsymbol{S^W}$	$\cup_{d=6}^{7} \mathrm{S_d}$: weekend shifts
$\boldsymbol{S_d^L(G)}$	{(t, l)}: Set off all shifts on day d ending in the last G hours of that day
$\boldsymbol{S_{d+1}^E(G)}$	{(t, l)}: Set off all shifts on day $d+1$ starting in the first G hours of that day
$\boldsymbol{B}$	$\{(\mathrm{t}, \mathrm{l})\}$: Set of all feasible structured flex work blocks, starting at hour t with length l

Parameters

The used parameters are described in Table 2.

Table 2: Parameters

PARAMETER	DEFINITION
DEMAND AND COST	
$\boldsymbol{P}$	minimum service level
$\boldsymbol{D_t}$	nursing demand required at time t based on chosen service level
$\boldsymbol{C_t}$	costs per nursing hour at time t
$\boldsymbol{C_t^M}$	costs of nursing hour when using mertid at hour t
$\boldsymbol{C^O}$	cost of nursing hour when using overtime
$\boldsymbol{C^B}$	base rate of costs per hour
$\boldsymbol{H}$	number of hours in a fulltime contract
SHIFT SETTINGS	
$\boldsymbol{X_{t^s,l,t}}$	$1\ if\ (t - t^s\ mod\ T) \in \{0,1, \dots, \mathrm{l}-1\}, 0$ otherwise. This is used to determine whether a shift which start at t^s and has length l covers hour t. It also wraps connect the Sunday evening hours with the Monday morning hours
$\boldsymbol{Z_{t,i}}$	1 if i care nurses are planned at time t, 0 otherwise
CONTRACT SETTINGS	
$\boldsymbol{F_k^{max}}$	maximum number of nurses with contract type k
$\boldsymbol{F_k^{min}}$	minimum number of nurses with contract type k
$\boldsymbol{\bar{k}}$	minimum average contract percentage
$\boldsymbol{F^{max}}$	maximum number of FTEs
FLEXIBLE PARAMETERS	
$\boldsymbol{R^F}$	ratio of flexible tasks to total demand
$\boldsymbol{R^{SF}}$	ratio of structured flexible tasks to free flexible tasks
$\boldsymbol{B^L}$	minimum length of structured flex work block
$\boldsymbol{B^S}$	earliest hour of the day a structured flex work block can start
$\boldsymbol{B^E}$	latest hour of the day during which structured flex work can be done
LIMITS AND CONTROLS	
$\boldsymbol{O}$	maximum ratio of overtime compared to planned hours
$\boldsymbol{M}$	maximum ratio of mertid compared to planned hours
$\boldsymbol{N}$	maximum number of nurses present at the same time
$\boldsymbol{G}$	minimum rest between two shifts
SERVICE LEVEL RESULTS	
$\boldsymbol{E_{t,n}}$	expected understaffing at time t while planning n nurses
$\boldsymbol{A_{t,n}}$	actual service level at time t while planning n nurses

Decision variables

The used decision variables are described in Table 3.

Table 3: Decision variables

DECISION VARIABLE	EXPLANATION
$\boldsymbol{s_{t,l}}$	number of shifts starting at time t with length l
$\boldsymbol{r_t}$	understaffing reduction by using flex hours at time t
$\boldsymbol{n_t^C}$	planned nurses for care at time t
$\boldsymbol{n_t^F}$	planned nurses for flexible tasks at time t
$\boldsymbol{n_t^{SF}}$	planned nurses for structured flexible tasks at time t
$\boldsymbol{b_{t,l}}$	number of structured flex blocks starting at time t and have length l
$\boldsymbol{f_k}$	number of nurses hired with contract percentage k
$\boldsymbol{u_t^M}$	number of understaffing hours covered by mertid at time t
$\boldsymbol{u_t^O}$	number of understaffing hours covered by overtime at time t

Objective

We use a bi-objective optimization approach with a lexicographic structure. This means we prioritize one part of the objective over the other. In the first stage, the model minimizes total staffing-related costs. In the second stage, given cost-optimal solutions from stage one, the model optimizes the allocation of flexible nursing tasks to maximize their contribution to mitigating potential understaffing in care delivery.

Objective 1: The objective is to minimize costs.

$$\mathrm{f}_1 = \sum_{i=0}^{l-1} C_{t^s+i} \sum_{t^s \in T^s} \sum_{l \in L} s_{t^s,l} \qquad \text{(staffing costs)}$$

$$+\sum_{t \in T} (C_t^M u_t^M + C^O u_t^O) \qquad \text{(mertid and overtime costs)}$$

$$+(\sum_{k \in K} R_k f_k - \sum_{t \in T} (n_t^C + n_t^F + n_t^{SF}))C^B \qquad \text{(pay for unused hours)}$$

Objective 2: The objective is to allocate flexible nursing tasks in a way that maximizes their ability to mitigate potential care understaffing.

$$\mathrm{f}_2 = \sum_{\mathrm{t \in T}} r_t \qquad \text{(reward flex when care shortfall)}$$

Bi-objective problem:

Min $\mathrm{f}_1(x)$, max $\mathrm{f}_2(x)$

Subject to: all model constraints explained in the following section.

Constraints

The constraints related to picking number of nurses are shown in Table 4. Here we assure that at each hour t, the sum of care, free-flex, and structured-flex nurses does not exceed the number of nurses actually scheduled on the shifts (1). Also, we make sure that there are never more nurses at the department at the same time than allowed (2). With (3) we make sure that the right number of care nurses gets selected. We will use this binary variable in a later stage. Lastly, we make sure that only one number of care nurses gets selected (4).

Table 4: Number of nurses constraints

	CONSTRAINT	EXPLANATION
1	$\boldsymbol{n_t^C + n_t^F + n_t^{SF} \leq \sum_{t^s \in T^s} \sum_{l \in L} X_{t^s,l,t} s_{t^s,l} \quad \forall t \in T}$	Planned nurses from shift coverage
2	$\boldsymbol{\sum_{t^s \in T^s} \sum_{l \in L} X_{t^s,l,t} s_{t^s,l} \leq N \quad \forall t \in T}$	Max nurses at one time
3	$\sum_{i=0}^{N} i\boldsymbol{Z_{i,t} = n_t^C \quad \forall t \in T}$	Assign right binary value for selected nurses
4	$\sum_{i=0}^{N} \boldsymbol{Z_{i,t}} = 1 \quad \forall \boldsymbol{t \in T}$	At most one number of nurses selected

Constraints related to service level and mertid/overtime trigger are shown in Table 5. Here we make sure that the demand following from the chosen service level is met by planned nurses (5). Secondly, we ensure that planned mertid and overtime covers the expected understaffing for the chosen service level (6). Thirdly, we ensure that the total mertid is capped by the contract types (7). This means that someone with an 80% contract can only have 20% of the 35.5 hours as metid.

Table 5: Service level, overtime, mertid constraints

	CONSTRAINT	EXPLANATION
5	$\sum_{i=0}^{N} \boldsymbol{A_{i,t} \geq P \quad \forall t \in T}$	Service level shortage not possible
6	$\boldsymbol{u_t^M + u_t^O \geq \sum_{i=0}^{N} E_{i,t} \quad \forall t \in T}$	Cover expected understaffing
7	$\boldsymbol{\sum_{t \in T} u_t^M \leq \sum_{k \in K} (1-k)\, H f_k}$	Mertid constraint

Constraint related to flexible tasks are shown in Table 6. Here the total free-flex hours are calculated as a fraction of the total nursing demand (8). Secondly, we calculate the structured flex hours (9). Thirdly, we calculate the number of nurses needed to cover the structured flex hours (10).

Table 6: Flexible tasks constraints

CONSTRAINT	EXPLANATION

8	$\sum_{t \in T} n_t^F \geq \sum_{t \in T} D_t R^F$	Flexible task constraint
9	$\sum_{(s,l) \in B} l b_{s,l} \geq \sum_{t \in T} D_t R^F\, R^{SF}$	Structured flexible task constraint
10	$n_t^{SF} \geq \sum_{(s,l) \in B} \sum_{t=s}^{s+l-1} b_{s,l} \quad \forall t \in T$	Link structured flex to # nurses

Constraint related to smart planning of flexible tasks are shown in Table 7. These constraints ensure that the amount of flex work counted toward mitigating understaffing at each time period does not exceed the number of available flexible nurses (11), nor can it exceed the actual shortfall between the maximum observed demand and the scheduled care nurses (12). This helps prioritize flex task placement in periods with risk of understaffing.

Table 7: Smart planning constraints

	CONSTRAINT	EXPLANATION
11	$r_t \leq n_t^F + n_t^{SF} \quad \forall t \in T$	Not more than available flex
12	$r_t \leq D_t^{max} - n_t^C \quad \forall t \in T$	Not more than actual shortfall

The constraint related to the contracts are shown in Table 8. Here we bound the number of nurses for each contract type (13). We also make sure there is a limit of the total number of FTEs used (14). Lastly, we enforce a minimum average contract percentage across all hires (15).

Table 8: Contract constraints

	CONSTRAINT	EXPLANATION
13	$F_k^{min} \leq f_k \leq F_k^{max} \quad \forall k \in K$	Contract limit per type
14	$\sum_{k \in K} k f_k \leq F^{max}$	Cap on max FTEs
15	$\sum_{k \in K} k f_k \geq \bar{k} \sum_{k \in K} f_k$	Minimum average contract percentage

The constraint related to the weekends are shown in Table 9. Here we make sure the number of Saturday and Sunday shifts are the same (16). This is in line with the rule that when someone works a weekend, they work both the Saturday and the Sunday. Then we also limit the weekend shifts to one-third of the total staff and to one-fourth for the people that work 12-hour shifts (17, 18). This is in line with the rule that nurses have to work at most one in three weekends and at most one in four if they work 12-hour shifts.

Table 9: Weekend constraints

	CONSTRAINT	EXPLANATION
16	$\sum_{(t^s,l)\in S_6} s_{t^s,l} = \sum_{(t^s,l)\in S_7} s_{t^s,l}$	Equal # shifts Saturday and Sunday
17	$\sum_{(t^s,l\neq 12)\in S_6} s_{t^s,l} + \frac{4}{3}\sum_{(t^s,l=12)\in S_6} s_{t^s,l} \leq \frac{1}{3}\sum_{k\in K} f_k$	At most 1/3, ¼, extra on Saturdays
18	$\sum_{(t^s,l\neq 12)\in S_7} s_{t^s,l} + \frac{4}{3}\sum_{(t^s,l=12)\in S_7} s_{t^s,l} \leq \frac{1}{3}\sum_{k\in K} f_k$	At most 1/3, ¼, extra on Sundays

The constraints related to the used hours are shows in Table 10. Here we ensure the scheduled shift hours do not exceed the contracted capacity (19). Secondly, we restrict to at most one shift per nurse per day (20). We do this by saying the number of shifts that start at a certain day should be smaller than the number of people hired. We also make sure at most five shifts can be worked per person hired (21) and that there is at least G hours rest between two shifts (22). It is important to note that these constraints are estimates as the model never assigns specific nurses to specific shifts. But still an estimate of the number of people needed to fill the shifts can be made.

Table 10: Shift constraints

	CONSTRAINT	EXPLANATION
19	$\sum_{t^s\in T^s}\sum_{l\in L} s_{t^s,l}\, l \leq H\sum_{k\in K} k f_k$	Shift hours within contract capacity
20	$\sum_{(t^s,l)\in S_d} s_{t^s,l} \leq \sum_{k\in K} f_k \quad \forall d\in D$	At most 1 shift per day per person
21	$\sum_{t^s\in T^s}\sum_{l\in L} s_{t^s,l} \leq 5\sum_{k\in K} f_k$	At most 5 shifts per week
22	$\sum_{(t,l)\in S_d^L(G)} s_{t,l} + \sum_{(t,l)\in S_{d+1}^E(G)} s_{t,l} \leq \sum_{k\in K} f_k \quad \forall d\in D$	At least G hours rest between shifts

The constraints related to overtime and mertid ratios are shows in Table 11. Here we bound total mertid (23) and overtime (24) to fractions of the planned care hours.

Table 11: Overtime and mertid ratio constraints

CONSTRAINT	EXPLANATION

23	$\sum_{t\in T} u_t^M \leq M \sum_{t\in T} n_t^C$	Mertid limit
24	$\sum_{t\in T} u_t^O \leq O \sum_{t\in T} n_t^C$	Overtime limit

The constraints related to non-negativity are $n_t^C, n_t^F, n_t^{SF}, s_{t^s,l}, f_k, u_t^M, u_t^O, b_{t,l}, t_t^S, w_6, w_7 \;\geq 0 \;\; \forall t, t^s, l$

Estimation model

The solution method we use consists of two steps. First it solves the model without allowing for extra weekend shifts. In this step the estimation parameters for the shift length and the costs for the extra weekend shifts are determined. This results in an assumption that the extra weekend shifts will cost the average off all other weekend shifts. In the second step, we allow for the extra weekend shifts and use the precalculated parameters. We use a two-step approach so that the average length and cost of weekend shifts can be pre-computed as fixed parameters, avoiding any bilinear or piecewise logic. Then the main model only needs two extra integer variables, keeping the model fully linear and fast to solve.

To be able to do this we introduce the following parameters:

Table 12: New decision variables estimation model

DECISION VARIABLE	EXPLANATION
w_6	extra Saturday shift to use
w_7	extra Sunday shifts to use

Table 13: Estimation parameters estimation model

ESTIMATION PARAMTERS	EXPLANATION
$A_6^C = \frac{\left(\sum_{(t^s,l)\in S_6} s_{t^s,l} \sum_{i=0}^{l-1} C_{t^s+i}\right)}{\sum_{(t^s,l)\in S_6} s_{t^s,l}}$	Average shift costs on Saturday
$A_7^C = \frac{\left(\sum_{(t^s,l)\in S_7} s_{t^s,l} \sum_{i=0}^{l-1} C_{t^s+i}\right)}{\sum_{(t^s,l)\in S_7} s_{t^s,l}}$	Average shift costs on Sunday
W^S	Additional weekend shift allowed per 3-week cycle
W^R	Multiplier applied to the cost of extra weekends
W	Boolean flag to run the model with or without extra weekends

When these are estimated, the model will be run again, with changed constraint and a changed objective. The change in constraints 17 and 18 allows more freedom when planning weekend shifts, based on the number of extra weekends shifts available. Constraint 25 and 26 are new and make sure that the extra shifts used do not exceed the number of extra shifts available.

Table 14: New constraint estimation model

CONSTRAINT	EXPLANATION

17	$\sum_{(t^s,l\neq 12)\in S_6} s_{t^s,l} + \frac{4}{3}\sum_{(t^s,l=12)\in S_6} s_{t^s,l} \leq \frac{1}{3}\sum_{k\in K} f_k + \frac{w_6}{3}$	At most 1/3, ¼, extra on Saturdays
18	$\sum_{(t^s,l\neq 12)\in S_7} s_{t^s,l} + \frac{4}{3}\sum_{(t^s,l=12)\in S_7} s_{t^s,l} \leq \frac{1}{3}\sum_{k\in K} f_k + \frac{w_7}{3}$	At most 1/3, ¼, extra on Sundays
25	$3 * w_6 \leq W^S$	Extra shifts used lower than available
26	$3 * w_7 \leq W^S$	Extra shifts used lower than available

Objective stage 1: The objective is to minimize costs.

$min \sum_{t^s\in T^s} \sum_{l\in L} s_{t^s,l} \sum_{i=0}^{l-1} C_{t^s+i}$ (staffing costs)

$+ \sum_{t\in T} C_t^M u_t^M + C^O * u_t^O)$ (mertid and overtime costs)

$+(\sum_{k\in K} R_k f_k - \sum_{t\in T}(n_t^C + n_t^F + n_t^{SF}))C^B$ (pay for unused hours)

$+ W^{rate}(w_6 A_6^C + w_7 A_7^C)$ (extra weekend costs)

Actual service level and expected understaffing

Before running the optimization, we translate our historical demand percentiles into a function that tells us, for any given number of nurses n at hour t, what percentile of demand they would satisfy. Concretely, we take our demand-percentile data and the corresponding required nurse counts and build a piecewise-linear curve (p). That curve gives the number of nurses needed to meet percentile p at hour t.

To find the actual service level achieved by scheduling direct-care nurses at hour t, we invert the demand curve to identify the percentile p at which exactly n nurses are required. Because the curve is piecewise-linear, this inversion corresponds to linearly interpolating between the two known data points between which the number of nurses falls. The result is a number between 0 and 1 that tells us the fraction of historical demand covered by those nurses.

$A_{t,n_t^C} = f_t^{-1}(n_t^C)\ \forall t \in T$, where $f_t(p)$ is a piecewise linear function defined by known data points $(p, D_{t,p})$ for all service levels and A_{t,n_t^C} is determined by interpolating the percentile p such that $f_t(p) = n_t^C$

Even when we plan to meet a chosen service percentile, there can be times where demand exceeds scheduled nurses. To quantify the average shortfall across all such times, we compute expected understaffing as follows:

1) Identify the under-percentile range: Look at every historical percentile p where p exceeds the actual service level. These are the situations in which demand would exceed the number of planned nurses.
2) Measure shortfall per percentile: For each percentiles p, determine how many extra nurses would have been needed.

3) Weight by probability: Each percentile p represents a probability mass, the chance that arrivals fall between percentiles. We take the gap between p and the actual service level as the approximate probability of that shortfall scenario.
4) Aggregate: Multiply each shortfall amount by its scenario probability and sum across all percentiles. The result is the average number of nurse-hours by which care would be understaffed at hour t. In practice, this is equivalent to a weighted sum over the demand percentiles. For each percentile above the actual service level, multiply the difference in required nurses by the probability weight for that percentile-range, and then sum them all.

Imagine we have plan 12 nurses at a certain time and that this corresponds to an actual service level of 85%. To reach a service level of 90% we would need 13 nurses, to reach 95% we would need 14 nurses and to reach 100% we would need 16 nurses. In this case you have a 5% chance of being understaffed by one nurse, a 5% chance of being understaffed by two nurses, and a 5% chance of being understaffed by three nurses. The expected understaffing therefore is: $0.05 * 1 + 0.05 * 2 + 0.05 * 3 = 0.3$ nurses.

4. St. Olav's Hospital case

In this section we make the general approach more specific to the case of St. Olav's hospital. We determine the settings which are later used in the experiments. We start with the demand modelling, then the optimization model.

4.1 Demand modelling

As mentioned in the previous sections, the demand modelling is highly dependent on the available data. Either we base nurse-to-patient ratios and the duration of each phase on historical data, or we need to make assumptions.

The first step was to decide on the phases of the patient care. In literature we find that a patient's stay can be divided into phases. Wundavalli et al. (2019) note that support activities constitute 28.92% of nursing work. Duma & Aringhieri (2020) propose five phases: triage, physician visit, tests and care (duration varies), revaluation, and discharge. Ajmi et al. (2015) alternatively identify three main phases: arrival and initial assessment, (re)orientation and treatment, and patient destination. Combining the literature with the expert opinion from the emergency department we assume the following main phases: triage, treatment, discharge, and general care during stay, as is visualized in Figure 7.

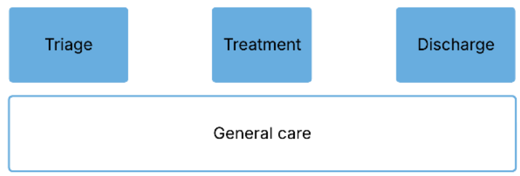

Figure 7: Phases in ED care

To determine the nurse workload, it is important to know how much time each of these phases take. As we have limited data on this, we need to make assumptions. Wundavalli et al. (2019) measured triage at 6 minutes, with both discharge/transfer and admission averaging 15 minutes. Duma & Aringhieri (2020) similarly report 5 minutes for triage, a variable period for tests and care, 10 minutes for revaluation, and 1 minute for discharge. Ruffing et al. (2014) found that, across six treatment categories the frequency-weighted average treatment time was 17.6 minutes. The nurses at the ED explained that they assume triage to take 15 minutes. The length of stay depends on various characteristics, but on average is 195.7 minutes (Karaca et al., 2012). When looking at the available data, we have information on the duration of the triage and the length of stay. Combining this with the expert opinions we assume the following:

- Treatment: 18 minutes
- Revaluation & discharge: 13 minutes
- Triage: We base it on available data, otherwise 15 minutes.
- Length of stay is based on available data.

The next step is to know how much care is needed for each of these phases. This is expressed in a nurse-to-patient ratio. Meaning, if it is one, we need exactly one nurse during the activity, if it is less than one the nurse can still perform other tasks during this activity. When it is more than one, we need multiple nurses during this activity. We have no data on the care needed during the care phases, therefore we need to make assumptions.

After discussing with the emergency department, we assume the following to be true:

- Triage: 1
- Treatment: 0.4

- Discharge: 1
- General care during stay: 0.3

The care intensity depends on certain patient characteristics. In an ideal data setting, we would combine regression analysis with a patient classification system. Patient types would be defined based on characteristics which significantly impact nursing hours. Based on the literature presented in Section 3.2.1 we include the following patient characteristics:

- Age: 0-9, 10-19, 20-59, 60-79, 80+
- Gender: male / female
- Triage level: blue, green, yellow, orange, red
- Clinical diagnosis: all possible diagnosis
- Arrival by ambulance: yes / no
- Need for hospital admission: yes / no

As the data of St. Olav's is limited, we only use age and triage level as characteristics. We assess how a patient's age affects nursing hours by defining age groups and corresponding adjustment factors. Iordache et al. (2020) report a positive correlation between age and direct care needs, using a 22-minute baseline per patient. They found that patients under 10 require an average of 10 minutes per stay, while those over 80 require 48.6 minutes. Therefore, we assume patients younger than 10 need on average 55% less care and the patients older than 80 need on average 120% more care. These factors affect the needed care intensity. Based on this we assume the following factors to be true for the age groups:

- 0-9: 0.45
- 10-19: 0.61
- 20-59: 1
- 60-79: 1.8
- 80+: 2.2

We have the same question for triage levels. The goal is to assign a factor to each triage level to reflect its deviation from average care. Clopton & Hyrkäs (2020) show a linear relationship between triage level and workload across five levels: Resuscitation, immediate and intensive has factor 1.7, Emergent, very urgent has factor 1.3, urgent, moderate cases have factor 1, semi-urgent has factor 0.7, and non-urgent has factor 0.3. Based on this we assume the following weights for our triage level:

- Blue: 0.3
- Green: 0.7
- Yellow: 1
- Orange: 1.3
- Red: 1.7

Combining age and triage factors with the nurse-to-patient ratios gives us 25 unique care paths, each representing a different care intensity. Using these care paths and the phase durations, we calculate the required nursing time per hour. We do this by calculating the nursing needed for each 15-minute interval. Since we have a full year of data and are scheduling for one week, we use 52 weekly data points per hour. We remove outliers using the IQR method and apply Winsorizing to limit their effect (Dash et al., 2023). Finally, we calculate hourly percentiles, which serve as input for the shift optimization model. It makes sense to look at the percentile per hour instead of per week as we want to guarantee a service level for each moment of the week. When working with percentiles per week, it could be the case that you have a 95% service level at the beginning of the week, but only an 85% service level at the end of the week. On average this would still be 90%, but that is not how the service level is intended.

At the end of the demand modelling, we find the number of nurses needed for each hour of the week. We have chosen a rather safe approach for rounding to integers and round down when the decimal is .25 or lower and round up when the decimal is higher.

4.2 Shift optimization model

This section outlines the components of the optimization model used to create the shift schedule. We begin with the assumptions needed, followed by the adaptations to the mathematical model which are needed for St. Olav's case.

Model assumptions

Due to incomplete information, it is necessary to make some assumptions to enable the use of the model. In this section we discuss all assumptions made. We start by discussing how to handle the difference in nursing costs during the week. Then we look into the difference in costs for unplanned hours.

Different nursing costs per hour of week

Nursing wages vary depending on the time of week. Internal research at St. Olav's has provided detailed cost estimates, accounting for both social costs and expected leave. We assume the following cost structure:

- Base hourly rate: 520 NOK/hour
- Evening premium (15-22): +32%
- Night premium (22-07): +40%
- Weekend premium (Saturday and Sunday): +29.5%
- Premiums are additive

These costs are used as penalties in the model. The same research showed that when considering factors such as sick leave and other forms of non-working time, it is best to assume a fulltime workweek results in 35.5 hours per week.

Different costs for unplanned hours

Unplanned staffing comes at a higher cost. The same study differentiates between:

- Mertid: additional hours for nurses not yet at a full-time workload
- Overtime: extra hours for those already scheduled at 100%

Mertid happens when a nurse has a part-time contract. The extra hours worked until the number of hours like in a full-time contract are reached, fall in the "mertid" category. As an example, if a nurse has an 80% contract, the first 20% of extra hours worked are mertid, the rest is overtime. The mertid base rate is 420NOK and follows the same premiums as the normal hourly costs. Overtime is estimated to be 100% higher than the base rate, which results in 1040NOK per hour.

Percentage of flex work

The ED mentioned there are also tasks that are not patient related. These tasks are more flexible in terms of timing than patient-related tasks. Wundavalli et al. (2019) note that support activities constitute 28.92% of nursing work. Some of these tasks can be performed at any time, but some cannot. We differentiate between two types of flexible tasks: free flexible tasks and structured flexible tasks. Structured flexible tasks are more structured in the sense that they can only be performed during specific hours and have a minimum duration. We assume a 50/50 division between free flexible tasks and structured flexible tasks.

Adaptations to mathematical formulations

For the case of St. Olav's we assign certain values to part of the parameters and sets. This is what we will describe here. The table 15 and 16 below show what information is given as input.

Table 15: Input sets

SET	EXPLANATION	SETTING
$\boldsymbol{T}$	Time	{0, 1, …, 167}
$\boldsymbol{D}$	Days	$\{1, 2, 3, 4, 5, 6, 7\}$
$\mathbf{K}$	Contract percentages	$\{0.2, 0.4, 0.6, 0.8, 1\}$

Table 16: Input parameters

PARAMETER	EXPLANATION	SETTING
$\mathbf{D_t}$	Demand for care nurses at time t	Explained in Section 5.2.1.
$\mathbf{C_t}$	Cost for an hour work at time t	Explained in the Section 5.2.2 under Model assumptions.
$\boldsymbol{C_t^M}$	Cost for an hour work as mertid at time t	Explained in the Section 5.2.2 under Model assumptions.
$\boldsymbol{C^O}$	Cost for an hour work as overtime	Explained in the Section 5.2.2 under Model assumptions.
$\boldsymbol{C^B}$	The base rate cost per hour work before premiums are applied	520
$\boldsymbol{H}$	The number of hours in a fulltime contract	35.5
$\boldsymbol{F^{max}}$	Max number of employees	1000000
$\boldsymbol{B^L}$	The minimum length of a structured flex block	2
$\boldsymbol{B^S}$	First hour structured flex blocks can start	8
$\boldsymbol{B^E}$	Last hour structured flex blocks can end	16
$\boldsymbol{N}$	The maximum number of nurses present at the same time	0
$\mathbf{E_{t,n}}$	Expected understaffing at time t when n nurses are planned	Explained in Section 5.1.2.
$\mathbf{A_{t,n}}$	Actual service level at time t when n nurses are planned	Explained in Section 5.1.2.

$\mathbf{D_t}$ is given as an input and represents the demand for care at each hour t. This is the results from the demand modelling and is visualized in Figure 8.

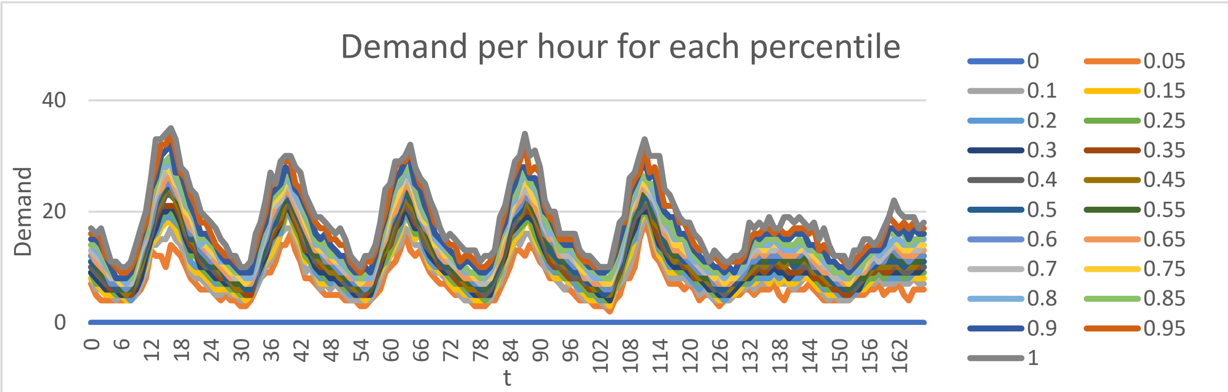

Figure 8: Demand per hour for each percentile

For the experiments we use the 90th percentile as this is in line with the literature explain in Section 3.1.2 under Demand modelling methods.

5. Experiments

The aim of this research is to find the optimal way of scheduling shift for the emergency department in St. Olav's. We start with technical experiments, which are used to determine which technical settings result in the best performing mathematical model. Next, we look at the managerial experiments which allow us to compare different instances and base conclusion on.

5.1Technical experiments

In this section we perform experiments that validate and tune the model on hand before we start with the managerial experiments.

All experiments were run on a laptop with an Intel® Core™ i7-9750H processor, 16 GB RAM, and Windows 11. Optimization models were solved using Gurobi 12.0.1 with Python 3.11.1. With the current settings, the model declares an optimum within 0.01% of true optimality. Running the model with 64 experiment runs in less than 15 minutes, which means less than 15 seconds per run. Because this is already a very small runtime it is not interesting to use methods for decreasing it.

5.2Managerial experiments

In real life, staffing decisions go far beyond just "make the model run fast." We must find out which settings positively influence the shift schedule. Based on this we give advice to the planning department. In this section, we show a series of "what-if" experiments that mirror those everyday choices in our model. For each policy we explain how we set it up, what we measure and what happens when we change the settings. We start with looking at the results of changing one setting at a time and later we combine multiple settings in one experiment. Over different experiments we will change the following sets and parameters to see the impact: start times, shift lengths, minimum average contract percentage, maximum overtime percentage, maximum mertid percentage, extra weekend rate and extra weekends. We also experiment with the length of stay of patients to see the impact on the costs.

As the goal is to find settings which improve on the current situation, we defined a "base case" which resembles the current way of planning in the ED. Currently, they work with 5 shifts which are all 8 hours. These have the following settings:

- Start times: 6, 12, 14, 20, 22
- Shift lengths: 8
- Min average contract: 0.8
- Max overtime: 0.1
- Max mertid: 0.1
- Flexible task ratio: 0.3
- Structured flex task ratio: 0.5
- Structured flex block minimum length: 2
- Structured flex block hours: 8-16
- Extra weekends: 0
- Extra weekend rate: 1

In Figure 9 we show both the demand curve on Monday and the standard shifts. It is meant to show which shift cover which part of the day. All shifts are 8 hours, but some start in the evening and continue into the morning.

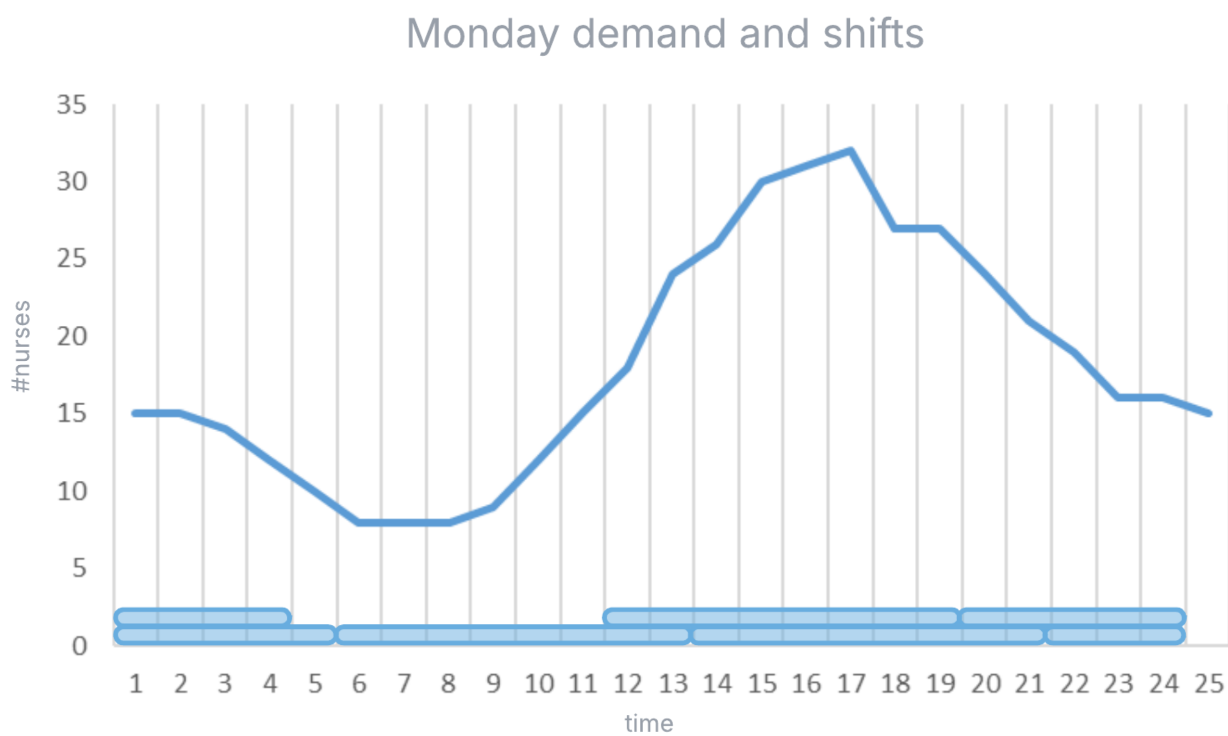

Figure 9: Demand curve and shifts

The next ten subsections dive into the experiments. The first section explains the tool we used for performing the experiments. For the first 10 experiments, we chose to display the total cost of the shift schedule, as this is an important decision-making factor. The total costs consist of the costs for planning the shift hours, the expected mertid and overtime costs, the extra weekend costs, and the costs for the unused contract hours. We also show a variable called "staffing costs". These consist of the nursing costs during the week. This only includes the hours in which effective work is planned, so not the idle hours, only the productive hours. This puts the total costs better into perspective, as this allows us to compare the costs which are really needed with the costs we eventually make. To get more context to the situation we also show the approximated number of FTEs needed to fulfil the schedule, the number of shifts planned, the idle hours and the unused contract hours. Idle hour are the hours which are planned when scheduling shifts, but do not have a specific task assigned to it. This is often the result of translating the separate tasks into workable shifts. Unused contract hours are the hours which are available based on the number of people hired but are not planned in the shift schedule because they are not needed.

The first ten experiments are about trying different rules regarding the scheduling of shifts. The optimization model uses these rules and finds an optimal solution which is in line with them. By performing the experiments, we see how each rule impacts the outcomes.

In the last experiment, we see how the shift schedules optimized via different rules perform when comparing them to new demand data. We then look at the cost of planning that shift schedule, the cost of understaffing, and the total costs. This way we compare the overall performance of each shift schedule.

5.2.1 Shift scheduling tool

To perform the experiments, we developed a tool. It takes input from the user and based on this performs the optimization and finds the optimal shift schedule together with some output measures. The tool can also be used to compare different settings. This allows the user to see what happens to the schedule when applying different scheduling rules. Appendix A shows three screenshots from the tool.

5.2.2 Shift lengths

Currently the nurses mainly work 8-hour shifts. We are interested in analysing what happens to the shift schedule when we allow for different sets of shift lengths. To get an idea of how sensitive shift schedules are to this decision, we ran "what-if" tests. We held all settings constant and then varied the shift lengths. The results are summarized in Table 17. Figure 10 shows the total costs and the idle time for each instance.

Table 17: Shift length experiment outcomes

INSTANCE	ALLOWED SHIFT LENGTHS	STAFFING COSTS	ESTIMATE FTE	TOTAL SHIFTS	IDLE HOURS	UNUSED CONTRACT HOURS	TOTAL COST
1	*8*	2,403,086	108	452	382	218	2,534,456
2	6, 8	2,344,337	108	451	301	298	2,517,197
3	8, 10	2,350,878	105.6	418	292	210.8	2,478,533
4	6, 8, 10	2,341,794	104.8	421	305	184.4	2,455,635
5	8, 12	2,337,093	104	396	308	158	2,437,154
6	8, 10, 12	2,340,385	105.6	426	327	206.8	2,465,821
7	6, 8, 10, 12	2,338,164	104	423	305	150	2,434,065

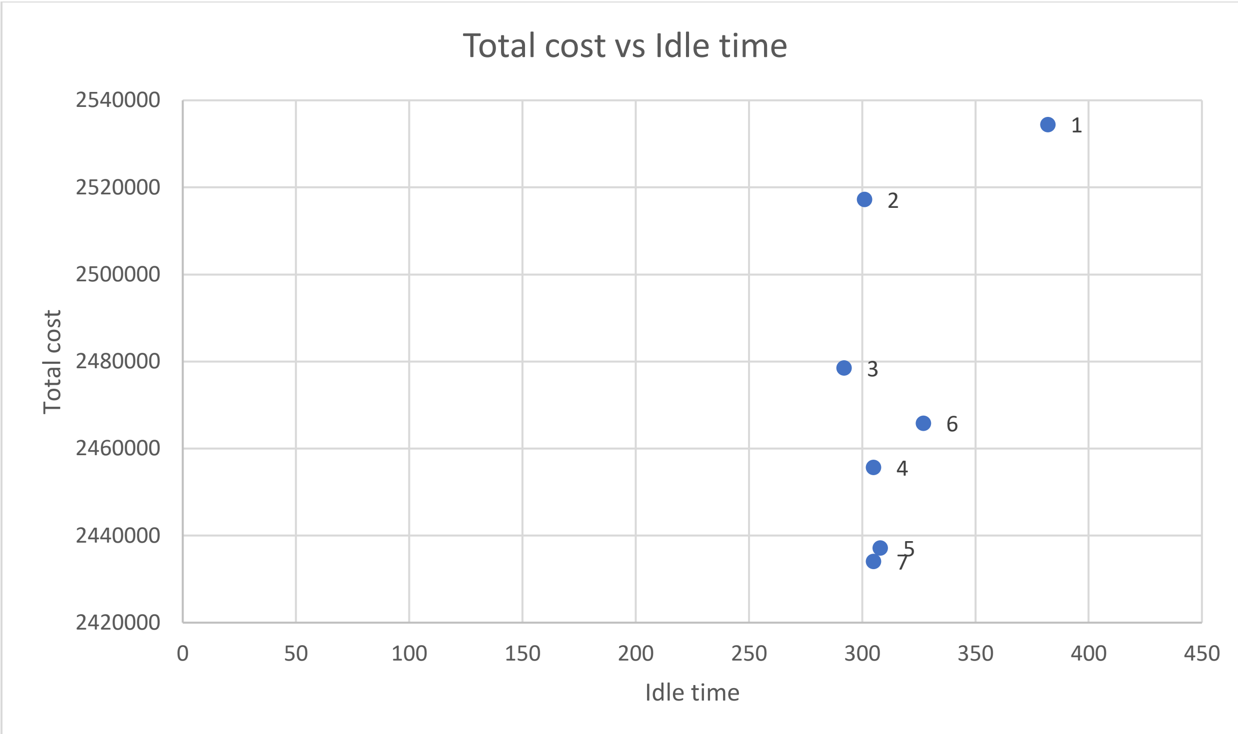

Figure 10: Shift length experiment outcomes

The results indicate that allowing a greater variety of shift lengths generally leads to improved staffing efficiency. Compared to the baseline scenario where only 8-hour shifts are permitted, introducing longer or more flexible shifts significantly reduces staffing costs, and the total number of shifts needed. It also reduces the FTE requirement and idle and unused hours, suggesting better alignment between staff availability and demand. These results suggest that incorporating a mix of shift lengths could be a cost-effective and resource-efficient strategy. We need to keep in mind that the effect of the shift lengths might change when we vary both the start times and the shift length at the same time. Besides, keep in mind that there are also non-monetary downsides to implementing different shifts lengths. Nurses might be less willing to work longer shifts or be less productive during longer shifts.

For the next experiments we have selected a subset of the above instances to perform further experiments with. These are described and named in Table 19.

Table 18: Shift length names

INSTANCE	ALLOWED START TIMES	NAME
1	8	SL1
5	8, 12	SL2
6	8, 10, 12	SL3

5.2.3 Start times

By default, nurses begin their shifts at one of five fixed times. To see how smoothing or tightening that window affects costs and workload, we ran seven "what-if" scenarios in which we only changed the set of allowed start times. All other settings remained the same as the base case settings. The results are summarized in Table 18. Figure 11 shows the total costs and the idle time for each instance.

Table 19: Start times experiment outcomes

INSTANCE	ALLOWED START TIMES	STAFFING COSTS	ESTIMATE FTE	TOTAL SHIFTS	IDLE HOURS	UNUSED CONTRACT HOURS	TOTAL COST
1	6, 12, 14, 20, 22	2,403,086	108	452	382	218	2,534,456
2	6, 14, 22	2,551,692	115.2	478	569	265.6	2,707,705
3	6, 10, 12, 14, 17, 20, 22	2,343,221	105.6	441	319	212.8	2,471,778
4	6, 10, 12, 14, 20, 22	2,353,166	108	442	284	298	2,526,027
5	6, 12, 14, 17, 20, 22	2,392,213	108	450	355	234	2,531,794
6	5, 7, 9, 12, 15, 17, 19, 21	2,475,304	129.6	464	492	888.8	2,955,381
7	6,10,12,14,22	2,418,198	115.2	452	373	473.6	2,682,685
8	6,8,10,12,14,22	2,370,441	115.2	442	325	553.6	2,676,213
9	7,9,12,15,18,23	2,335,952	108	442	321	298	2,508,796
10	7,9,12,15,18, 20,23	2,330,749	108	442	315	298	2,503,610
11	7,12,15,20,23	2,340,572	108	442	293	298	2,513,539
12	6, 8, 12, 14, 17, 20, 22	2,350,265	108	442	326	298	2,523,227
13	6, 8,10, 12, 14, 17, 20, 22	2,336,857	105.6	442	362	212.8	2,465,449

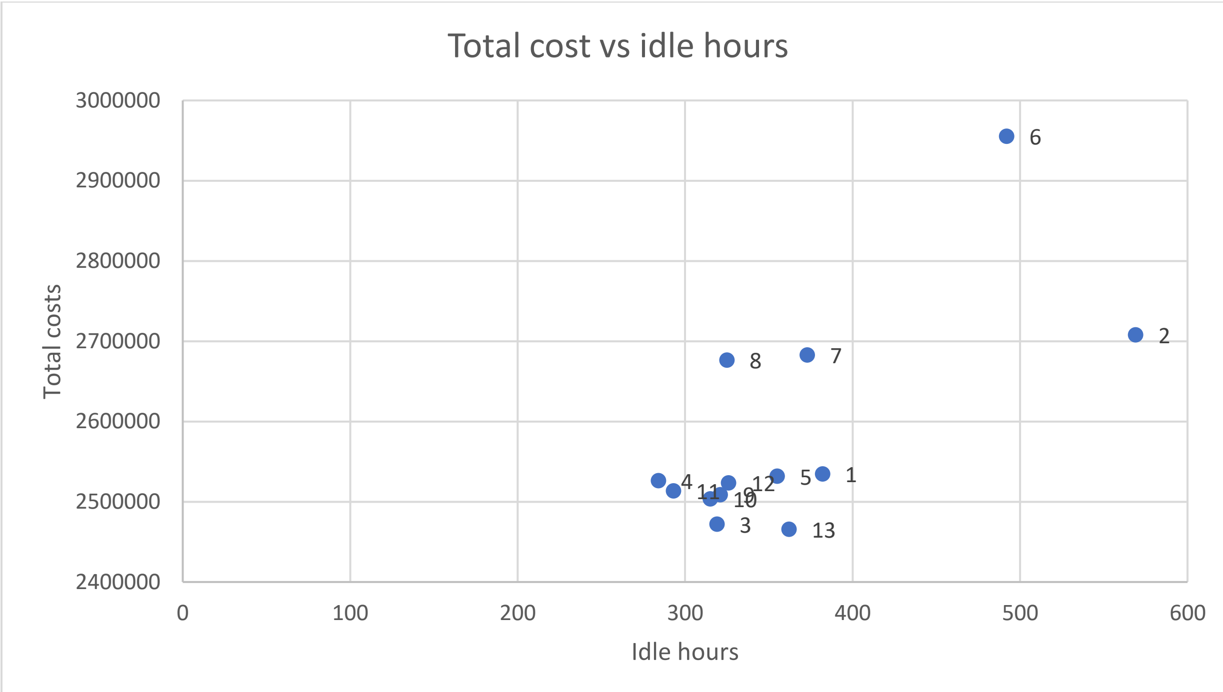

Figure 11: Start times experiment outcomes

The data shows that adding more start times can improve the schedule, but only when placed strategically. Various cases show lower costs as well as less idle time. We see that instances 3 and 13 show the best cost results.

For the next experiments we have selected a subset of the above instances to perform further experiments with. These are described and named in Table 19.

Table 20: Start time names

INSTANCE	ALLOWED START TIMES	NAME
1	6, 12, 14, 20, 22	ST1
3	6, 10, 12, 14, 17, 20, 22	ST2
13	6, 8,10, 12, 14, 17, 20, 22	ST3

5.2.4 Shift length and start times

In the real world, shift lengths are not picked in isolation but goes together with when shifts start. Longer blocks with tightly clustered start times might smooth demand in some hours but leave gaps elsewhere. To see these interactions, we ran tests where we varied with both the allowed shift lengths and the start times. By fixing the other variables, we see which combinations are efficient. The results are summarized in Table 20. Figure 12 shows the total costs and the idle time for each instance.

Table 21: Shift length and start times experiment outcomes

INSTANCE	SHIFT LENGTH, START TIMES	STAFFING COSTS	ESTIMATE FTE	TOTAL SHIFTS	IDLE HOURS	UNUSED CONTRACT HOURS	TOTAL COSTS
1	ST1, SL1	2,403,086	108	452	382	218	2,534,456
2	ST1, SL2	2,341,794	104.8	421	305	184.4	2,455,635
3	ST1, SL3	2,337,093	104	396	308	158	2,437,154
4	ST2, SL1	2,343,221	105.6	441	319	212.8	2,471,778
5	ST2, SL2	2,329,569	99.8	361	332	6.9	2,351,057
6	ST2, SL3	2,335,107	99.6	367	343	1.8	2,353,943
7	ST3, SL1	2,336,857	105.6	442	362	212.8	2,465,449
8	ST3, SL2	2,322,949	99.8	348	340	6.9	2,344,438
9	ST3, SL3	2,329,532	99.6	348	353	1.8	2,348,369

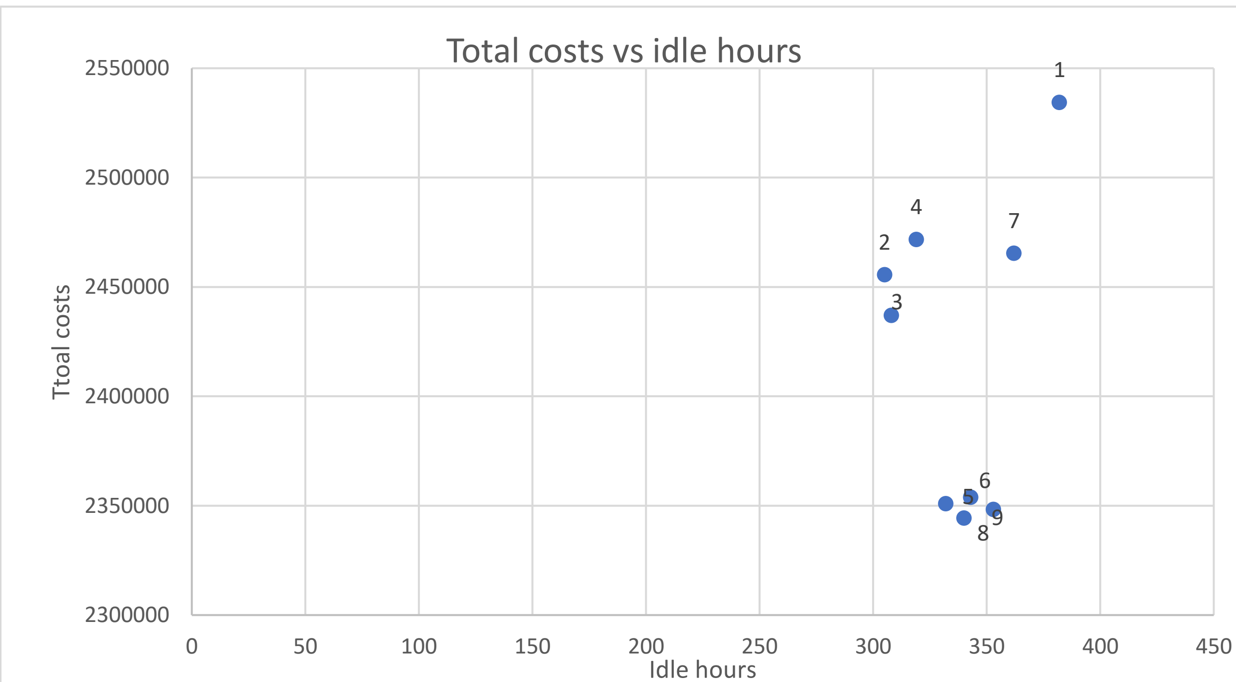

Figure 12: Shift length and start times experiment outcomes

From these results we see that a combination of multiple start times and shift lengths results to a more cost-efficient shift schedule. Especially the flexibility in shift length impacts the total costs. In the figure this is shows by instance 1, 2, 3, 4, and 7 being high up in the figure, which is because they allow for relatively little scheduling flexibility. The instances only working with 8-hour shifts are 1, 4, and 7. Instance 3 only allows five start times. The other instances allow for more start times and/or more shift lengths. The difference in costs between instances 5, 6, 8, and 9 are rather small. The non-monetary downside which can come along with certain settings are not represented in these results. Many different shift types might be experienced as a downside to the nurses and therefore impact how they experience their work in a negative way. Therefore, instances 5 and 8 are considered better.

Based on these outcomes we name some settings which we will use in later experiments. Instead of repeating the allowed shift lengths and start times, we give them names to improve readability. These names are summarized in Table 21.

Table 22: Setting names

NAME	START TIMES	SHIFT LENGTHS
ST1, SL1	6, 12, 14, 20, 22	8
ST2, SL1	6, 10, 12, 14, 17, 20, 22	8
ST3, SL1	6, 8, 10, 12, 14, 17, 20, 22	8
ST1, SL2	6, 12, 14, 20, 22	8, 12
ST2, SL2	6, 10, 12, 14, 17, 20, 22	8, 12
ST3, SL2	6, 8,10, 12, 14, 17, 20, 22	8, 12
ST1, SL3	6, 12, 14, 20, 22	8, 10, 12
ST2, SL3	6, 10, 12, 14, 17, 20, 22	8, 10, 12
ST3, SL3	6, 8,10, 12, 14, 17, 20, 22	8, 10, 12

5.2.5 Mertid

It is possible for nurses to work part-time contracts. We assume that nurses that choose to work part-time have their reasons not to work 100%. Therefore, working too high a percentage of mertid can negatively affect the nurses. Which is why we are interested in seeing the effect of capping the maximum mertid ratio. To see how allowing different mertid percentages affects the outcomes we have set up related experiments. A maximum mertid of 10% means that at most 10% of the total scheduled hour can be mertid. The results are summarized in Table 22.

Table 23: Mertid ratio experiment outcomes

MAXIMUM MERTID RATIO	STAFFING COSTS	ESTIMATE FTE	TOTAL SHIFTS	IDLE HOURS	UNUSED CONTRACT HOURS	TOTAL COSTS
0	2,387,486	108	449	354	242	2,535,866
0.05	2,403,086	108	452	382	218	2,534,456
0.1	2,403,086	108	452	382	218	2,534,456
0.15	2,403,086	108	452	382	218	2,534,456
0.2	2,403,086	108	452	382	218	2,534,456

The results suggest that enabling a modest level of mertid can lead to a slight reduction in staffing costs and unused hours. However, increasing mertid beyond this threshold yielded diminishing returns, with FTE and idle hours remaining fairly constant. While mertid improves flexibility and reduces unused capacity, its added value becomes limited beyond a certain level. Part of the reason for this result is that we are working with a relatively high service level, therefore we expect to meet most demand and we do not expect to be understaffed much. Therefore, the expected need for mertid is rather low. The difference in costs between allowing mertid and not allowing it can be explained by mertid being replaced by overtime now, which is more expensive.

5.2.6 Overtime

Overtime can be used to handle unexpected demand, but it comes at premium rates. To understand its true value how allowing more or less overtime affects the outcomes we performed five experiments. All other variables stayed constant. The results are summarized in Table 23.

Table 24: Overtime ratio experiment outcomes

MAXIMUM OVERTIME RATIO	STAFFING COSTS	ESTIMATE FTE	TOTAL SHIFTS	IDLE HOURS	UNUSED CONTRACT HOURS	TOTAL COSTS
0	2,403,336	108	452	378	218	2,534,636
0.05	2,403,086	108	452	382	218	2,534,456
0.1	2,403,086	108	452	382	218	2,534,456
0.15	2,403,086	108	452	382	218	2,534,456
0.2	2,403,086	108	452	382	218	2,534,456

Allowing some overtime provides a small cost benefit. Idle and unused contract hours also show small improvements. However, the gains are relatively small and further increases in the overtime ratio show diminishing returns. Similarly to the mertid experiments, the results make sense because we are working with a high service level and are not expecting to use much overtime. A non-monetary downside of using overtime is the strain on the nurses. This should be balanced with the benefit of having operational flexibility.

5.2.7 Average contract percentage

Staff consists of a mix of full-timers and part-timers. At St. Olav's everyone works one out of three weekends, no matter the contract type. Therefore, when filling the weekends shifts in the weekend, having part time contracts might be very valuable. We varied the minimum average contract percentage to see how contract mix affects the outcomes. This helps pinpoint the right blend of full- and part-time staff for stable, cost-effective coverage. The results are summarized in table 24. Figure 13 shows a visualization of how the FTE estimate changes depending on the contract percentage.

Table 25: Average contract percentage experiment outcomes

AVERAGE CONTRACT %	STAFFING COSTS	ESTIMATE FTE	TOTAL SHIFTS	IDLE HOURS	UNUSED CONTRACT HOURS	TOTAL COSTS
10	2,365,454	100	444	308	5.1	2,386,007
20	2,365,454	100	444	308	5.1	2,386,007
30	2,362,610	100	443	293	6	2,383,630
40	2,362,610	100	443	286	6	2,383,630
50	2,362,610	100	443	307	6	2,383,630
60	2,374,414	100.4	445	303	4.2	2,394,498
70	2,376,436	100.6	446	318	3.3	2,397,403
80	2,424,958	108	456	418	186	2,539,578
90	2,410,200	121.6	453	382	692.8	2,788,850
100	2,409,160	135	453	382	1168.5	3,039,320

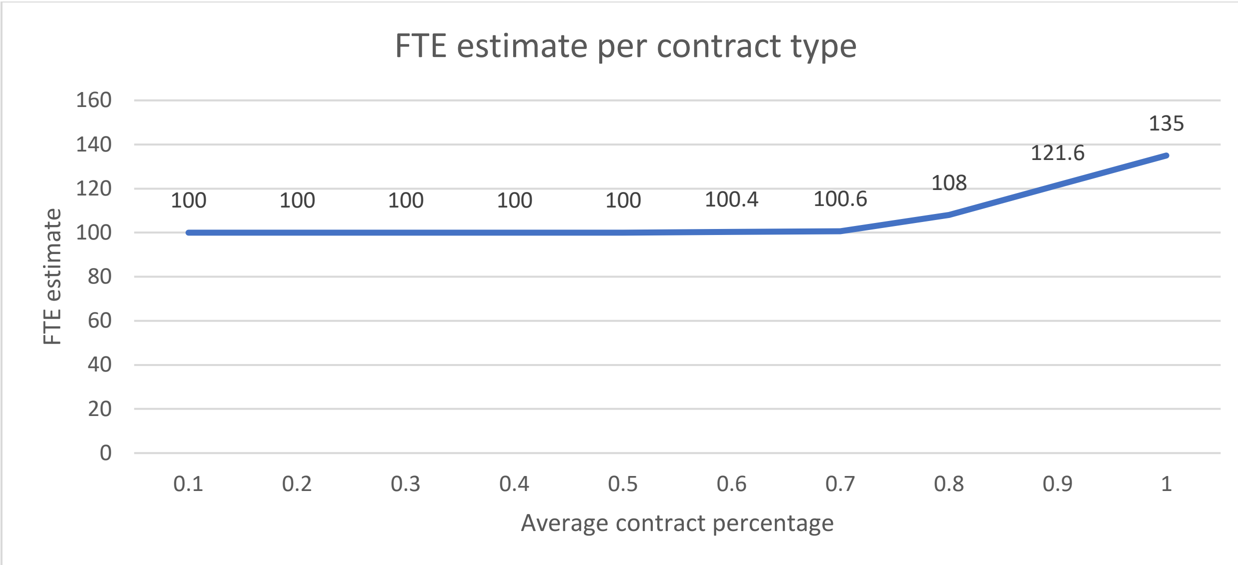

Figure 13: Average contract percentage experiment outcomes

In the results we see that the lower the contract percentage, the lower the costs, which stagnates between ten and twenty percent. This can be explained by the fact that the weekend shifts are "difficult" to schedule. With current rules, every nurse can work at most one out of three weekends, no matter the contract percentage. This means to fill the weekend shifts we need a certain number of nurses. If we would only hire people with a high contract percentage we end up with too many leftover hours during the week. We can see this in the table by the increasing number of unused contract hours, when the contract percentage increases. We see an increase in the number of FTEs, because the number of people needed to fill the weekend shifts stays the same, but we increase the contract percentage, which results in more FTEs. It is good to keep in mind that there can also be a non-monetary downside to having a lower contract percentage. An example is that the efficiency might be lower because people spend more time on "updating" each other, or the quality of care might be lower because of more switching staff.

Another interesting way to look at the different contract percentages, is to see how the costs evolve when allowing for more flexibility. We have looked at this "price path" for six different contract types. The first step of the price path is the base setting, then we allow for more start times while still having only 8-hour shifts. Lastly, we allow also 12-hour shifts. In the Table 25 we see how the total costs changed when allowing for more flexibility. We also displayed this in Figure 14 for better visualization.

Table 26: Cost flow per average contract percentage

AVERAGE CONTRACT %	ST1, SL1	ST3, SL1	ST3, SL2
60	2,394,498	2,407,246	2,379,288
70	2,397,403	2,369,494	2,361,109
80	2,539,578	2,465,449	2,348,239
90	2,788,850	2,712,992	2,361,794
95	2,914,988	2,845,029	2,448,639
100	3,039,320	2,959,061	2,555,561

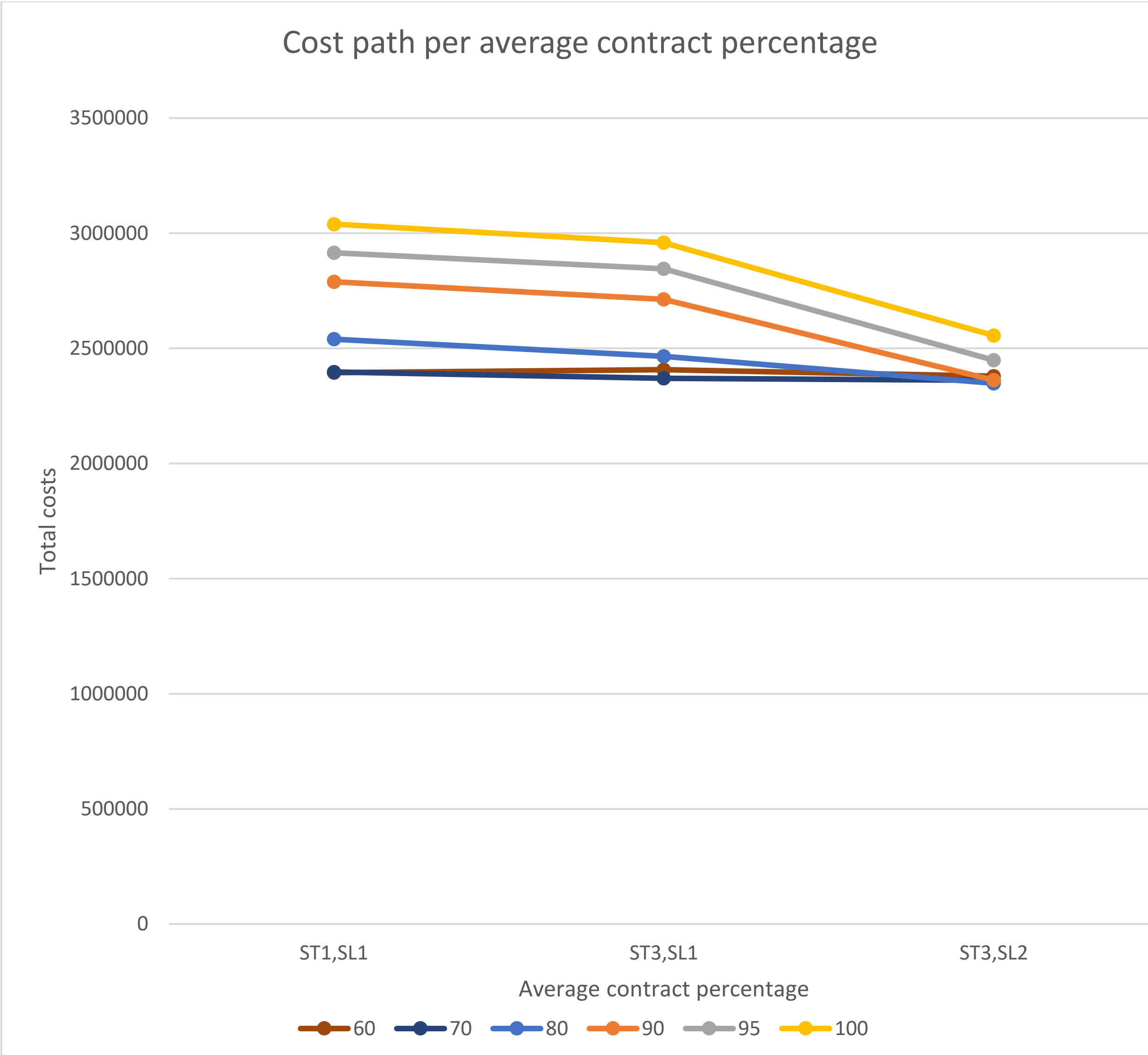

Figure 14: Cost flow for different contract percentages

From Figure 14 we can see that when comparing a high contract percentage, and therefore little flexibility, with a lower contract percentage, and therefore higher flexibility, the lower contract percentage performs better. But when adding flexibility in the scheduling options, the differences get smaller and the gap between higher and lower contract percentages gets smaller. The decreasing differences can be explained by the weekend shifts being the bottle neck. We get access to relatively more weekend shifts, if we attain a lower contract percentage. When allowing for more flexibility we create other ways of covering the weekend shifts, like using only long shifts in the weekend and short shifts during the week. There can also be non-monetary downsides of using a lower average contract percentage, like nurses spending more time on updating each other. Based on these results using an average contract percentage of 90% of lower is beneficial.

5.2.8 Length of Stay

How long patients stay in the ED affect the nurse workload. Based on the expert opinion at St. Olav's a non-efficient flowthrough of patients into other departments of the hospital results in an unnecessary long stay at the ED, which in turn affects the workload for the nurses. Therefore, it is of interest to see how much can be won when this flowthrough improves. Therefore, we look at different length of stays and how they affect the outcomes. Table 26 summarized the outcomes and Figure 15 shows the 90th percentile demand curves for each change in length of stay.

Table 27: Length of stay experiment outcomes

LENGTH OF STAY INCREASE	STAFFING COSTS	ESTIMATE FTE	TOTAL SHIFTS	IDLE HOURS	UNUSED CONTRACT HOURS	TOTAL COSTS
+5%	2,478,221	110.4	466	427	191.2	2,592,375
0%	2,403,336	108	452	375	218	2,534,279
-5%	2,316,652	105.6	434	359	268.8	2,474,185
-10%	2,232,209	98.4	420	363	133.2	2,319,140
-15%	2,148,063	98.4	403	340	269.2	2,304,560

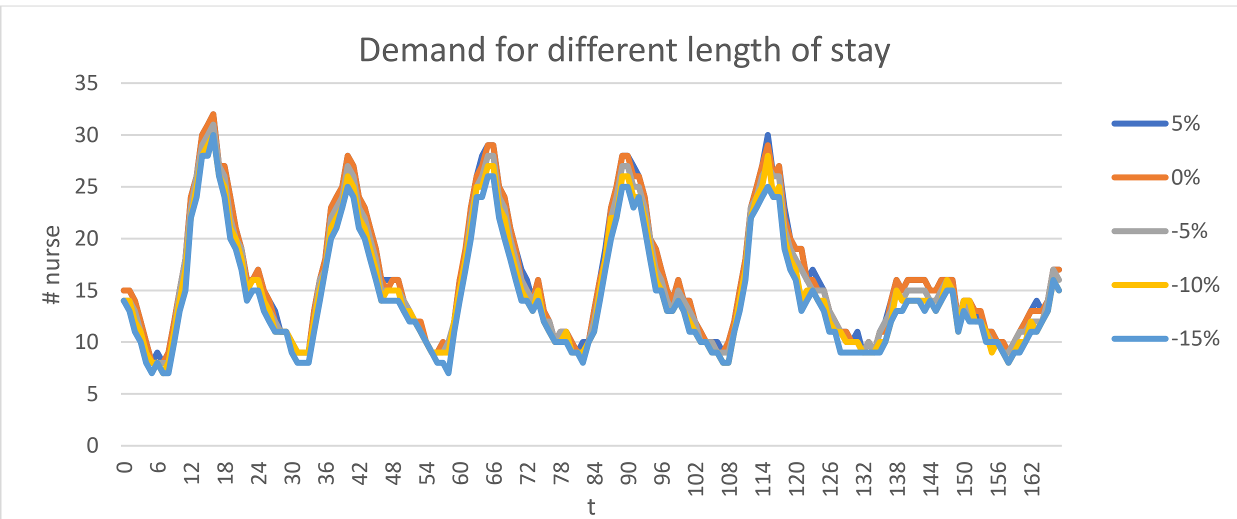

Figure 15: Demand for different length of stays

In these results we see that the length of stay directly affects the total costs. This is explained by having less work and therefore needing less nurses. Based on this we conclude that it would be worth investigating a better connection between departments, to make more efficient use of capacity.

5.2.9 Extra weekends

Weekend coverage at St. Olav's is tight. Nurses can work at most one out of every three weekends, and sometimes even only one in four, which often leaves gaps in service or forces expensive overtime. One way to improve this situation could be to check whether nurses are willing to work an extra weekend in exchange for a higher pay during this extra weekend. This comes down to "buying" more weekend capacity. But how valuable are those extra shifts, and how do they improve schedule efficiency? To investigate this, we run a set of experiments where we vary the number of additional weekend shifts nurses are willing to take on and the premium rate for those shifts. The normal wage is multiplied with the premium rate to find the new costs. All other settings are the same as in the base case. The results are summarized in Table 27. This table shows the number of extra weekend shifts available against what premium. Figure 16 shows how the number of FTEs needed changes depending on the premium costs for the extra weekend shifts.

Table 28: Extra weekend experiment outcomes

INSTANCE	EXTRA SHIFTS, EXTRA COSTS	STAFFING COSTS	EXTRA WEEKEND COSTS	ESTIMATE FTE	TOTAL SHIFTS	IDLE HOURS	UNUSED CONTRACT HOURS	TOTAL COSTS
1	0, 0	2,403,086	0	108	452	382	218	2,534,456
2	1, 1	2,409,680	11044	105.6	451	376	140.8	2,500,950
3	1, 2	2,420,641	22089	105.6	451	376	140.8	2,511,758
4	1, 3	2,440,505	33134	105.6	453	394	124.8	2,523,302
5	1, 4	2,443,355	44178	105.6	451	365	140.8	2,534,471
6	1, 5	2,398,635	0	108	451	374	226	2,534,464
7	2, 1	2,425,425	22089	103.2	452	385	47.6	2,468,078
8	2, 2	2,447,515	44178	103.2	452	385	47.6	2,490,413
9	2, 3	2,470,020	66268	103.2	452	382	47.6	2,512,723
10	2, 4	2,487,533	88357	103.2	451	362	55.6	2,534,346
11	2, 5	2,403,086	0	108	452	382	218	2,534,456
12	3, 1	2,410,402	33134	100.8	447	329	2.4	2,429,551
13	3, 2	2,445,123	66268	100.8	447	319	2.4	2,464,271
14	3, 3	2,478,257	99402	100.8	447	328	2.4	2,497,674
15	3, 4	2,491,777	88357	103.2	452	380	47.6	2,534,429
16	3, 5	2,403,086	0	108	452	385	218	2,534,456

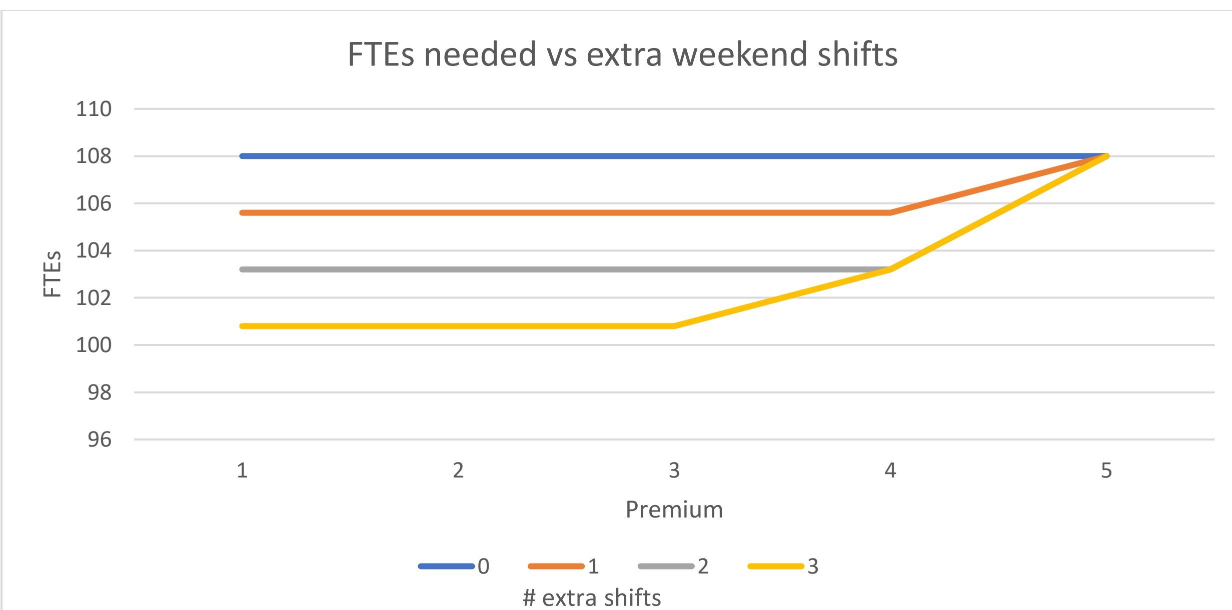

Figure 16:Extra weekend shifts experiment outcomes

While the baseline configuration offers no additional shifts, experimental setups allowing more weekend work with increased premiums show potential for reducing total staffing costs and improving efficiency. The results show that there is a threshold premium beyond which the optimal solution no longer includes additional weekend shifts. The results show that for one extra worked weekend we are willing to pay between 4 and 5 times more than the normal wage. When comparing the costs between the first two instances we see that one extra weekend saves us 1.32% of the costs. These findings suggest that extra weekend shifts can be a strategic tool, but the cost-benefit ratio must be managed. In the outcomes we can also see that the number of total shifts does not change. This is because the model plans the needed number of shifts to meet the wished demand. All the hours which we do not need fall under unused hours. In the results we can see the number of shifts stay the same while the number of unused hours decreases when we get access to extra weekend shifts.

5.2.10 Longer shift only in weekend

Using shifts longer than 8 hours can have non-monetary downsides, like being more tiring for the nurses. Therefore, it is worth investigating what happens with the results if we only allow long shifts in the weekend compared to allowing them throughout the whole week. We experiment with this for several start time and shift length settings. The rest of the settings are the same as the base case. The results are summarized in Table 28 and Figure 17 visualizes the cost comparison. The dark blue columns show the costs when only 8-hour shifts are allowed. The orange bars show SL3 where 10- and 12-hour shifts are only allowed in the weekend. The grey bars show SL3 when the 10- and 12-hour shifts are allowed throughout the whole week. The yellow bars show SL2 when the 12-hour shift are allowed in only the weekend. The light blue bars show SL 2 when the 12-hour shifts are allowed throughout the whole week.

Table 29: Longer shift weekend experiment outcomes

INSTANCE	NAME, LONGER SHIFTS WEEKDAYS?	STAFFING COSTS	ESTIMATE FTE	TOTAL SHIFTS	IDLE HOURS	UNUSED CONTRACT HOURS	TOTAL COSTS
1	ST1, SL1	2,459,491	108	469	0	82	2,509,391
2	ST1, SL2, N	2,404,308	104.8	437	0	32.4	2,431,302
3	ST1, SL2, Y	2,435,919	104.8	463	0	4.4	2,446,698
4	ST1, SL3, N	2,400,466	104	421	0	0	2,410,818
5	ST1, SL3, Y	2,420,844	104	460	0	2	2,431,358
6	ST2, SL1	2,414,391	105.6	463	0	44.8	2,447,066
7	ST2, SL2, N	2,306,816	99.8	379	0	2.9	2,325,906
8	ST2, SL2, Y	2,319,273	99.8	436	0	2.9	2,338,358
9	ST2, SL3, N	2,303,839	99.8	386	0	0.9	2,321,802
10	ST2, SL3, Y	2,319,083	99.8	437	0	0.9	2,337,032
11	ST3, SL1	2,421,089	105.6	466	0	20.8	2,440,482
12	ST3, SL2, N	2,298,460	99.8	363	0	2.9	2,317,589
13	ST3, SL2, Y	2,313,015	99.8	436	0	2.9	2,332,163
14	ST3, SL3, N	2,296,226	99.8	361	0	0.9	2,314,052
15	ST3, SL3, Y	2,313,412	99.8	437	0	0.9	2,331,316

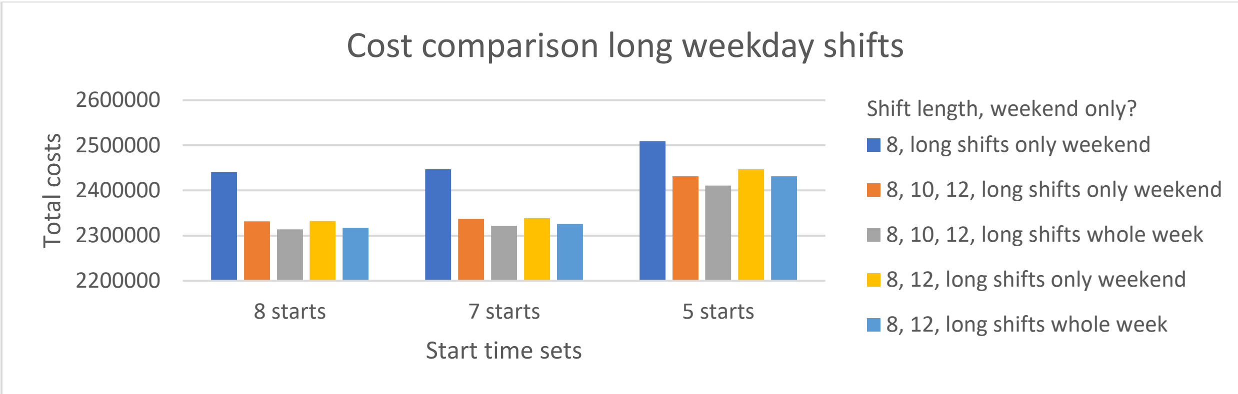

Figure 17: Longer shift weekend experiment outcomes

From these results we can see that allowing shifts longer than 8 hours throughout the whole week has the best results. But only allowing them during the weekends gives less flexibility but still has a big benefit. A non-monetary benefit of only allowing longer shifts during the weekends can be that nurses will be more satisfied with this than also working long shifts during the weekdays.

5.2.11 Service levels

It is interesting to see what happens to the results when choosing different service levels. The service level is the minimum demand that needs to be met by the care nurses. The lower the service level, the bigger the risk off understaffing. In this section we look at what shift schedule is found when applying lower service levels. For this experiment we have also taken out the maximum mertid and overtime percentage constraints, so we can see what happens to the staffing decisions without still limiting planning less nurses because of the allowed overtime and mertid use. All the other settings are the same as the base case. The results can be found in Table 29. Figure 18 shows us how the costs for unused hours and mertid+overtime change depending on the service level.

Table 30: Service level experiment results

INSTANCE	SERVICE LEVEL	STAFFING COSTS	FTE ESTIMATE	MERTID COSTS	OVERTIME COSTS	UNUSED HOUR COSTS	TOTAL COST
1	1	2,719,886	132	0	0	306800	3,026,686
2	0.9	2,403,086	108	13805	4204	113360	2,534,456
3	0.8	2,141,032	98.4	36780	5355	139984	2,323,152
4	0.7	2,029,560	91.2	50527	6090	90272	2,176,449
5	0.6	1,923,501	81.6	70671	7875	416	2,002,463
6	0.5	1,861,475	79.2	79020	8750	1872	1,951,118

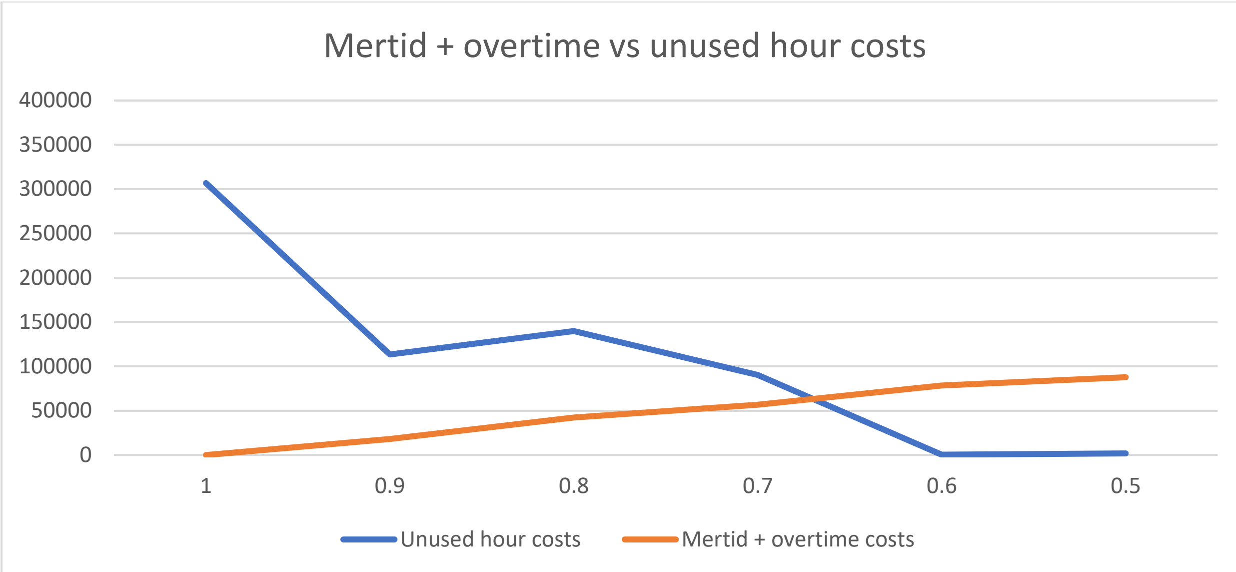

Figure 18: Mertid and overtime vs unused hour costs

The results are logical because by choosing a lower service level, less staff is planned beforehand, and this is compensated for by using mertid and overtime. Mertid and overtime are more costly per hour than planned nurse hours, but they also do not come with a contract. In this case you only pay for the hours you really need, which is beneficial compared to hiring an additional nurse and thereby increasing unused contract hours. The decreasing costs for unused hours are visible when using a lower service level. We see that between a service level of 60% and 70% the mertid and overtime costs start to be higher than the unused hours costs. We also see that the number of FTEs needed declines. In real-life it is not realistic to rely on a lot of mertid and overtime. It is already difficult to find people willing to work extra, this problem will only get bigger when decreasing the service level, as you then plan even less people and are understaffed more often.

To see how outcomes differ per settings we have done the same experiments with different settings for the start times and shift lengths. Figure 19 shows the difference in costs for extra needed staff and the costs for unused hours.

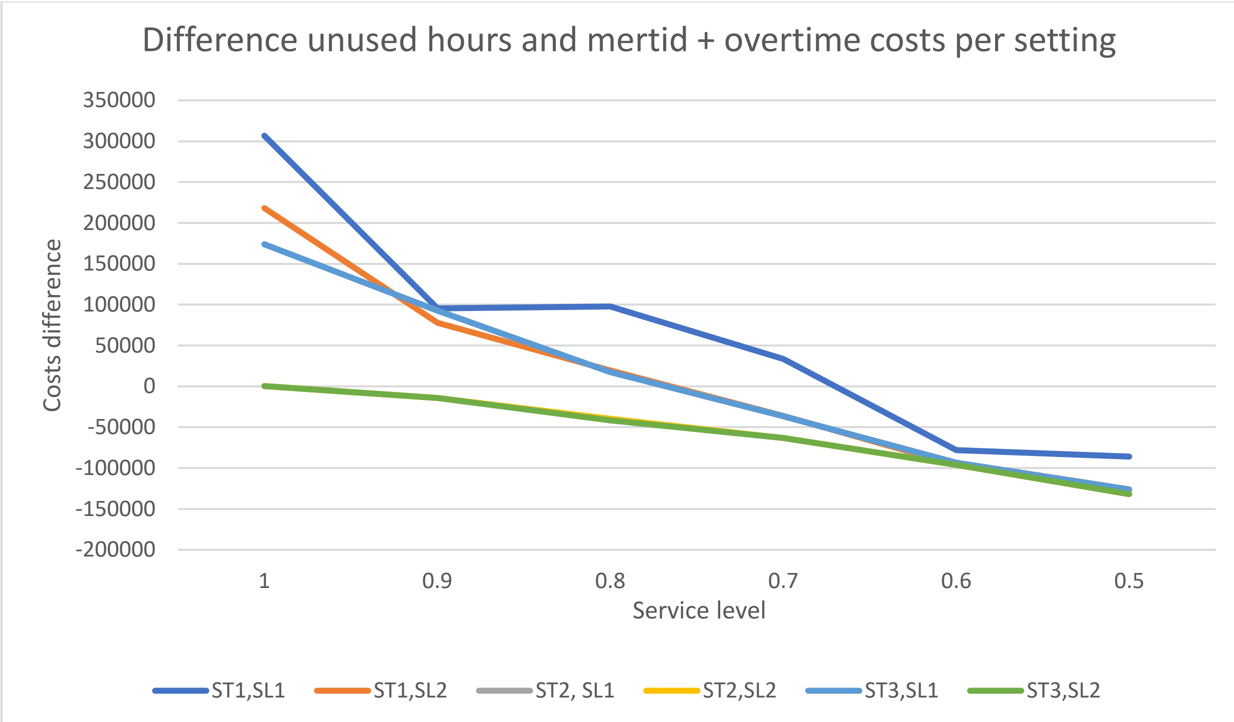

Figure 19: Difference unused hours and mertid + overtime per setting

In Figure 19 we can see that more flexible settings approach the point where the overtime and mertid costs are higher than the unused hour costs more quickly. This makes sense because shift schedules planned with more flexible rules, fit the demand curve better and therefore have less unused hours. From this we can conclude that when working with more flexible settings, a lower contract percentage becomes less beneficial. This is because now there are different ways to tackle the bottleneck of filling the weekend shifts. In Figure 19, the light blue and grey line follow a similar path which results in the grey line no being visible. The same goes for the green and yellow line.

Currently, the downside of a lower service level is only expressed by the cost of the expected understaffing. However, it is important to note that relying too heavily on reactive staffing introduces more risks than are captured by monetary measures. Lower service levels may lead to higher stress for the nurses, higher turnover, reduced continuity of care. This negatively affects patient outcomes. Griffiths et al. (2020) explains that planning staffing at the 90th percentile of demand, comparable to a 90% service level in this study, effectively reduces understaffing without incurring prohibitive costs. This suggests that while lower service levels may appear cost-effective in a model, they may not be the long-term result in the real-world

5.2.12 Performance new data

It is interesting to see how the shift schedules of certain settings perform on a different dataset. To check robustness, we find the optimized shift schedule for each setting and compare this against the demand data of the first months of 2025. We track how often the ED would have been short on staff which needs to be paid for by use of overtime or mertid. The understaffing at a certain hour is the difference between the needed nurses and the planned care nurses, the planned idle time, the planned free flex nurses, and the planned structured flex nurses. We include the flex nurses in this calculation because it is easier to reschedule these tasks to other moments. If the need for care nurses was higher than planned and we use flex task hours for it, these need to be rescheduled. If it is not possible to do this in other idle hours, these will also go into mertid or overtime. The found results are summarized in Table 30 and Figure 20 visualizes how the total costs are split over planning costs and understaffing costs.

Table 31: New data experiment outcomes

INSTANCE	(SHIFT LENGTH), (START TIMES)	AVERGAE SERVICE LEVEL	US CARE + IDLE + TOTAL FLEX	COST OF UNDERSTAFFING	SCHEDULE COST	TOTAL COST
1	ST1, SL1	98.5	193	116,283	2,521,492	2,637,775
2	ST2, SL1	97.8	287	171,484	2,453,589	2,625,074
3	ST3, SL1	97.9	284	169,901	2,447,300	2,617,201
4	ST1, SL2	98	272	164,498	2,437,498	2,601,996
5	ST2, SL2	97.9	279	168,163	2,337,655	2,505,818
6	ST3, SL2	97.8	290	174,496	2,330,332	2,504,828
7	ST1, SL3	97.9	287	172,702	2,419,095	2,591,798
8	ST2, SL3	98.1	249	149,412	2,345,365	2,494,778
9	ST3, SL3	97.8	300	179,795	2,334,174	2,513,969

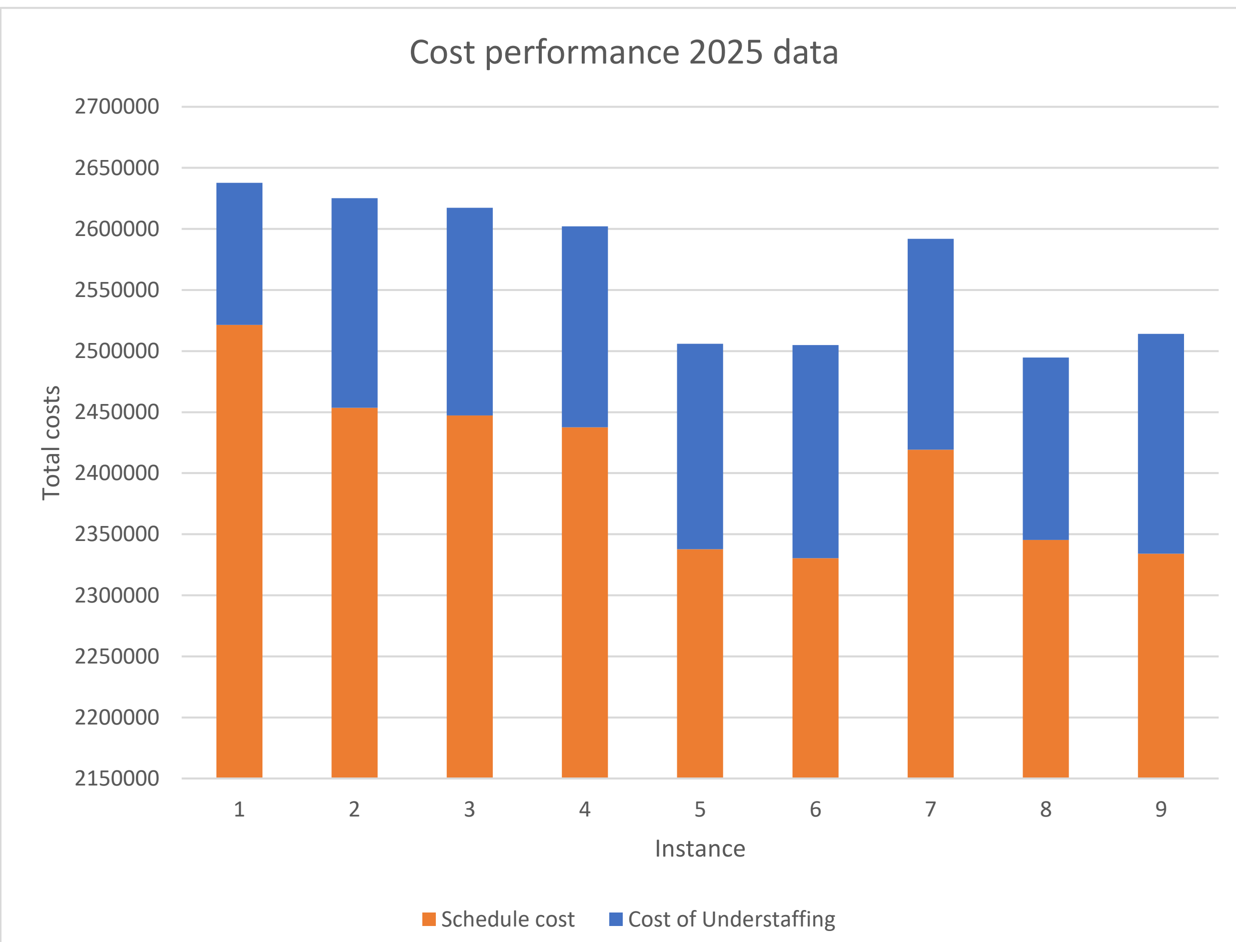

Figure 20: New data experiment outcomes

From this data we can see that the understaffing costs get higher when the schedule has more flexibility. This is logical because now the shifts can be better fitted to the expected demand. This means that there is less idle time planned, which can make up for unexpected increases in the demand. The figure also shows quite clearly that the planning costs of the flexible shifts are much lower and make up for these understaffing costs. The setting with the lowest total cost is instance 8, but the difference between instances 5,6, 8, and 9 are quite small. Therefore, the decision between these instances can be based on preference. Instances 8 and 9 allow three different shift types, which is likely not preferred by the nurses.

5.3Summary and conclusion experiments

The experiments in this chapter have given us many insights. We know that more flexible shift settings results in a shift schedule which better fits the demand curve. This is beneficial for the planning costs but also comes with higher risks of understaffing when using a different demand set. Still, the benefits of the lower scheduling costs are greater than the downside of higher unexpected staffing needs. The bottleneck of the shift scheduling are the weekend shifts. We see in the experiments that more flexibility allows us to fill their weekends shifts more easily. We see that the settings like lower contract percentage make more sense when using fewer flexible settings. The same goes for the service level settings and the impact of having shifts longer than eight hours during the whole week instead of only weekends.

We conclude that allowing 12-hour shifts throughout the whole week combined with 7 start times decreases costs with 7.31% compared to the base case. Only allowing the 12-hour shifts during the weekend still reduces the costs with 6.83%. Allowing 12-hour shifts throughout the whole week combined with 8 start times results in a 7.65% cost improvement, compared to 7.06% cost reduction when only allowing 12-hour shifts in the weekend. For each of these settings the needed FTEs decrease with 7.6%. Besides allowing more flexibility, our experiments show that it could be of interest to allow more weekend shifts by paying an extra premium. Getting 1 person per week to work an extra weekend saves 1.32% of the costs.

6. Advice on implementation

Before the nurse shift scheduling tool can be implemented within the ED, there are several steps that need to be taken. When it is ready to be implemented, there is more needed than only technical deployment. It requires a structured change management approach (Kotter, 1995), sustained stakeholder engagement (Bryson, 2004), and iterative improvement (Jenny et al., 2015). Drawing on the impression of the organization, the demand modelling insights, and the prototype development, this chapter outlines a roadmap to successful use of this research.

Despite robust demand modelling and prototype development, realizing the full potential of the tool depends on five interrelated streams of work: Researching the demand modelling further so it resembles reality, configuring the tool to match the requirements of the ED, making sure the potential users are engaged, improve user experience by using a iterative improvement process, and organizing pilot trials with targeted feedback collection to look for possibilities of applying this research to different departments as well.

Improving demand modelling

The current way of modelling the demand is not in the amount of detail that is needed when using the tool in real-life. This is also discussing in Chapter 7. The data collection of the ED needs to be improved after which the patient classification system can be expanded. When the demand modelling is of good enough quality, the tool can be used to replace current shift scheduling. Depending on the preferences of the ED, the tool can be used on itself to determine the shift schedules, but it is also possible to first investigate integrating it with the nurse rostering process.

Aligning settings with department preferences

Building on the experimental analyses, the project team can now select the preferred rule sets and scheduling parameters most appropriate for the ED. These include shift lengths, start times, and service-level targets. During this phase, it is crucial to refine the tool's usability by simplifying input screens and clarifying configuration options. Enhancing the interface at this stage will make it easier for stakeholders to test scenarios and provide meaningful feedback, which accelerates the customization of the tool.

Engaging emergency department stakeholders

Simultaneously, we must create a sense of ownership among the clinicians who will rely on the tool daily. A series of workshops to which nurse managers, the persons responsible for planning, physicians, and quality officers are invited to review drafts of the user interface and to experience the underlying logic. Rather than presenting a polished product, these sessions encourage stakeholders to voice concerns and suggest modifications. The list of enhancements that emerges from each workshop ensures that following development cycles are aligned with ED practice, making sure the department is involved which minimizes the resistance to change.

Refining usability through iterative development

Once the core configuration is established and stakeholder opinions are collected, the team transitions to usability testing with a functional prototype. Over three cycles, the people responsible for planning complete scheduling tasks while observers record task completion times, navigation errors, and usability ratings. These mixed-method evaluations show friction points such as non-intuitive use of words or workflows. Following each cycle, the development team implements the spotted improvements. This ensures that the interface evolves in direct response to user needs, resulting in an intuitive and efficient tool.

Embedding change and scaling up

In the final phase, Kotter's emphasis on consolidating gains and anchoring new approaches guides the transition from pilot to routine practice (Kotter, 1995). Pilot successes, such as a measurable reduction in scheduling time, are shared in the department to create a positive attitude towards the change. Detailed training materials, like video tutorials and concise quick-reference guides, are consolidated in an online repository to lower adoption barriers. Governance of usage metrics and configuration maintenance is handed over to the IT department, embedding oversight into existing operational structures. This balanced approach maintains consistency with the established methodology while allowing users to tailor the tool to its situation (Hiatt, 2006).

7. Conclusions & recommendations

In this chapter, we draw together the principal findings of this thesis and reflect on how they address the original research questions: *"How can we develop a tool that can make the shift scheduling process more efficient and that can optimize the shift schedules in the emergency department at St. Olav's Hospital?"* We start with summarizing the conclusions about the experiments and the performance of the tool in Section 7.1. Then, we continue with recommendations to St. Olav's ED about their shift scheduling process in Section 7.2. After this, we discuss the contributions this research has resulted in. In Section 7.4 we discuss the limitation to this research. We finish with the recommendations for further research in Section 7.5.

7.1Conclusions

Through a two-stage framework, percentile-based demand modelling followed by mixed-integer programming for shift assignment, we have demonstrated that allowing different settings regarding scheduling shifts can positively impact the efficiency of the schedule.

First, we showed that transforming historical arrival data into hourly nurse-demand percentiles at a chosen service level finds nurse workload. By linking direct-care coverage to these percentiles, we link risk-tolerance and staffing levels.

Second, our shift-optimization model successfully translates those hourly targets into a weekly shift schedule. With the tool it is possible to include real-world features like flexible start times, multiple shift lengths and contract-mix caps. The two-stage solution approach keeps the model both fast and interpretable. The tool can be used to compare different settings for shift scheduling. This enables the ED to easily consider different scheduling rules. Besides, the ED can easily use this tool to find an optimal shift schedule when the preferred settings are clear.

The managerial experiments revealed actionable insights for St. Olav's ED:

- Allowing mixes of 8 hour and 12-hour shifts combined with seven start times saves 7.31% of the costs, while for eight start times this is 7.65%.
- Paying extra to create extra weekend capacity can be well worth the investment. One extra weekend shift saves 1.32% of the total costs.
- We provided data which allowed the ED to make educated decisions on whether to allow longer shifts only in the weekend or throughout the whole week. Looking at the settings which allows 8- and 12-hour shift with seven start times the cost improvement goes from 7.31% to 6.83% when allowing long shifts only in the weekend. For the settings with eight start times the cost improvement goes from 7.65% to 7.06%.
- Choosing to allow more start times and shift lengths, either only in the weekend or throughout the whole week, results in a needed FTE decrease of 7.6%.
- A lower average contract percentage helps fill weekend gaps, but this effect fades when the settings for the schedule get more flexible. The data shows that with independent of which settings are chosen, the average contract percentage should not be above 90%.

Together, these findings confirm that a data-driven, optimization-backed scheduling tool can deliver better care coverage with more efficient staffing. Our research provides the ED with valuable data to support decision-making regarding the shift scheduling. Once these decisions are made, we have provided a prototype tool that can be used to find the optimal shift schedule.

7.2Recommendations

To translate our findings into improved nurse scheduling and cost-efficiency, we recommend the following actions:

- Base rostered capacity on the 90% demand percentile. This ensures that service-level targets are met in most hours while avoiding the excessive overstaffing that comes from planning to the absolute peak.
- Complement standard 8-hour shifts with 12-hour shifts for the weekends.
- Increase the number of shifts start options to the eight start times.
- Begin systematically collecting additional patient characteristics and treatment-duration data so it is possible to improve the patient classification system and make the demand modelling more realistic.
- Look into whether nurses are willing to work extra weekends in return for an extra premium.
- Develop the tool further, so its use becomes more beneficial.
- Use an average contract percentage of 90% or lower.

7.3Contributions

This thesis adds to both the theory and practice of nurse shift scheduling in high-variability clinical settings through three contributions.

4) We introduce a two-stage framework that translates historical arrival patterns into percentile-based staffing targets and then applies mixed-integer programming to assign shifts. By combining demand forecasting with optimization, our approach extends the shift scheduling models and offers a template adaptable to other situations.
5) We demonstrate that certain combinations of shift lengths, specifically mixes of 8- and 12-hour blocks with staggered start times, outperform traditional fixed-length shift schedules. Experiments show these flexible configurations yield the service-level coverage with lower total nurse-hours than standard 8-hour schedules, underlining the efficiency gains achievable by tailoring shift settings to demand patterns.
6) We deliver a functional prototype scheduling tool, which can be built further on when this research is continued. Currently, it can already be used to compare what different settings do with the efficiency of the shift schedule.

7.4Limitations and further research

While our two-stage scheduling framework delivers clear operational benefits, several limitations and suggestions for future work remain. First, the underlying demand model assumes uniform treatment durations and service-time ratios across all patient arrivals. This is an oversimplification that is too simple to represent true variability in care pathways. It is interesting to further research ways of letting the demand modelling better represent reality. When doing this, the model could become more valuable for planning the actual shift schedule, while now the main value is comparing the shift settings.

Besides the limitations regarding the data quality, it is also good to discuss the limitations regarding the amount of data. Currently, the tool bases its schedule on data of 2024. Literature explains that predicting ED demand can be done based on historical demand. Although the quality would be better when it is based on more data than a single year. Therefore, it would be interesting to keep gathering data and use this as input to increase the reliability and see what multiple years of data does to the results. Also, when there is more data available, it would be interesting to test different forecasting techniques. This could for example show that there is an increasing trend in number of patients which is not taken into account when assuming similar demand for subsequent years.

On the staffing side, our objective focuses almost exclusively on monetary costs and service-level targets, without accounting for individual nurse preferences, shift-fairness, or fatigue. It would be interesting if in the future soft preferences and fairness metrics can be included in the MIP to produce schedules that also take nurse experience into account.

Because we use a fixed week design, we do not take seasonal factors into account. It would be interesting to research how demand expectations are different from week to week and automatically adapt the schedule based on this. This would allow the tool to recommend different schedules per week or season. Likewise, exploring a similar setting with a granularity smaller than hours might show cost or coverage-efficiency gains, though at the expense of larger model size.

This research has been focusing on the shift schedule. The next step is assigning specific nurses to these shifts, nurse rostering. It might be interesting to include this next step into the model. This way it is possible to use more specific constraints which yields a more accurate estimation of staffing needs in practice.

One of the recommendations is to implement the use of 12-hour shifts. This thesis does not explore potential non-monetary downsides, such as fatigue and productivity, which is important for future research.

Finally, as length-of-stay reductions can lower required staffing and the length of stay is highly affected by the connection between the ED and other departments, it is interesting to investigate how to connect the expectations for the ED to the schedules of the other departments in the hospital.

8. Bibliography

Aickelin, U., Burke, E. K., & Li, J. (2009). An evolutionary squeaky wheel optimization approach to personnel scheduling. *IEEE Transactions on Evolutionary Computation*, *13*(2). https://doi.org/10.1109/TEVC.2008.2004262

Aickelin, U., & Dowsland, K. A. (2004). An indirect genetic algorithm for a nurse-scheduling problem. *Computers and Operations Research*, *31*(5). https://doi.org/10.1016/S0305-0548(03)00034-0

Ajmi, I., Zgaya, H., Gammoudi, L., Hammadi, S., Martinot, A., Beuscart, R., & Renard, J. M. (2015). Mapping patient path in the Pediatric Emergency Department: A workflow model driven approach. *Journal of Biomedical Informatics*, *54*. https://doi.org/10.1016/j.jbi.2014.12.004

Akkermans, A., Post, G., & Uetz, M. (2021). Solving the shift and break design problem using integer linear programming. *Annals of Operations Research*, *302*(2). https://doi.org/10.1007/s10479-019-03487-6

Aschan, H., Junttila, K., Fagerstrôm, L., Kanerva, A., & Rauhala, A. (2009). RAFAELA™ Patient Classification System as a tool for management. *Studies in Health Technology and Informatics*, *146*. https://doi.org/10.3233/978-1-60750-024-7-478

Awadallah, M. A., Khader, A. T., Al-Betar, M. A., & Bolaji, A. L. aro. (2011). Nurse rostering using modified harmony search algorithm. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7077 LNCS*(PART 2). https://doi.org/10.1007/978-3-642-27242-4_4

Azaiez, M. N., & Al Sharif, S. S. (2005). A 0-1 goal programming model for nurse scheduling. *Computers and Operations Research*, *32*(3). https://doi.org/10.1016/S0305-0548(03)00249-1

Bellanti, F., Carello, G., Della Croce, F., & Tadei, R. (2004). A greedy-based neighborhood search approach to a nurse rostering problem. *European Journal of Operational Research*, *153*(1). https://doi.org/10.1016/S0377-2217(03)00096-1

Beswick, S., Hill, P. D., & Anderson, M. A. (2010). Comparison of nurse workload approaches. *Journal of Nursing Management*, *18*(5). https://doi.org/10.1111/j.1365-2834.2010.01124.x

Bilgin, B., De Causmaecker, P., Rossie, B., & Vanden Berghe, G. (2012). Local search neighbourhoods for dealing with a novel nurse rostering model. *Annals of Operations Research*, *194*(1). https://doi.org/10.1007/s10479-010-0804-0

Bryson, J. M. (2004). Why strategic planning is more important than ever. *Strategic Planning for Public and Nonprofit Organizations*.

Burke, E. K., De Causmaecker, P., Petrovic, S., & Berghe, G. Vanden. (2006). Metaheuristics for handling time interval coverage constraints in nurse scheduling. *Applied Artificial Intelligence*, *20*(9). https://doi.org/10.1080/08839510600903841

Cardoen, B., Demeulemeester, E., & Beliën, J. (2010). Operating room planning and scheduling: A literature review. *European Journal of Operational Research*, *201*(3). https://doi.org/10.1016/j.ejor.2009.04.011

Clopton, E. L., & Hyrkäs, E. K. (2020). Modeling emergency department nursing workload in real time: An exploratory study. *International Emergency Nursing*, *48*. https://doi.org/10.1016/j.ienj.2019.100793

Coffman, J. M., Seago, J. A., & Spetz, J. (2002). Minimum nurse-to-patient ratios in acute care hospitals in California. *Health Affairs*, *21*(5). https://doi.org/10.1377/hlthaff.21.5.53

Daniel Wright, P., Bretthauer, K. M., & Côté, M. J. (2006). Reexamining the nurse scheduling problem: Staffing ratios and nursing shortages. In *Decision Sciences* (Vol. 37, Issue 1). https://doi.org/10.1111/j.1540-5414.2006.00109.x

Dash, C. S. K., Behera, A. K., Dehuri, S., & Ghosh, A. (2023). An outliers detection and elimination framework in classification task of data mining. *Decision Analytics Journal*, *6*. https://doi.org/10.1016/j.dajour.2023.100164

de Grano, M. L., Medeiros, D. J., & Eitel, D. (2009). Accommodating individual preferences in nurse scheduling via auctions and optimization. *Health Care Management Science*, *12*(3). https://doi.org/10.1007/s10729-008-9087-2

Di Gaspero, L., Gärtner, J., Kortsarz, G., Musliu, N., Schaerf, A., & Slany, W. (2007). The minimum shift design problem. *Annals of Operations Research*, *155*(1). https://doi.org/10.1007/s10479-007-0221-1

Duma, D., & Aringhieri, R. (2020). An ad hoc process mining approach to discover patient paths of an Emergency Department. *Flexible Services and Manufacturing Journal*, *32*(1). https://doi.org/10.1007/s10696-018-9330-1

Ernst, A. T., Jiang, H., Krishnamoorthy, M., & Sier, D. (2004). Staff scheduling and rostering: A review of applications, methods and models. *European Journal of Operational Research*, *153*(1). https://doi.org/10.1016/S0377-2217(03)00095-X

Gräff, I., Goldschmidt, B., Glien, P., Klockner, S., Erdfelder, F., Schiefer, J. L., & Grigutsch, D. (2016). Nurse staffing calculation in the emergency department - Performance-oriented calculation based on the Manchester Triage System at the University Hospital Bonn. *PLoS ONE*, *11*(5). https://doi.org/10.1371/journal.pone.0154344

Griffiths, P., Saville, C., Ball, J., Jones, J., Pattison, N., & Monks, T. (2020). Nursing workload, nurse staffing methodologies and tools: A systematic scoping review and discussion. *International Journal of Nursing Studies*, *103*. https://doi.org/10.1016/j.ijnurstu.2019.103487

Gutjahr, W. J., & Rauner, M. S. (2007). An ACO algorithm for a dynamic regional nurse-scheduling problem in Austria. *Computers and Operations Research*, *34*(3). https://doi.org/10.1016/j.cor.2005.03.018

Hiatt, J. (2006). ADKAR: A model for change in business, government and our community. In *Model for change in business, government, and our community*.

Hulshof, P. J. H., Kortbeek, N., Boucherie, R. J., Hans, E. W., & Bakker, P. J. M. (2012). Taxonomic classification of planning decisions in health care: a structured review of the state of the art in OR/MS. In *Health Systems* (Vol. 1, Issue 2). https://doi.org/10.1057/hs.2012.18

Hurst, Keith. (2003). *Selecting and applying methods for estimating the size and mix of nursing teams: a systematic review of the literature: summary*. Nuffield Institute for Health.

Iordache, S., Elseviers, M., De Cock, R., & Van Rompaey, B. (2020). Development and validation of an assessment tool for nursing workload in emergency departments. *Journal of Clinical Nursing*, *29*(5–6). https://doi.org/10.1111/jocn.15106

Jenny, P., Rogers, Yvonne, & Helen, S. (2015). Interaction Design: Beyond human-computer interaction (4th ed). In *John Wiley & Sons Ltd* (Vol. 176).

Karaca, Z., Wong, H. S., & Mutter, R. L. (2012). Duration of patients' visits to the hospital emergency department. *BMC Emergency Medicine*, *12*. https://doi.org/10.1186/1471-227X-12-15

Kotter. (1995). Leading Change; Kotter. (ثق ثقثقثق)*73* ,⍰⍰⍰⍰⍰⍰.

Li, J., Burke, E. K., & Qu, R. (2012). A pattern recognition based intelligent search method and two assignment problem case studies. *Applied Intelligence*, *36*(2). https://doi.org/10.1007/s10489-010-0270-z

Maier-Rothe, C., & Wolfe, H. B. (1973). Cyclical scheduling and allocation of nursing staff. *Socio-Economic Planning Sciences*, *7*(5). https://doi.org/10.1016/0038-0121(73)90043-8

Ozcan, Y. A. (2009). Quantitative methods in health care management: techniques and applications. In *Technometrics*.

Özean, E. (2005). Memetic algorithms for nurse rostering. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *3733 LNCS*. https://doi.org/10.1007/11569596_51

Parr, D., & Thompson, J. M. (2007). Solving the multi-objective nurse scheduling problem with a weighted cost function. *Annals of Operations Research*, *155*(1). https://doi.org/10.1007/s10479-007-0202-4

Purnomo, H. W., & Bard, J. F. (2007). Cyclic preference scheduling for nurses using branch and price. *Naval Research Logistics*, *54*(2). https://doi.org/10.1002/nav.20201

Ruffing, T., Danko, T., Weiss, C., Winkler, H., & Muhm, M. (2014). Behandlungsdauer orthopädisch-unfallchirurgischer Patienten in der Notaufnahme. *Der Chirurg*, *85*(10). https://doi.org/10.1007/s00104-014-2739-5

Sadjadi, S. J., Soltani, R., Izadkhah, M., Saberian, F., & Darayi, M. (2011). A new nonlinear stochastic staff scheduling model. *Scientia Iranica*, *18*(3 E). https://doi.org/10.1016/j.scient.2011.05.017

Smet, P., Bilgin, B., De Causmaecker, P., & Vanden Berghe, G. (2014). Modelling and evaluation issues in nurse rostering. *Annals of Operations Research*, *218*(1). https://doi.org/10.1007/s10479-012-1116-3

St. Olav's Hospital. (2024, December 17). *Vision, values, and goals - St. Olav's Hospital* .

Topaloglu, S., & Ozkarahan, I. (2004). An implicit goal programming model for the tour scheduling problem considering the employee work preferences. *Annals of Operations Research*, *128*(1–4). https://doi.org/10.1023/B:ANOR.0000019102.68222.df

Van Den Bergh, J., Beliën, J., De Bruecker, P., Demeulemeester, E., & De Boeck, L. (2013). Personnel scheduling: A literature review. *European Journal of Operational Research*, *226*(3). https://doi.org/10.1016/j.ejor.2012.11.029

Warner, M. (2006). *Patient flow: reducing the delay in healthcare delivery* (R. W. Hall, Ed.). Springer.

Williams, S., & Crouch, R. (2006). Emergency department patient classification systems: A systematic review. In *Accident and Emergency Nursing* (Vol. 14, Issue 3). https://doi.org/10.1016/j.aaen.2006.06.002

Wundavalli, L. T., Kumar, P., & Dutta, S. (2019). Workload Indicators of Staffing Need as a tool to determine nurse staffing for a high volume academic Emergency Department: An observational study. *International Emergency Nursing*, *46*. https://doi.org/10.1016/j.ienj.2019.06.003

Yilmaz, E. (2012). A mathematical programming model for scheduling of nurses' labor shifts. *Journal of Medical Systems*, *36*(2). https://doi.org/10.1007/s10916-010-9494-z

Appendix

A. Shift scheduling tool

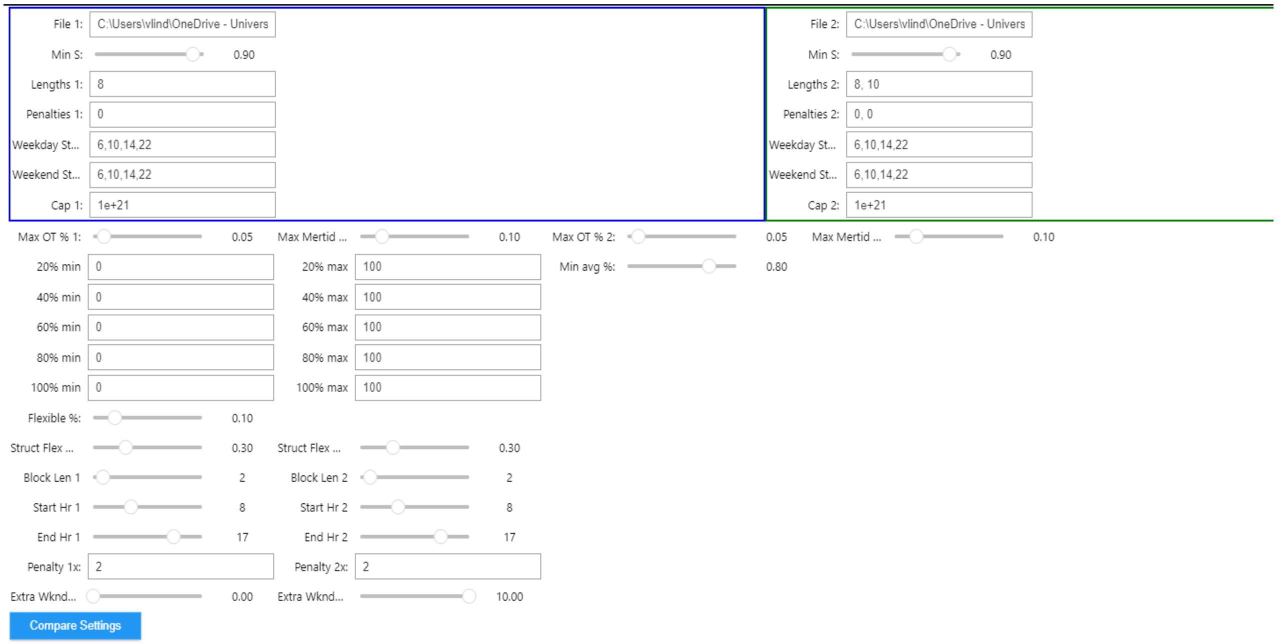

Figure 21: Input interface

Setting 1

Total Cost: 2509921.76

Staff Cost: 1839294.60

Mertid: 6943.86 (12.3h, 4.5% of 272.0h)

Overtime: 1009.17 (1.4h, 1.1% of 136.0h)

Shift Penalties: 0.00

FTEs: 100.80

Weekend Shifts: 84(100.0% of allowed)

Total Available Staff Hours: 4032.0

Unused Contract Hours: 1312.0 (32.5%)

Total Used Staff Hours: 2960.0

Structured Flex Blocks: 25

Structured Flex Hours: 72.0

Expected Understaffing: 13.72h

Staff per Contract

20%: -0

40%: -0

60%: 13

80%: 100

100%: 13

DIFFERENCE (2 - 1)

Total Obj: -126949.75 (-5.1%) ✅

Total Staff Costs: -70603.00 (-3.8%) ✅

Total Mertid: 2405.13 (+34.6%) ❌

Total Overtime: -116.67 (-11.6%) ✅

Fte: -8.00 (-7.9%) ✅

Setting 2

Total Cost: 2382972.01

Staff Cost: 1768691.60

Mertid: 9348.99 (16.6h, 6.4% of 260.6h)

Overtime: 892.50 (1.3h, 1.0% of 130.3h)

Shift Penalties: 0.00

FTEs: 92.80

Weekend Shifts: 84(100.0% of allowed)

Total Available Staff Hours: 3712.0

Unused Contract Hours: 1106.0 (29.8%)

Total Used Staff Hours: 2846.0

Structured Flex Blocks: 22

Structured Flex Hours: 72.0

Expected Understaffing: 17.88h

Staff per Contract

20%: 1

40%: 4

60%: -0

80%: 100

100%: 11

Figure 22: Output measures shift scheduling tool

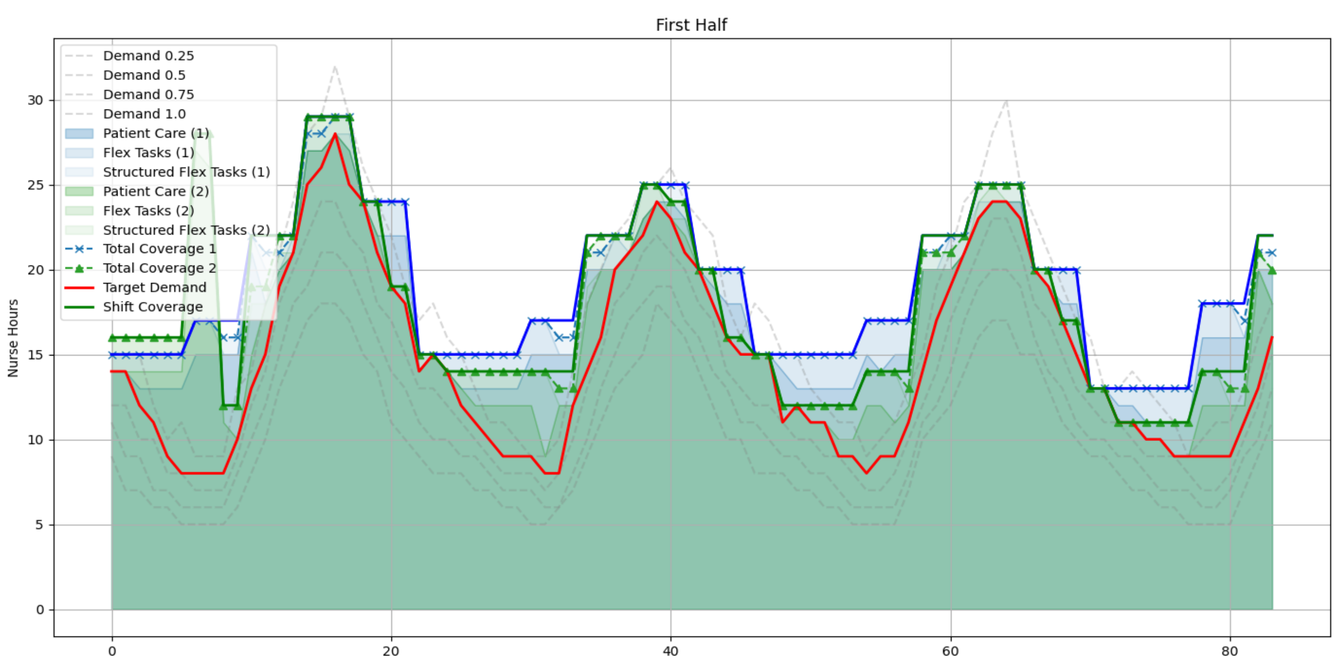

Figure 23: Planned manpower and demand curve shift scheduling tool