A Two-Stage Nurse Scheduling Approach for Residential Care Organizations in the Netherlands

K. de Greef

Industrial Engineering and Management, University of Twente

Faculty of Behavioural, Management and Social Sciences

Supervisors from University of Twente
dr. D. Guericke
dr. ir. A.G. Leeftink

Supervisor from Nedap Healthcare
R. Zaagman

June 16, 2023

Management summary

Introduction: Dutch residential care organizations are facing a shortage of nursing staff, while the demand for elderly care will rise due to the ageing population. This puts pressure on care organizations and their employees, resulting in job dissatisfaction and higher turnover rates. Research has shown that employee satisfaction can be enhanced by high-quality schedules and scheduling processes, which is referred to as the Nurse Scheduling Problem (NSP). This process should consider nurses' preferences, promote work-life balance, incorporate fairness aspects, meet coverage requirements, and comply with labor legislations and organizational standards. Despite the complexity, these schedules are still manually created by the planner, which is a time-consuming task and subjective to fairness. Therefore, the goal of this research is to *develop an automated scheduling method that meets the requirements of residential care organizations to enhance nurses' job satisfaction and fairness.*

This research has been conducted at Nedap Healthcare, where they have developed a software suite that enables planners to assign shifts to nurses manually. Nedap can use the results of this research to extend their software further and better support their customers.

Methods: In this research, a two-stage scheduling method is developed. To define the constraints, eight care organizations are interviewed, and a comprehensive literature review is conducted on previous approaches to solving the NSP and fairness-enhanced scheduling.

In the first stage, a tactical schedule is constructed, which focuses on preferences and allows for a better work-life balance by repeating the schedule over multiple periods. In the second stage, the operational schedule is generated, which focuses on meeting the periodic coverage requirements and incorporates planned absenteeism. To ensure predictability, a percentage of the tactical schedule, referred to as the *flexibility parameter*, is retained in the operational schedule.

A simulated annealing (SA) algorithm is used to optimize both schedules. The objective is to minimize the penalties resulting from violations of time- and organizational-related soft constraints (TRCs and ORCs) while evenly distributing the encountered penalties among the nurses. For the tactical schedule, an initial feasible solution is used as input for the SA, which is generated using a constructive heuristic that satisfies all hard constraints. Subsequently, the periodically planned absenteeism is removed from the tactical schedule, and the remaining is used as input to generate the operational schedule.

Parameter tuning on the tactical schedule is performed to determine the values for the SA algorithm, and the weights for the soft constraints are determined based on the preferences that result from the interviews.

Results: To assess the performance of the method, we compare the results with the current practice from three case studies, followed by a flexibility analysis and, at last, a sensitivity analysis. Al results showed a trade-off between meeting more coverage requirements or minimizing the increase in TRC violations. The proposed method results in a lower objective for two out of three case studies. It decreases the operational objective by 96%-98% and 53%-65% for the small and medium case study, respectively. This results from the decrease in TRC and ORC violations. The results have shown that the method can reduce the number of tactical and operational TRC violations by 79%-82% and 17-76%, respectively, compared to the current performance of the small case study. The number of TRC violations for the other two case studies decreased for the best operational solution by 1% and 23% for the medium and large case studies, where the number of TRC violations decreased by 22% in the best solution for the large case study. Noteworthy is that except for a single violation in the medium case study, the constraint for CWD and CNS is not violated. Additionally, the method reduces the number of ORC violations in the best tactical schedule for the small and medium case study by 76% and 65% compared to the current performance, resulting from a reduction in undercoverage during the week and weekend. Regarding the ORC violations of the operational schedule, the total of UQ shifts can be reduced by 47%-53% for the medium case study compared to the current performance while resulting in valid schedules. Unfortunately, the method did not find valid operational schedules for the large case study as there are still remaining open shifts. This is mainly due to the small ratio of available minutes and demand in minutes, reducing the flexibility in shift assignments while minimizing violations of the constraints.

Additionally, a flexibility analysis is performed to assess the influence of the flexibility parameter on the outcome of the operational schedule in terms of fairness and the number of violations. A flexibility parameter

of 0.0 resulted in the fairest and best schedules for the small and medium case studies. It resulted in the smallest increase in TRC violations, being 25% and 20% for the small and medium case study, respectively. However, the outcomes are less stable regarding TRC violations, reflected by the varied outcome for the objective value. In contrast, a flexibility parameter of 0.4 resulted in the best objective value for the large case study, but none of the flexibility experiments resulted in valid schedules. The results showed that the largest decrease in ORC violations results when using a flexibility parameter of 0.2, being 93% for the medium case study and 62% for the large case study, and a flexibility of 0.4 for the small case study with a decrease of 90%. The results showed that resolving more undercoverage comes at the cost of an increase in TRC violations and UQ shifts, as the goal is to meet demand. As indicated, the fairest schedule is obtained using a flexibility parameter of 0.0, among others, due to the fact that the RM is distributed more fairly, indicated by the smaller variation.

At last, a sensitivity analysis was performed on the weights assigned to the TRC and the flexibility parameter. The results showed that increasing w_{FRO} both reduces the number of FRO and RT violations, whereas increasing w_{RT} resulted in more FRO violations. Increasing w_{RM} resulted in less undercoverage in both tactical and operational schedules but came at the cost of more TRC violations and increased UQ shifts. The results showed that to obtain stable solutions, the weight of the flexibility parameter should be equal to or higher than ten.

Combined with the results of current performance, the flexibility analysis, and sensitivity analyses, we can conclude that the method can generate valid operational schedules without open shifts for the small and medium case studies. This does come at the cost of an increase in TRC violations and UQ shifts. Unfortunately, the method did not generate a valid schedule for the large case study, as the remainder of the shifts were not assigned. Noteworthy is that the method does not violate the constraint of CWD and CNS in most cases, with a single exception for a few experiments.

Conclusion and Discussion: The proposed method has the potential to support residential care organizations in generating tactical and operational schedules in a short amount of time. While it does not outperform current practice in terms of all TRC and ORC violations, it provides insight into the effect of the flexibility parameter on the outcome of the operational schedule and which can be implemented to provide nurses with a more predictable and fair schedule in practice. Allowing no flexibility resulted in the fairest schedules for the small and medium case study. However, allowing flexibility provides opportunities to decrease the number of ORC violations. There were several limitations to the study due to a lack of data and assumptions and simplifications that had to be made. These limitations and the parameter values chosen affect the quality and outcome of the schedule. The performance of the method is evaluated using three case studies. However, it is unknown what the priorities and individual agreements were when the schedules were created in practice. Therefore, the number of TRC violations can be misclassified in the current performance. Additionally, in practice, the priorities and preferences differ per organization and should be considered by fine-tuning the weights and performing a new parameter-tuning process. In this method, the assumption was made that the priorities were equal for all three case studies. Therefore, future research is needed to validate the performance in real-world settings, determine the parameter values for the operational schedule, and identify the appropriate weights for the soft constraints before implementing the proposed method in practice. Finally, by considering the suggested recommendations, Nedap can use the findings of this research to support their customers through an automated scheduling process that includes fairness and promotes a better work-life balance, enhancing nurses' job satisfaction.

Contents

1 Introduction **6**

1.1 Research background . . . 8

1.1.1 Nurse scheduling process . . . 8

1.1.2 Scheduling methods . . . 9

1.2 The scope of this research . . . 11

1.3 Research framework . . . 12

2 Nurse Scheduling for Residential Care in Practice **14**

2.1 Terminology . . . 14

2.2 Scheduling in practice . . . 15

2.2.1 General scheduling process . . . 16

2.2.2 Sequence shift scheduling . . . 17

2.3 Objectives and Constraints . . . 17

2.3.1 Hard and Soft constraints . . . 18

2.4 Conclusion . . . 19

3 Literature Review **20**

3.1 Fairness in nurse scheduling . . . 20

3.1.1 Quantification of fairness . . . 21

3.2 Approaches to incorporate fairness . . . 22

3.2.1 Fairness as constraint . . . 22

3.2.2 Fairness as objective . . . 23

3.3 Contribution of this research . . . 24

3.4 Conclusion . . . 26

4 Proposed Method **28**

4.1 Hard and Soft constraints . . . 29

4.1.1 Hard constraints . . . 29

4.1.2 Soft Constraints . . . 30

4.1.3 Additional constraints operational schedule . . . 31

4.1.4 Objective function . . . 33

4.2 A two-staged scheduling approach . . . 34

4.2.1 Simulated annealing . . . 35

4.2.2 The initial feasible solution . . . 36

4.2.3 From tactical to operational schedule . . . 38

4.2.4 Neighbourhood structures . . . 38

4.2.5 Probability operators . . . 40

4.3 Quantification of the performance . . . 40

4.4 Assumptions for the method . . . 41

4.5 Normalization of the penalties . . . 42

4.6 Conclusion . . . 43

5 Case studies **45**
5.1 Data analysis and Assumptions . . . 45
5.2 Description case studies . . . 46
5.3 Current performance . . . 49
5.4 Conclusion . . . 50

6 Model Performance **51**
6.1 Parameter tuning simulated annealing . . . 52
6.2 Trade-off penalty nurses and organization . . . 53
6.3 Experimental results current performance . . . 55
6.4 Flexibility analysis . . . 68
6.5 Sensitivity analysis . . . 79
6.5.1 Sensitivity analysis weights TRCs . . . 79
6.5.2 Sensitivity analysis weights flexibility parameter . . . 83
6.6 Conclusion . . . 84

7 Conclusion and Discussion **86**
7.1 Conclusion . . . 86
7.1.1 Scientific contribution . . . 88
7.1.2 Practical contribution . . . 88
7.2 Discussion . . . 89
7.2.1 Discussion results . . . 89
7.2.2 Limitations to this research . . . 92
7.2.3 Lack of data . . . 92
7.2.4 Limitations of the method . . . 93
7.3 Recommendations . . . 94
7.4 Future research . . . 96

A Calculation Soft Constraints **105**
A.1 Calculations time-related soft constraints . . . 105
A.2 Calculations organisational related soft constraints . . . 109

B Algorithms **111**

C Parameter tuning SA **112**
C.1 Results Parameter Tuning . . . 112
C.1.1 Outcomes parameter tuning Small case study . . . 114
C.1.2 Outcomes parameter tuning Medium case study . . . 115
C.1.3 Outcomes parameter tuning Large case study . . . 117

D Experimental results case studies **119**
D.1 Experimental results small case study . . . 119
D.2 Experimental results medium case study . . . 121
D.3 Experimental results large case study . . . 123

List of terms

CNS Consecutive night shifts

CWD Consecutive working days

EOW Every other weekend

FRO Forward rotating order

NS Night shift

ORC Organization-related constraint

QLs Qualification levels

RM Remaining minutes

RT Rest time

SFS shift on - off - on

SA Simulated annealing

TRC Time-related constraint

UQ Underqualification

WS Weekend shift

Chapter 1

Introduction

Due to the ageing population in the Netherlands [1, 2], the need for elderly care will rise. This will lead to an increased demand for nursing staff, as they form the largest group in the health workforce [3]. However, the Dutch healthcare industry is facing the problem of a shortage of skilled professionals [4]. For residential care, this is expected to rise to a shortage of 51.900 care workers in 2031 [4]. Heydrich et al. [5] mention that one of the repeated reasons for this shortage is the lack of attractiveness of the nursing profession. Nurses must deal with a high workload, lack of autonomy, many regulations, and little appreciation [6]. This leads to less time for the patients, inefficient work, job dissatisfaction, and high turnover, and impacts nurses' mental and physical well-being. This affects the quality of care, as it depends on the quality and motivation of the employees. Therefore, Maenhout and Vanhoucke [7] states that the organizational support of employees should be addressed, which can be achieved through care organizations scheduling policies and processes.

Previous research has shown that the scheduling process and the quality of the constructed schedule are aspects that influence employee satisfaction [5, 7–11]. To create a schedule, nurses need to be assigned to shifts. In literature, this is referred to as the Nursing Scheduling Problem (NSP), which is an extensively studied subject [5, 8, 12–14]. For in-depth reviews on personnel scheduling, we refer to den Bergh et al. [15] and nurse scheduling to Burke et al. [16] and Ngoo et al. [8].

Ejebu et al. [3] show that shift patterns are often organized in ways that harm nurses' health and well-being, their job performances, and the care they provide to patients, all impacting employee satisfaction. Therefore, it is important to construct high-quality schedules to increase job satisfaction. A way to achieve this in a schedule is by addressing employees' preferences, work-life balance, and sustaining autonomy [5]. Additionally, Wolbeck and Kliewer [17] and Uhde et al. [18] show that incorporating fairness within the scheduling process contributes to enhanced job satisfaction. However, it is challenging to create a high-quality schedule that covers all these requirements while meeting coverage requirements and complying with labor legislation and organizational standards.

The high-quality schedule allows nurses to meet their personal family needs while also satisfying the requirements for delivering qualitative care [19]. Perfection is however unlikely, resulting in that, in practice, the delivery of quality patient care will be prioritized over meeting personal needs. Therefore, it is important when schedules are generated that nurses feel treated fairly and are satisfied with the process. Nelson and Tarpey [19] mention that the perception of being treated fairly is referred to as *organizational justice*. This can be divided into *distributive justice* and *procedural justice*, with the first being what people get and the latter how it was given to them [19]. Nelson and Tarpey [19] conclude that the perception of fairness for the actual work schedule and the process used to generate that schedule is essential for satisfaction with the assigned schedule and can eventually attract and retain nurses.

Despite the complexity of constructing perfect schedules, in practice, these are still often created manually by a planner. This is time-consuming, and there can be a major difference between the quality of the schedules [20]. Furthermore, the planner picks a schedule considered 'fair' for everyone. However, fairness remains vague and subjective as the planner's understanding of a fair schedule may divert from the nurses or other planners [18]. To objectify the fairness aspects, a support system that automates the scheduling process can be used [17]. Additionally, automating the process saves time for the planner, and better quality schedules can be generated [8, 16]. Previous research has focused on nurses that work in the hospital. However, there is only a limited amount of research focusing on residential care organizations [21–23], while these are under pressure due to the ageing population and the shortage of professionals. According to Hulshof et al. [23], the dynamics of residential care services, although on a slower time scale, are comparable to that of inpatient care services. The latter delivers care to patients who are admitted for treatment and/or care and stay for a minimum of one night. Residential care provides supervision and assistance in activities of daily living with medical and nursing services required for the elderly who can no longer stay at home. As patients stay for an extended period of time or the remainder of their lives, the emphasis is on providing supportive care rather than acute care. As a result, it is beneficial to have a smaller nurse-patient ratio in this setting compared with inpatient care. This allows nurses to have more time to meet the patient individual needs and provide personalized care. However, both settings require care 24-hour-a-day, where the shifts are divided into day, evening, and night. It results that most planning decisions are similar for both services [23]. Hence, literature on the NSP in hospitals can be used to develop an appropriate scheduling approach for residential care organizations.

The goal of this research is to develop an automated scheduling method that fulfills the specific requirements of residential care organizations in the Netherlands while including fairness aspects to improve nurses' job satisfaction.

1.1 Research background

This section provides the background information required to understand the research position and problem. First, the NSP is introduced in the context of the three planning and scheduling levels within the framework of healthcare operations management. Subsequently, a description is given of the different scheduling methods used in literature and practice.

1.1.1 Nurse scheduling process

According to the framework of Hans et al. [24], the NSP can be considered under resource capacity planning as can be seen in Figure 1.1 [25]. This process can be divided into three planning phases: strategic, tactical, and operational level [24]. As can be seen each level corresponds with different planning time horizons [26], which are correlated with the uncertainty on the different levels. The uncertainty decreases over time, e.g., in the morning is exactly known which nurses are available or sick.

Strategic level

The strategic level involves the long planning horizon based on aggregated information and forecasts [23, 24]. Examples are capacity expansion, case mix planning, capacity dimensioning, and workforce planning. On this level, there is a lot of uncertainty, as it is unknown which nurses are available during the whole year or how many patients need care. We refer to Hans et al. [24] for a thorough description. As these decisions are taken before the actual scheduling process, the strategic level is left out of scope.

Tactical level

The decision taken on the strategic level provides the basis for the tactical level. On this level, the operations and execution of care delivery processes are addressed [23, 24]. The length of this horizon lies between the strategic and operational planning horizon, as seen in Figure 1.1. Compared with the operational level, this level creates more flexibility, is less detailed, and has less demand and staffing certainty. There is more certainty compared to the strategic level, as the shifts and contract agreements are known, e.g., fixed free day, contract hours, and allowed to work a night shift. Also, temporary capacity expansions like overtime or hiring staff are possible on this level. Resource capacity planning includes block planning, staffing, division of the day into shifts, scheduling policies, and admission planning [24, 26]. This level will be included in the scope of this research as we focus on staff-shift scheduling.

Operational level

Mid and short-term decisions are made on the operational level. On this level, there is low flexibility, as decisions on higher levels have set the scope for the operational level decision-making. On the other hand, there is less uncertainty, as the elective demand is entirely known, and only emergency demand has to be forecast [23]. There is a distinction between *offline* and *online* operational planning. The first concerns *in advance* planning. Given the workforce, it can be specified which shift a nurse should work resulting in a schedule. This is also referred to as *nurse rostering* [26]. There is less uncertainty compared to the tactical schedule, as it is known from the year planning which nurses will be absent, and therefore advanced rescheduling can take place before the start of the period. As this research aims to design a method that automates the scheduling process, the offline operational planning level will be included in the scope of the research. The online operational level involves *reactive* decision-making to unexpected events on short-term, e.g., add-on scheduling or emergencies. The schedule created during the mid-term planning serves as a suitable plan, but over time the staffing requirements or availability can change, and nurses should be rescheduled [26]. E.g., when there is a shortage, the online decisions include using overtime, calling in a nurse on a free day, using the flex pool, or working with the shortage. There is little uncertainty, as it is known which patients are present and which nurses can and cannot work.

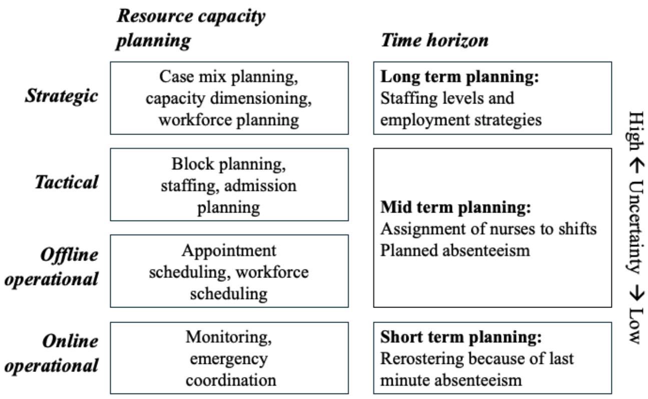


Figure 1.1: Resource capacity planning for health care planning and control in a general hospital based on the framework of Hans et al. [24], the time horizon concerning the three phases of nurse scheduling based on Rönnberg and Larsson [26], and the uncertainty over time.

1.1.2 Scheduling methods

As mentioned, the construction of high-quality schedules can contribute to enhanced job satisfaction. According to recent literature on the NSP [5, 8, 15, 26], on nurses' experiences around shift schedules [3], and on the influence of self-rostering [27], allowing employees to influence their schedule will contribute to this goal. This can be achieved by *preference scheduling*, with the most extreme form being *self-scheduling*, and *cyclic scheduling* [28]. The organizational procedure of this method can balance between *central scheduling* or *decentral scheduling.* In addition to these methods, an organization can choose to repeat a schedule for several periods with *cyclic scheduling.* Combining the previously mentioned methods with the cyclic method results in a central or decentral cyclic preference scheduling method.

During this research, the focus will be on these types of scheduling methods on the aforementioned online operational level. These methods will be briefly explained.

Preference and self-scheduling

With *preference scheduling* is meant that care workers may request shifts and days on or off, but the manager or planner is responsible for solving conflicts and constructing the final schedule [12]. This gives the employees the opportunity to influence their schedule and creates more autonomy. The disadvantage of this method is that after the final schedule is established, the employee finds out whether a preference can be fulfilled or not, and often it is no longer possible to integrate alternative preferences of the employee [5]. Wolbeck [28] mentions that from the three common scheduling categories that consider individual preferences, preference scheduling offers the greatest potential to generate a fair schedule. For a more elaborate description, we refer to Wolbeck [28].

When care workers cooperate to construct a schedule by signing up for shifts and solving conflicts together, it is called *self-scheduling* [12, 26]. According to van der Veen et al. [13], this enables, even more, to better cope with employees' preferences resulting in increased job satisfaction. Also, it leads to a reduction of the head nurse's scheduling time, increased belief in autonomy, and improved cooperation and teamwork [7]. Nonetheless, this method also has its drawbacks [7, 16, 26]. These schedules often do not match the staffing demand and reassignments must take place; schedules can be made at the convenience of the staff resulting in violating scheduling rules; there are no formal procedures for conflict-solving and because this is an active process it is time-consuming for the nurses. Since the outcome relies completely on the ability to cooperate and negotiate, the result can become unfair and eventually lead to conflicts[26].

Centralized and Decentralized scheduling

For preference scheduling, there are two organizational procedures, *centralized* and *decentralized scheduling.* Burke et al. [16] and Maenhout and Vanhoucke [7] describe the advantages and disadvantages of these two administrative procedures.

When this process is realized by centralized scheduling, one administrative department or employee of the organization constructs the schedules [7, 16]. The advantage of this method is that fairness can be better incorporated through consistent and objective application of policies and there is an opportunity for cost containment through better use of resources [7, 9, 16]. However, the disadvantages are that employees can have the feeling that the requirements of the team are not taken into consideration, schedules are unfair as they have no insight into the process, there is favoritism, or there is little employee autonomy.

Decentralized scheduling occurs when the head nurse or unit manager is responsible for generating the schedules. Here, there is more employee autonomy and personalized attention. However, this method does not guarantee fairness as there is less coordination and is time-consuming, and puts pressure on the head nurse or manager to create a popular and perfect roster [27]. Also, the quality of the schedule depends on the scheduling skills of the head nurse or manager.

Cyclic scheduling

An organization can choose to repeat a schedule for a predetermined number of periods, also called cyclic scheduling, or *fixed scheduling* [7, 16, 26]. The advantages are that the work is divided evenly, it is easy to manage from an administrative perspective and the schedules are known a long time in advance. The latter creates the opportunity for a better work-life balance. However, for the practical application, it has some drawbacks when an organization chooses to only use cyclic scheduling. These schedules are not flexible as they cannot address flexible work regulations, fluctuating demand, and personal preferences. This makes it difficult to deal with unexpected absenteeism. According to Kiermaier et al. [29], cyclic schedules offer a high degree of fairness and long-term predictability of days on and off. They introduce flexible cyclic rostering as means of accommodating limited weekly adjustments of employee schedules for the service industry. They showed that a reduction of undercoverage of more than 10% can be achieved with the proposed model.

The opposite is *non-cyclic scheduling*, or 'ad hoc' scheduling, which creates more flexibility by creating a unique schedule for each period [16, 26]. However, this is a time-consuming task for the planner. The other non-cyclic scheduling's drawbacks are opposed to the benefits of cyclic scheduling.

1.2 The scope of this research

As introduced already, healthcare organizations are facing the problem of a shortage of nurses. This can be addressed by creating high-quality schedules that incorporate fairness aspects, in order to increase job satisfaction. However, due to the many restrictions, it is a challenging and time-consuming task for the planner to devise a high-quality schedule for each employee. During this research, a method will be developed to support residential care in generating high-quality schedules that incorporate fairness aspects. To scope this research, the following will not be included. First, patients that need residential care live within the care organization, so the routing of the employees and planning of patients will be left out of scope. Furthermore, it is assumed that the decision on the strategic level, as well as the staffing demand and requirements, are known. Lastly, reactive decisions on the online operational level do not need to be considered within the proposed scheduling method. Section 2.2.1 provides a description of the scheduling process in practice, which will form the basis of the developed method and shift sequence used. To define the contribution of this research, a literature review is provided in Chapter 3 on current approaches to solving the NSP while including fairness aspects.

In this research, we will develop a two-stage method to construct nurse schedules for residential care organizations. First, we will construct a cyclic schedule, further referred to as the *tactical schedule*, which serves as a starting point. Within this tactical schedule flexibility is included, by scheduling nurses for 80% of their contract hours. The automated scheduling method ensures that the work is divided evenly, which contributes to increased fairness; and creates consistent schedules that are known in time, creating the opportunity for a better work-life balance. The tactical schedule will be used to construct the monthly period schedule, which will be further referred to as the *operational schedule*. In this schedule, all coverage requirements are met and nurses are allowed to be scheduled for 100% of their contract hours. However, as a result of periodic changes in demand, unexpected absenteeism, and additional requested preferences, adjustments need to be made to the assignments in the tactical schedule. To preserve the benefits derived from the implementation of a tactical schedule, it is important to have a certain percentage of flexibility in terms of permitted tactical reassignments. Throughout this research, we will refer to this as the *flexibility parameter*. In this research, experiments will be conducted to identify if allowing flexibility on the tactical schedule results in a high-quality and fair operational schedule and the optimal value of this parameter.

Three case studies will be used to evaluate the performance of the method. To increase the likeliness of impact, these will be conducted at Nedap Healthcare, which is the largest software developer for residential healthcare. They have developed a software suite called Ons. This application aims to simplify healthcare professionals' administrative tasks, thereby making more time available for actual care. Within this application, there is a scheduling module where shifts can be manually assigned to employees to create individual rosters. Eventually, Nedap Healthcare can implement the designed method in their software to better support the scheduling process of their customers.

1.3 Research framework

To facilitate residential care organizations in constructing high-quality schedules that incorporate fairness aspects, Nedap wants to extend its scheduling module with a new automated scheduling method. This research focuses on describing the criteria this method must meet, the design of the method, and validating the performance using multiple case studies. Therefore, the objective of this research is:

Develop a nurse scheduling method that supports Dutch residential care organizations in constructing fairness-enhanced tactical and operational schedules to sustain a better work-life balance and increase employee satisfaction.

The overall research question is: can we develop a method that supports residential care organizations to construct fair tactical and operational nurse schedules to sustain a better work-life balance and increase employee satisfaction?

To be able to achieve this objective, we formulate the following six research questions:

1. What is the current process of creating schedules in residential care?
2. What is fairness-enhanced scheduling according to literature and practice?
3. What is an appropriate method to generate fair schedules for residential care organizations?
4. How does the proposed method perform compared with the current performance?
5. Is including flexibility a good way to incorporate fairness in the operational schedule, and how does it impact the outcome of the schedule?
6. How does the proposed method perform when using different input data?
7. What are the recommendations when implementing this method in practice?

Questions 1 and 2 are answered using literature research, conducting interviews with multiple care organizations, and using the available knowledge within Nedap. The gained knowledge is used to formulate KPIs and criteria for the method. Based on the requirements defined by practice and the evaluated literature, a new automated scheduling method is designed to answer question 3. To evaluate the performance of the method compared with the current practice, data is used from different case studies and the results are analyzed with respect to the formulated KPIs, which will provide the answer to question 4. To answer questions 5 and 6, experiments are conducted to evaluate the influence of the flexibility parameter and the sensitivity of the method for different parameter settings and input data. At last, by critically reflecting on the results of the thesis question 6 is answered.

Chapter 2

Nurse Scheduling for Residential Care in Practice

This chapter gives an introduction to the scheduling process in residential care organizations. It provides a brief overview of the terminology used throughout this research and the nurse scheduling process in practice. Hereafter, the objectives and constraints from practice are introduced. To understand the process and determine the criteria for constructing a schedule in practice, we interviewed eight care organizations.

2.1 Terminology

This section gives an overview of the nurse scheduling terminology used in practice that will be used throughout this research.

- **Shifts.** As healthcare organizations deliver care around the clock, the day is divided into multiple shifts. These are the periods where work activities take place. Usually, there are three shifts, morning, evening, and night [30]. For full-time nurses, these have a length of 8 hours.
- **Clients.** The clients of care organizations are the people who receive care during the planning horizon. They receive intramural care as they live within the building of the care organization.
- **Nurses.** The nurses deliver care to the clients. Nurses can have different qualification levels (QLs) obtained from training or education, enabling them to work specific shifts. If a shift is assigned to the nurse, the qualification requirement must be fulfilled. Nurses have their own contracts where general tasks and agreements are defined, including working hours, vacation days, rest days, or not working night shifts.

- **Minimum staffing levels.** The needed care of clients determines the minimum staffing levels, which is the minimum number of nurses needed for certain shifts on a certain day [31]. When this minimum is not reached, there is *undercoverage*.
- **Coverage requirements.** These are the specific QLs and skills requirements corresponding with each shift. Each organization has defined *hour types*, which represent a specific skill or task required for a shift or assigned to a nurse, e.g., one hour type indicates that a nurse can administer medication or that he is classified to change clothes. The shifts are associated with a set of hour types that defines the minimum hour types needed. If a nurse has more hour types assigned than needed, he or she is *overqualified* for that specific shift. On the other hand, if he or she is missing an hour type, he or she is *underqualified*.
- **Responsible shift.** Shifts with a specific QL need to be present 24 hours a day within the organization.
- **Tactical schedule.** A standard schedule that is repeated every 4-8 weeks, providing the nurses with a predictable schedule and enabling them to have a better work-life balance. It includes, among others the fixed free day and should not violate any law legislation.
- **Operational schedule.** This is the final schedule consisting of 4-8 weeks, where all shifts are assigned, and there are no remaining conflicts. Each period nurses can request additional wishes for free days, which do not have to be fulfilled by the planner. This operational schedule is shared 4-8 weeks prior to the start of the corresponding period.
- **Year planning.** It provides an overview of predictable absenteeism throughout the year for a nurse, e.g., holidays, courses, and pregnancy.
- **Contract hours.** These are the agreed working hours per period. When a nurse works more or less than the agreed contract hours, there can be additional *plus* or *min hours*, respectively. In the upcoming period, these are compensated by assigning the nurse to fewer or more shifts. If a nurse has hours assigned, we refer to them as a nurse or *regular* nurses. When a nurse does not have hours assigned but does work for the specific organization, we refer to them as *intra-organizational flex* nurses. Otherwise, we refer to them as *extra-organizationl flex* nurses.
- **Fixed free day.** Each nurse can request a regular day off from Monday until Friday at 4:00 AM. The planner must fulfill this request in order to comply with law legislation [32].

2.2 Scheduling in practice

Maenhout and Vanhoucke [7] mentions that with the use of organizations' scheduling policies, the organizational support of nurses should be addressed. Although these policies can differ between organizations, the main premise of the policies is meeting the patient's demand for care [33–39]. To guarantee the continuity of care, minimum coverage requirements are specified for each shift on each day [7]. The goal is to assign nurses in such a way that the coverage requirements are met while meeting other requirements and keeping in mind the aspects of the three stakeholders: the clients, the organization, and the nurses. Leading are the law legislations from the collective employment agreement (CEA) [32] and the Working Hours Act (WHA)[40]. Additional scheduling requirements and agreements are outlined in the scheduling policies of the organization, which will be further explained below.

2.2.1 General scheduling process

Each organization has a different approach to generating operational schedules, where nurses are assigned to shifts or otherwise will be off duty. In this thesis, we focus on the process of the organizations that schedule centrally and use a tactical schedule to generate the operational schedule. The steps to construct both schedules are further described and are visualized in Figure 2.1.

First, a capacity plan is made where an estimation is made on the required number of hours per function per day to meet the demand for care. Hereafter, the staffing demand and coverage requirements are determined for each shift. These requirements are used to construct the tactical schedule. For a time period of four consecutive weeks, shifts are assigned to regular nurses. Depending on the organization, these nurses are scheduled for all or a fraction of their contract hours. The tactical schedule considers the fixed free day and complies with labor legislation. When finalized, the tactical schedule is repeated for a predetermined number of periods. On an annual or half-year basis, the expected absenteeism, such as holidays, education, and pregnancy, are inventoried and combined into a year planning.

Subsequently, the operational schedule is constructed 12 weeks prior to the corresponding period. Each nurse can request incidental wishes that do not have to be fulfilled by the planner. The tactical schedule is used as starting point, and the planned absenteeism from the year planning is gathered for the designated period. Shifts assigned in the tactical schedule that conflict with the planned absenteeism will be removed from the tactical schedule. This leads to unassigned shifts and, thus, not meeting the coverage requirements. The primary task of the planner is to adapt the schedule in a way that resolves these unassigned shifts. One approach is to allocate the open shifts to employed nurses who have min-hours. Alternatively, when the solution cannot be found in the own team or organization, the planner can choose to export the shifts to an external flex pool. Finally, the final operational schedule is shared with the nurses, providing them with their assigned shifts and work schedule 4-8 weeks prior to the start of the corresponding period depending on the organization. This schedule is realized and carried out at the start of the designated period, as seen in Figure 2.1.

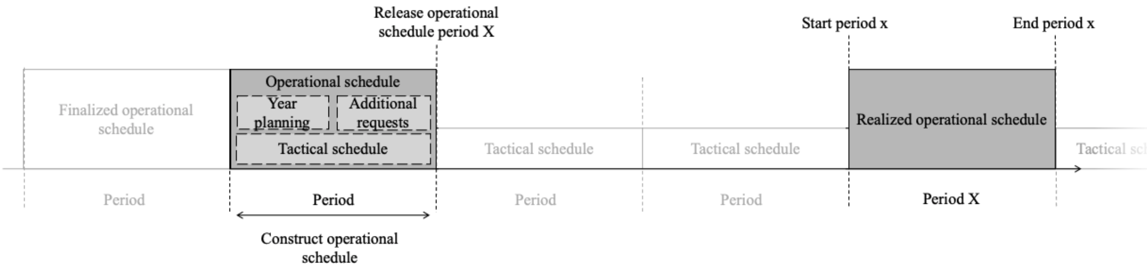


Figure 2.1: Outline of the overall scheduling process when combining the tactical and operational schedule.

2.2.2 Sequence shift scheduling

In practice, the shifts are assigned in a specific sequence in the tactical and operational schedule. This sequence differs between organizations but also between planners within the same organizations. The most common sequence is assigning the night shifts, followed by either weekends and responsible shifts or vice versa, and then the remaining weekly shifts. This sequence will be applied in this research to find a first solution, as explained in Section 4.2.2. A different approach was used by one of the interviewed organizations. Here, they distinguished the night shifts that occur Monday (Mo) until Thursday (Th) from those on Friday night. First, they assign the Mo-Th night shifts, followed by the weekend shifts. Subsequently, they assign the Friday night shift, preferably to nurses working the successive weekend. Lastly, the remaining shifts are assigned.

2.3 Objectives and Constraints

Throughout the scheduling process, many rules have to be taken into consideration. These arise from law legislation, work contracts, organizational standards, qualifications, and availability. It is a challenging task for planners to satisfy all these rules. Therefore, organizations have classified the rules into hard and soft rules, which are equivalent to hard and soft constraints discussed in literature [30, 31]. There are different objectives for the tactical and operational schedules. Whereas the tactical schedule prioritizes satisfying the nurses, the operational schedule aims to meet all coverage requirements. The tactical schedule aims to create a high-quality individual schedule for each nurse, taking into account their preferences and individual contract agreements. It aims to have no law violations and meet the coverage requirements as much as possible. As the schedules are repeated over a period of time, it provides opportunities for a better work-life balance. Within these schedules, nurses are scheduled for at most 80% of their contract hours. The goal of the operational schedule is to deliver the required care in each period. Therefore, all shifts must be assigned while complying with labor legislation and keeping in mind the preferences of the nurses. In practice, the planners aim to achieve fairness by distributing the weekend, night, and day shifts evenly and having a fair distribution of plus and min hours in both schedules. Also, they aim to have a fair distribution of fulfilled additional wishes in the operational schedule.

2.3.1 Hard and Soft constraints

Based on the interviews, the hard and soft constraints from practice are identified. Those are in line with law legislation and the organizations' scheduling policies [32–39], and some will be briefly explained. We make a distinction between the constraints for the tactical schedule and the operational schedule, as the operational schedule allows some exceptions. First, the constraints for the tactical schedule are explained, followed by the constraints and exceptions for the operational schedule.

The most common hard constraints for the tactical schedule are:

- Every nurse should have one fixed free weekday.
- QLs of the shifts are fulfilled.
- A nurse can work only one shift per day.
- Shifts are assigned in a forward rotating order.
- Nurses cannot work on the agreed planned absenteeism from the year planning.
- There is enough rest time between shifts and days.
- A nurse cannot work a night shift when they are older than 55 or younger than 18 or due to personal contract agreements.

The soft constraints for the tactical schedule are that a nurse:

- Can work a maximum of three consecutive night shifts.
- Can be scheduled for a maximum of five days in a seven-day schedule.
- Has a maximum of 10 plus or min hours;
- Has as many free consecutive nights as possible.
- Has a free night around a weekend off.

Before constructing the operational schedule, nurses can request additional wishes, such as free days. These are additional soft constraints as it is not mandatory to meet these requests. The constraints for the tactical schedule also apply to the operational schedule. Nonetheless, the following exceptions can be made:

- An nurse can work more than one shift in a day, due to e.g. unexpected absenteeism.
- The length of a shift can be extended to 12 hours instead of the max of 10 hours.

- Once every four weeks an exception may be made on the forward rotation constraint.
- In agreement, a nurse works more than five days a week.
- In agreement, a nurse works more than three consecutive night shifts.
- In agreement, a nurse can work the night shift if they are older than 55.

2.4 Conclusion

This section provides an answer to the research question: *what is the current process of creating schedules in residential care?* Eight Dutch residential care organizations are interviewed which leads to an overview of the terminology used in practice and that will be used throughout this research. Additionally, based on the interviews the scheduling process, corresponding objectives, and the hard and soft constraints from practice are identified. Fairness is achieved by an equal distribution of weekend, night, and day shifts, plus and min hours, and the fulfillment of additional wishes in the operational schedule. This lays the foundation for the proposed method that takes into account the priorities and constraints from practice.

Chapter 3

Literature Review

In this section, we present a literature overview on solution approaches to solve the NSP and fairness aspects in nurse scheduling. In recent years, there has been an increased interest in considering fairness aspects in personnel scheduling [17]. Wolbeck and Kliewer [17] provides a review of personnel scheduling approaches that consider fairness aspects. They mention that the effects of schedules perceived as unfair are: decreased job satisfaction, lower job performance, bickering, increased absenteeism, increased turnover and triggering of labor strikes [41]. To prevent these, the allocation of human resources should consider fairness aspects. It has been put forward by Warner [42] that fairness can be used as a quality measure but has not been explicitly addressed in former research on nurse and employee scheduling [16, 43]. First, we will discuss research that has proposed mathematical approaches to solve the general NSP. Additionally, an overview of recent literature and approaches that incorporate fairness aspects in employee and nurse scheduling is provided. Table 3.1 provides an overview of the studies that have proposed solution methods for the NSP and the position of this research.

3.1 Fairness in nurse scheduling

Wolbeck [28] mentions two distinct angles from which fairness can be assessed, being the *perspective of fairness* and the *time horizon of fairness.* The latter can be divided into *short-term* and *long-term fairness* [44]. Often fairness aspects are considered in one planning period for nurse scheduling and, therefore, only ensure short-term fairness. The allocation of resources is measured over a longer time period or at the end of the cycle to ensure long-term fairness [44]. Shi et al. [44] state that short-term fairness has a more significant impact on the quality of service, and long-term fairness is more important when resources are scarce.

Within the perspective of fairness, there is a distinction between *group* and *individual fairness* [28]. According to Shi et al. [44], group fairness is achieved when all individuals are treated equally and the outcomes are distributed across all individuals. Individual fairness evaluates the situation from a self-centered point of view and compares if one individual is treated unequally compared to other individuals [45]. Within nurse scheduling, it is more challenging to access individual fairness, as each nurse may have a different perspective of fairness.

Uhde et al. [18] aims to understand better what determines the fairness of a nurse shift schedule and how systems can support fair planning. They concluded that the oversimplified concept of fairness as equality does not capture nurses' understanding of fairness. On a general level, equality should be the goal. Schedules should meet a similar number of wishes, free weekends, and similar abilities to include preferences. However, when conflict arises due to, e.g., overlapping preferences, this should be solved on a need-basis. A computer could support this by finding the conflicts, presenting legal solutions, and indicating how and when the nurses should resolve the conflicts in advance.

3.1.1 Quantification of fairness

In practice, the quality of schedules differs because they are evaluated by the decision-maker based on experience and understanding of fairness. Therefore, the research of Wolbeck [28] aims to quantify an objective function that includes fairness aspects. The general goal is to find a feasible solution that satisfies all hard constraints. In addition, a nurse-specific penalty score is added that indicates how high the penalization is for violating soft constraints. The higher the score the more dissatisfied the employee is with the schedule. We refer to Wolbeck [28] for an overview of the literature's most commonly used fairness objectives that consider the two angles concerning fairness.

To measure group fairness, Jain et al. [46] have introduced the *Jain's index*. This index lies in the range of $1/R$ to 1, with R being the number of resources. An index of 1 indicates the best group fairness because all resources are treated 100% equally. This index can be used to measure the fairness of a resource allocation scheme and can be applied to any resource-sharing or allocation problem. According to Jain et al. [46], this fairness measure fulfills the four properties they have defined: population size independence, scale and metric independence, boundedness, and continuity [46]. Their research shows that the fairness measures proposed in the literature, e.g., variance, coefficient of variation, and min-max ratio, do not fulfill these properties. However, Burget and Rudová [47] questions the use of this index because of its dependency on the relative sizes of individual penalties. They state that a higher value of Jain's index does not mean better fairness but can relate to a uniform degradation of all penalties.

Wolbeck [28] concludes that there are several ways to integrate fairness aspects into personal scheduling. First, they conclude that one objective of fair personal scheduling should be distributing the workload evenly among the employees to enhance group fairness. Additionally, they conclude that considering individual preferences is essential to give employees autonomy in the process and thus increase the positive acceptance of the schedule. Moreover, they state that when an objective does not aim at a fair distribution of preference fulfillment, the approach does not adequately reflect fairness. The research concludes that the essential type of request is the request for the shift on/off and, therefore, an essential objective in fair scheduling. Furthermore, Wolbeck [28] states that short-term fairness is inherently considered in each fair schedule, which is different for long-term fairness. Therefore, to increase satisfaction, long-term fairness should be a mandatory objective. Ideally, the generation of the schedules is based on different objectives and is further evaluated.

3.2 Approaches to incorporate fairness

This section discusses previous research that incorporates fairness aspects within nurs scheduling. Based on the literature, a distinction can be made between research that uses constraints to ensure fairness and research that defines fairness as an objective. The studies that use fairness as a constraint or objective are also included in Table 3.1.

3.2.1 Fairness as constraint

When it comes to nurse scheduling, previous models make a distinction between coverage constraints and time-related constraints [43]. The latter is noteworthy in a fairness context, e.g., balancing working hours or weekend work among full-time nurses. Maenhout and Vanhoucke [7] guarantees fairness between nurses within the monthly schedule by considering time-related constraints, e.g., the minimum and the maximum number of weekends. Whereas Burke et al. [30] ensures fairness by balancing the working time.

Hadwan and Ayob [48, 49] incorporate fairness by a soft constraint attempting to distribute the workload and days off evenly. Hadwan and Ayob [49] introduce a semi-cyclic shift pattern approach (SCSPA), where the night shift patterns are allocated cyclically, followed by allocating the morning and evening shifts in a non-cyclic manner. They compare the performances with their previous work, where they used a non-cyclic shift pattern approach (NCSPA) [48]. Both studies propose a two-stage model to solve a real-world NSP. First, a constructive heuristic method is used to find feasible shift sequence patterns. In addition, SA is applied to optimize the constructed feasible solution, with the objective to minimize the deviation from the eight goals due to violating the soft constraints. Using the SCSPA, two benefits are gained. First, the number of shift patterns decreases, reducing construction time. Second, allocating the night shift patterns fairly become more manageable. Their results show that the proposed model can meet all hard and soft constraints of the hospital's rostering system.

3.2.2 Fairness as objective

To determine the quality of the solution, various objectives are applied in literature, as seen in Table 3.1. A commonly used objective in nurse rostering is the weighted sum, $MinWs$, which minimizes two parts [43, 50]. These are the cost of assigning a nurse to a given shift and the coverage violations that occurred by any over- or understaffing. However, Ouelhadj et al. [43] and Lavygina et al. [51] show that this objective does not result in fair solutions.

Lavygina et al. [51] mentions that if a solution quality is measured by the simple $MinWs$ of constraint violations for each employee, an optimizer may produce solutions in which some employees suffer a highly disproportionate share of these violations. This results in unfair distribution between individual schedules, which is also highlighted by Ouelhadj et al. [43]. Ouelhadj et al. [43] emphasized the importance of a fair distribution of contractual violations among nurses, as it has a direct impact on their satisfaction and overall job satisfaction [50]. Therefore, these and other studies introduce new objectives to include fairness without the expense of another nurse.

Lavygina et al. [51] considers fairness as an additional objective, defined as the deviation of individual workers' schedule constraint violation penalties, i.e., minimizing the standard deviation of penalties. To evaluate the performance, they use the problem introduced in their earlier study [52], which was solved using a construction heuristic and optimizing it with SA. The aim was to find a schedule that minimizes the hours worked in violation of the soft constraint. The results of Lavygina et al. [51] indicate that optimizing multi-objectives that minimize the total sum of penalties and incorporate the proposed fairness objective leads to better schedules than considering only one of them as a single objective [28, 51].

Another objective to ensure fairness is the $min - max$ or $max - min$ objective, introduced by Ouelhadj et al. [43], Smet et al. [53], and Constantino et al. [54]. Using this objective, the quality of the worst individual schedule determines the overall solution quality, ensuring that the nurses' schedules will not be improved at the expense of the worst individual schedule [43].

Ouelhadj et al. [43] and Smet et al. [53] both use the $min - max$ fairness-based objective where violations of time-related constraints are penalized using a self-scheduling approach. Ouelhadj et al. [43] uses a cooperative meta-heuristic agent-based framework to incorporate fairness in nurse rostering. Their research showed that good values of the new objective correspond to rosters that are fairer than those found by optimizing $MinWs$ without aggravating the quality of the roster [43]. Smet et al. [53] concluded that with the new objective, the quality of the individual rosters varies less, thus producing fairer solutions. However, the result is not consistent for all instances. They recommend optimizing the new objective while improving the original $MinWs$ objective without decreasing the quality of the worst individual schedule.

The $max - min$ objective is applied by Constantino et al. [54], where they introduce a new variant of the NSP called the *nurse scheduling with balanced preference satisfaction* (NSBPS). The total preference satisfaction is evenly distributed to ensure fairness by considering individual preferences, where the minimum individual satisfaction is maximized. The total preference is expressed by the sum of preference satisfaction considering each shift assigned to the nurse in the schedule.

Martin et al. [50] uses the previous work of Ouelhadj et al. [43] to examine four fairness objective functions that distribute penalties for time-related constraints and individual requests for shifts equally among all nurses [17, 50]. They use Jain's index to evaluate the relative fairness of their solutions. It results that using the mean deviation as an objective outperforms the others, followed by the $min - max$ objective. Their study also concludes that the $MinWs$ performs the worst regarding fairness.

Osman et al. [20] and Tsaia and Leeb [55] introduce another approach, where they try to evenly distribute days off among all nurses. Osman et al. [20] develops a two-phase heuristic with an objective of fair distribution of staff at various shifts and compliance with constraints. The quality of the schedule is determined with a fairness measure being the standard deviation between nurses' days off. They conclude that their proposed algorithm can ensure fairness.

A two-stage model is designed by Tsaia and Leeb [55] to solve the NSP. The first model is designed to identify the optimal solution of a complete off-shift table, which is the optimal vacation schedule for the next month. This schedule is generated using a self-scheduling approach. The algorithm checks for regulation violations and schedules vacations fairly. To ensure fairness, the objective of the first model is to minimize the variance of days off on Saturdays, Sundays, and holidays. Hereafter, the second model tries to complete the entire schedule using a GA. Using a case study, they show that their approach reduces the workload for generating the schedule and increases the nurses' satisfaction by providing vacation fairness and incorporating self-scheduling [55].

3.3 Contribution of this research

As mentioned, the NSP is a widely studied subject, and many approaches have been proposed to model and solve the problem. We refer to Burke et al. [16] and Ngoo et al. [8] for an overview. Mathematical programming approaches, such as linear programming, dynamic programming, and constructive or improvement heuristics, have been widely applied to solve the employee scheduling problem [15]. However, solving the NSP is complex, challenging, and time-consuming due to the high constraint density [16]. Osogami and Imai [56] have proven that the NSP is NP-hard due to many hard constraints that must be fulfilled and the soft constraints that must be considered to construct a schedule. So, to solve large real-world problems, meta-heuristics can be applied [56]. As seen in

Table 3.1, the meta-heuristic simulated annealing (SA) has been widely applied in the literature to solve the NSP. Jafari and Salmasi [57] discuss that generating an initial solution is a difficult task due to various types of constraints considered. Therefore, they conclude that a meta-heuristic, such as SA, that only needs one initial solution is more appropriate to solve the NSP. Within the literature, different approaches are used to construct an initial solution, e.g., constructive heuristic [48, 49, 52], mixed integer programming-based heuristic [58], and based on nurses' preferences [59]. Turhan and Bilgen [58] and Ceschia et al. [60] test the performance of the SA on available datasets and showed that the SA method outperforms most of the techniques. Both studies first construct an initial feasible solution which is optimized using the SA algorithm. The use of SA in a real-world setting is shown by Hadwan and Ayob [48, 49], Lavygina et al. [52], Jafari and Salmasi [57], Lin et al. [59]. These studies take nurses' preferences into account, and additional factors such as hospital policies, labor laws, and governmental regulations are considered in the method.

The researches show that the SA algorithm converges to good-quality schedules in a short period of time. Therefore, we will apply SA in this research to support residential care organizations constructing fairness-enhanced schedules.

Moreover, recent research has focused on incorporating fairness in nurse scheduling, either by including fairness aspects as constraints or as an objective. Ouelhadj et al. [61] and Smet et al. [53] use a self-scheduling method to take nurses' preferences into account. However, as mentioned, this method takes a lot of time for nurses, and fairer schedules can be constructed using preference or cyclic scheduling [28, 29]. Martin et al. [50] takes requests into account by adding them as a soft constraint. Nonetheless, besides an equal workload distribution, considering individual preferences is essential to give employee autonomy and should therefore be taken into account as an objective in fair scheduling [17]. Most proposed methods start by constructing a schedule based on an empty schedule. In order to give nurses the opportunity to improve their work-life balance, schedules must be known in time. This can be achieved by using a tactical schedule. Yet, only Hadwan and Ayob [49] uses a semi-cyclic scheduling approach. However, their method used a single objective where they minimized the deviation from the violations of the soft constraints and did not take preferences into account.

In order to bridge the gap between recent research and the requirements from residential care organizations from practice, we aim to develop a two-stage method that distributes the workload evenly and takes individual preferences into account to ensure fairness while meeting coverage requirements and law legislation. First, in order to sustain nurses with a better work-life balance, a tactical schedule will be constructed where nurses will be scheduled for 80% of their contract hours in order to create flexibility. Subsequently, this schedule will be used as input to construct the final operational schedule, where the predetermined leaves and days off from the year planning are processed. To sustain predictability and a stable schedule, a certain percentage of the tactical schedule must remain unchanged. Therefore, after removing the shifts that conflict with the planned absenteeism, we allow certain flexibility on reassignments of the tactical schedule to construct the final operational schedule, which meets the periodic coverage requirements. To the best of our knowledge, we are the first who combines a tactical and operational schedule to solve the real-world NSP, while simultaneously exploring the potential influence of flexibility on achieving fairness within the schedule. By incorporating flexibility within the tactical schedule, we aim to generate high-quality fair operational schedules to allow for a better work-life balance and increase employee satisfaction.

3.4 Conclusion

The literature found provides answer to the question *What is fairness-enhanced scheduling according to literature?*. In the literature, the concept of fairness can be categorized into two perspectives: group and individual fairness. Additionally, fairness can also be viewed in terms of time horizon, with short-term and long-term fairness being the two categories. Previous research concluded that besides a fair distribution of the workload, individual preferences should be considered. Within nurse scheduling, there are two approaches to incorporating fairness. First, fairness can be considered as a constraint, which is often addressed through time-related constraints by balancing the number of days off or balancing the workload. Second, fairness is considered as an objective. Several objectives are introduced that ensure that nurses' schedules will not be improved at the expense of another schedule.

To determine an appropriate method for solving the NSP in this research, a literature review was conducted on previous approaches. It showed that SA is an effective method for constructing high-quality schedules quickly for real-world problems. Hence, SA will be applied in this research and serves as the starting point for addressing the third research question. The findings from literature and practice allow us to formulate the contribution of this research, which is to develop a two-stage method that constructs a tactical and operational fairness-enhanced schedule that includes flexibility to sustain employees with a better work-life balance and increase job satisfaction while meeting the requirements from practice.

Table 3.1: Overview of literature that has proposed approaches for achieving fairness when solving the NSP.

			Burke et al. [62]	Burke et al. [30]	Bövarsdóttir et al. [12]	Ceschia et al. [60]	Constantino et al. [54]	Hadwan and Ayob [48]	Hadwan and Ayob [49]	Jafari and Salmasi [57]	Kletzander and Musliu [63]	Lavygina et al. [52]	Lavygina et al. [51]	Lin et al. [59]	Martin et al. [50]	Maenhout and Vanhoucke [7]	Osman et al. [20]	Ouelhadj et al. [61]	Oyeleye et al. [64]	Smet et al. [53]	Soriano et al. [31]	Tsaia and Leeb [55]	Turhan and Bilgen [58]	Wolbeck and Kliewer [17]	This research
Objective		Min. sum of soft constraints violations	x	x	x	x		x	x		x	x	x	x	x	x			x		x	x	x		x
		Min Max individual preferences					x																		
		Min Max individual soft constraint violations																x		x					
		Min coverage violations													x						x				
		Max individual preferences			x														x		x				
		Min weighted average																							
		Min overtime of employee																			x				
		Max preferences for working shifts and weekends off								x															
Constraints	Hard	Coverage requirements	x	x		x		x	x	x	x	x	x	x		x	x		x			x	x		
		One person per shift per day		x		x		x	x	x					x	x	x	x	x		x	x	x	x	x
		Skill category	x	x		x		x	x						x			x							x
		Shift succsion				x	x	x	x	x	x	x	x	x			x		x			x	x		
		Other						x	x	x				x	x				x		x	x	x		
	Soft	Personel constraints	x	x	x	x	x			x				x	x						x		x	x	x
		Organisational constraints	x			x				x				x	x			x			x		x		x
		Balancing the workload		x	x			x	x																x
		Work regulations	x			x	x		x	x	x	x	x		x	x		x	x						x
Fairness	As objective	Min deviation of individual workers schedule constraint violation					x						x		x									x	x
		Min deviation days-off															x								
		Min variance off days weekend and holidays																				x			
		Min max penalties									x				x			x							
	As constraint			x	x			x	x							x									x
Preferences		Day off	x		x	x		x	x	x	x	x	x			x	x			x	x	x	x	x	x
		Shift assignment			x	x	x			x	x	x	x	x						x			x	x	
		Weekend off								x	x					x				x					x
		Requested assignemnts	x		x										x		x	x		x		x	x		
Method	Metaheuristic	Tabu search		x											x			x							
		Simulated Annealing				x		x	x	x	x	x	x	x	x			x					x	x	x
		Local Search					x							x											
		Genetic Algorithm																	x			x			
		Other heuristic															x								
	MINLP																					x			
	MIP				x																x		x		
	CP																			x					
	Branch-and-price															x									
Two-phased						x	x			x						x	x					x		x	x
Scheduling Method	Preference		x	x		x	x	x	x	x		x	x	x	x	x			x		x		x	x	x
	Self													x				x		x		x			
	Cyclic								x							x								x	x

Chapter 4

Proposed Method

The studies of Hadwan and Ayob [48, 49], Lavygina et al. [52], Jafari and Salmasi [57], Turhan and Bilgen [58], Lin et al. [59], Ceschia et al. [60] have applied the meta-heuristic SA to solve the NSP. First, a constructive heuristic constructs an initial feasible solution, which is used as input for the SA algorithm to improve the solution. The violations of soft constraints are used as the quality measure of the schedule. The studies of Hadwan and Ayob [48, 49], Lavygina et al. [52], Jafari and Salmasi [57], Lin et al. [59] have been tested on real-world cases and have shown that the SA algorithm converged to good-quality schedules.

In this research, we propose a flexible two-stage algorithm to generate nurse schedules for residential care organizations. The steps and information considered are visualized in Figure 4.1. First, we construct a tactical schedule in which 80% of the regular nurses' contract hours are scheduled. Second, the tactical schedule is used as input to generate the operational schedules. A constructive heuristic is implemented to find an initial feasible tactical schedule, which is optimized using an SA algorithm. As the tactical schedule does not consider the planned absenteeism from the year planning in the scheduling process, assigned shifts that overlap with this absenteeism have to be removed from the tactical schedule. The remaining tactical schedule is used as input to generate the operational schedule using the adapted SA algorithm. The operational schedule schedules 100% of contract hours and allows for min hours if needed. Also, regular nurses and intra-organizational flex nurses are scheduled, where it prioritizes scheduling the first. The tactical schedule aims to construct fair schedules while meeting coverage requirements and law legislation, whereas the goal of the operational schedule is to meet the coverage requirement while retaining the tactical schedule assignments.

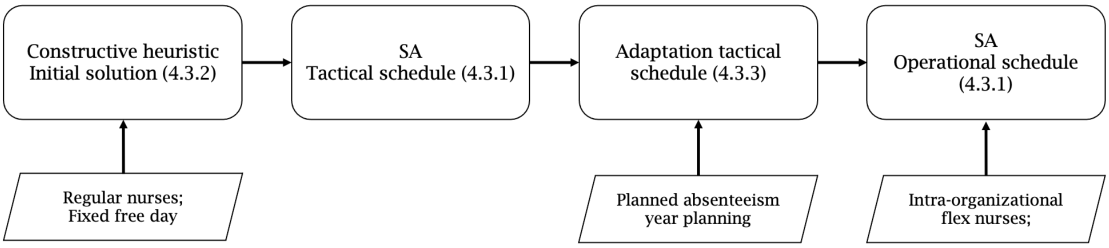


Figure 4.1: Overview of the two-stage scheduling process within this research and the information flow.

The proven ability of the SA to generate good-quality solutions is the major motivation to apply SA in this research. In the remainder of this chapter, we will introduce the hard and soft constraints and objective functions used for both the tactical and operational schedules and the additional constraints exclusively considered in the adapted SA algorithm when constructing the operational schedule. Hereafter follows the description of the constructive heuristic to find an initial solution, the neighbourhood operators, and the probabilities regarding the operators. Furthermore, the KPIs used to quantify the performance of the proposed method are introduced in Section 4.3. At last, the assumptions made regarding the method and the normalization of the penalties are described in Section 4.4 and 4.5. In line with Ceschia et al. [60], in this research, we discuss the static version of the NSP, where all information is known at the beginning of the planning horizon.

4.1 Hard and Soft constraints

To obtain a feasible schedule, all hard constraints must be satisfied. Violations of the soft constraints are allowed but penalized [25]. These are used to measure the quality of the solution. In the optimization approach, the objective is to minimize these penalties. We will describe the hard and soft constraints used in this research.

4.1.1 Hard constraints

For the NSP, we have defined the following four hard constraints (HC):

- **HC1: One shift per day.** Each nurse can work at most one shift each day or has a day off.
- **HC2: Fixed free day.** Nurses can request a fixed free day which should be assigned to them within the tactical schedule. This is seen as absenteeism and a nurse cannot be scheduled on this day. This HC is implemented to include nurses' preferences.
- **HC3: Qualification level 3.** For the shifts that require a qualification level 3 (QL3), only nurses with the minimum qualifications can be assigned. This is to satisfy the requirement

that during the whole day, a QL3 nurse should be present.

- **HC4: Minimum and maximum age night shifts.** According to law legislation, to be able to work a night shift, a nurse must be older than 18 and younger than 55 years.

4.1.2 Soft Constraints

The soft constraints are divided into *time-related constraints* (TRC) and *organizational-related constraints* (ORC). We aim to minimize the penalties occurred by violating the soft constraints. Appendix A provides a detailed description of the calculations of the penalties for violations of soft constraints.

4.1.2.1 Time-related soft constraints

The TRCs are based on the law legislations in the Netherlands [32, 40], which include violations of rest time, forward rotating order, consecutive working days and night shifts, and amount of working weekends. In addition, three constraints are added in order to prevent the pattern of assigning shifts 'on-off-on', to distribute the contract hours equally, and to have a fair distribution of night shifts. The calculations of the TRCs are based on the approach used in Lavygina et al. [52], which determines the missing rest hours according to law legislation, e.g., a nurse should get assigned eleven hours of rest between two shifts; when eight rest hours are assigned, the penalty equals three missing rest hours. The following TRCs are included in both tactical and operational schedule:

- **SC1: Rest time between shifts.** Nurses should get assigned a daily minimum rest time (RT) of 11 consecutive hours.
- **SC2: Forward rotating order.** Consecutively assigned shifts need to follow a forward rotating order (FRO), where the start time of the next shift is not earlier than the start time of the current shift.
- **SC3: Consecutive working days.** An employee should get 36 hours of RT after 5 consecutive working days (CWD).
- **SC4: Consecutive night shifts.** An employee should get 48 hours of RT after 3 or more consecutive night shifts (CNS).
- **SC5: Maximum of two weekends.** Nurses are limited to working two weekends (2W) in a four-week period.
- **SC6: Every other weekend.** Nurses should work every other weekend (EOW).

- **SC7: Forbidden patterns.** To maintain consistency in the nurses' schedule, we aim to prevent a nurse from having a free day between two consecutive shifts, which results in the pattern 'on-off-on' (SFS).

- **SC8: Remaining minutes.** Nurses should work according to their agreed contract hours. Additional overtime or remaining minutes (RM) should be distributed evenly among all nurses.

- **SC9: Ratio night day.** We aim to have an equal distribution of night and day shifts (RND). The penalty is calculated by determining the absolute difference in the ratio of night and total shifts and the ratio of the day and total shifts.

All SC penalties are measured in units of seconds, except SC8 and SC9, which are a ratio between [0,1]. Appendix A.1 provides a detailed description of the calculations of the penalties for violating the TRC. The violations of TRCs are used to incorporate fairness in the proposed method, as we aim to distribute the penalties among the nurses evenly, applying the objective function explained in Section 4.1.4 Equation 4.5.

4.1.2.2 Organisational-related soft constraints

The ORCs include undercoverage of shifts during the week, weekend, and night, missing hours of QL3, and the number of UQ shifts. In practice, coverage constraints are often relaxed by the planner if the staffing capacity is not fitting [65]. Therefore, we implement these as soft constraints. The penalties encountered are measured in units of minutes. In Appendix A.2, a detailed explanation of the calculations of the penalties for violating the ORC is provided.

- **SC10: Coverage constraints.** The goal of a care organization is to deliver the right care at the right time. Therefore, we aim to assign the minimum number of nurses to meet the staffing demand for each shift during the week, weekend, and night.

- **SC11: QL3 coverage.** A nurse with QL3 should be present 24 hours in the care organization.

- **SC12. Underqualification.** To deliver the right care, QLs should be taken into consideration. Shifts that are assigned to nurses that are missing the required QLs are denoted as underqualified (UQ).

4.1.3 Additional constraints operational schedule

When constructing the operational schedule, the agreed planned absenteeism from the year planning is taken into account. During agreed-upon planned absenteeism, e.g., holidays or education, a nurse can not work, and the assigned shifts in the tactical schedule are removed. In addition, each period, a nurse can request additional free days or weekends. These are not considered hard constraints but are taken into consideration in order to increase employee satisfaction. The following constraints are added for constructing the operational schedule:

- **HC5: Planned absenteeism.** A nurse cannot work during the planned absenteeism from the year planning. This HC is also used to meet nurses' preferences.
- **SC13: Percentage flexibility tactical schedule.** To provide nurses with predictable and stable schedules, a percentage of the tactical schedule should be preserved in the operational schedule. As this percentage is mandatory, a high weight is assigned such that it is not violated. This SC is also implemented to improve fairness in the operational schedule.

SC13 ensures that a predetermined percentage of the tactical schedule is preserved in the operational schedule. We refer to this as the *flexibility parameter*, which is the allowed amount that the operational schedule can deviate from shifts assigned in the tactical schedule after removing the planned absenteeism. For each nurse, we determine the number of tactical shifts rescheduled or removed within the operational schedule. An example is provided in Figure 4.2. Based on the remaining number of original tactical shift assignments in the operational schedule, we determine the preserved percentage and the penalty encountered. We provide an example of determining the penalty using the examples in Figure 4.2. We only consider the tactical shifts removed, using the operator Remove, and reassigned on different days, using SwapRandomDays. We assume that reassignments on the same day, using SwapSameDay, do not negatively impact the work-life balance and, therefore, are not considered when determining the remaining ratio or TS. As we aim to meet demand in the operational schedule, we only determine the number of changes from the original tactical schedule and do not consider the additional shifts added by the operator Add.

In the example in Figure 4.2, the black colored cells indicate the tactical shifts assigned, and we consider a flexibility parameter of 0.2, i.e., 80% of the tactical assignments should remain. In the tactical schedule, we count the number of tactical shifts indicated by TS. In the operational schedule, we count the remaining tactical shifts from the original schedule indicated by RTS. We determine the penalty by the following equation $max(0, 1 - flexibilityparameter - ratioremained)$.

The first example originally had four TS; there are only two RTS in the operational schedule. This results in a ratio of 0.5 remaining shifts, resulting in a penalty of 0.8 - 0.5 = 0.3. In the second example, there are five TS assigned in the tactical schedule and three RTS in the operational schedule, resulting in a ratio of 0.6 and a penalty of 0.2. Both examples did not have an additional tactical shift assigned, and the RTS occurred on the same day.

The third example has two TS assigned. However, in the operational schedule, these are both removed. In the operational schedule, the nurse has one additional shift assigned. Since it occurs on a different day than the original tactical schedule, it does not count as an RTS. Therefore, the ratio remaining is equal to zero, and a penalty of 0.8 is encountered. The last example was not assigned to shifts in the tactical schedule, resulting in a TS of zero. Therefore, the RTS also equals zero, as we only consider the remaining shifts assigned from the original tactical schedule. This instantly results in a penalty of zero, as there is no ratio to determine.

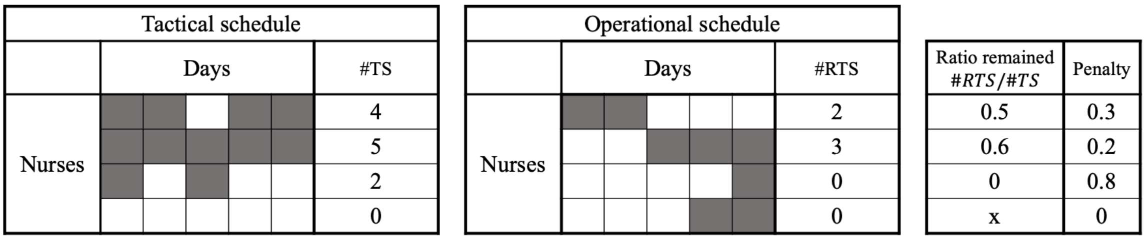


Figure 4.2: Example to determine the penalty for the flexibility of the tactical schedule. From left to right, the tactical schedule, operational schedule, ratio of remaining tactical shifts, and the penalty. The black-colored cells are the tactical shifts assigned. TS represents the number of tactical shifts in the original schedule and RTS is the remaining tactical shifts in the operational schedule.

4.1.4 Objective function

Within the objective function, we aim to minimize the total sum of penalties of the TRCs and ORCs to determine the quality of the solution. The mathematical formulation of the objective function is given in Equation 4.1. For each SC, a weight is assigned based on the priorities of the organization. Since the NSP is case-based and depends on the regulations and priorities of the care organizations, no standard weights can be given for the soft constraints [49]. We determine the weights according to the requirements of the interviewed organizations for scheduling 100% of the contract hours. A smaller weight indicates a lower priority, and vice versa [66]. According to Guericke [66], in a weighted sum approach, the desired metric is combined into a single linear function. So, after defining the priorities, the weight of each part of the objective function is scaled between a range of [0,1], using the softmax normalization as in Equation 4.2. The objective function consists of two weighted sums, p_{nurses} and $p_{organization}$, which individual weights, w_{nurses} and $w_{organization}$, are therefore scaled between [0,1] and sum up to one. Hereby we can determine the trade-off between the penalties for the nurses and the organization. The first part represents the penalties for the nurses and the other for violating the organizational rules, $p_{organization}$, see Equation 4.1.

$$minz = w_{nurse} * p_{nurses} + w_{organization} * p_{organization} \tag{4.1}$$

$$w_{sc} = \frac{w_{sc}}{\sum_{sc \in SC} w_{sc}} \tag{4.2}$$

Equation 4.3 described the penalty for the nurses, p_{nurses}, which is determined by the total sum of the TRC penalties and the fairness metrics used in Lavygina et al. [51], Equation 4.5. This non-linear model is chosen as the results of Lavygina et al. [51] and Martin et al. [50] showed that minimizing the deviation of individual workers' schedule constraint violation penalties is a good fairness measure and results in high-quality fair schedules. Additionally, we want to minimize the highest penalties. Using this objective, higher penalties are penalized more than lower penalties, contributing more to the overall fairness measure. If a linear model had been used, the differences would have been treated equally, and minimizing large disparities or outliers are not prioritized over others.

Each separate part of the p_{nurses} is assigned a weight and normalized, resulting in the sum of $w_{fairness}$ and w_{total} equals 1. The penalties are calculated by determining the nurses' individual penalties for violating the TRC, denoted by p_n^{trc}. This is calculated by Equation 4.4, where p_{trc} is the penalty for violating the corresponding TRC and has a unique weight w_{trc}.

$$p_{nurses} = w_{fairness} * fairness + w_{total} * \sum_{n=1}^{N} \sum_{trc=1}^{TRC} p_n^{trc} \tag{4.3}$$

$$p_n^{trc} = w_{trc} * p_{trc} \tag{4.4}$$

$$fairness = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (\sum_{trc=1}^{TRC} p_n^{trc} - \overline{p})^2} \tag{4.5}$$

where $\overline{p} = \frac{1}{N} \sum_{n=1}^{N} \sum_{trc=1}^{TRC} p_n^{trc}$ is the average penalty.

The second part of the objective function comprises the penalties associated with violating ORCs, each with its own weight w_{orc}. The total penalty $p_{organizational}$ is calculated using Equation 4.6.

$$p_{organisational} = \sum_{osc=1}^{OSC} w_{osc} * p_{osc} \tag{4.6}$$

To ensure a comparable scale of the penalties, the penalties are normalized using min-max normalization. This ensures the penalties have a consistent domain and weighting [66]. A worst-case value is determined for each SC by assuming the likelihood of its occurrence. These are discussed in Section 4.5.

4.2 A two-staged scheduling approach

In this section, we explain the general idea of SA, the constructive heuristic to find an initial feasible solution for the tactical schedule, and how the operational schedule is generated using the tactical schedule as input. Also, the neighbourhood structures used in the SA are explained. A general overview of the process can be found in Figure 4.1.

4.2.1 Simulated annealing

According to Kirkpatrick et al. [67], Henderson et al. [68], Michalewicz and Fogel [69], SA is a temperature-based local search meta-heuristic, which is inspired by the process to simulate the physical crystallization cooling procedure. By allowing hill-climbing moves, i.e. moves that aggravate the objective function, SA provides a way to escape from the local optima to find a near-global optimum [68].

To apply SA, an initial feasible solution is generated and used as the current solution. Hereafter, neighbor solutions are generated by implementing local operators on the current solution. At each iteration, the objective function of the current solution and the neighborhood solution is evaluated. Solutions that improve the objective function are always accepted, whereas worse solutions are accepted with a probability of escaping the local optima. This probability of accepting the worse solution depends on the temperature parameter, T_{start}. At high temperatures, rearrangements causing large changes in the objective function occur due to hill-climbing moves, referred to as diversification or exploration [67, 68]. In contrast, small changes occur at low temperatures where the probability concentrates on the set of locally optimal solutions, also referred to as intensification or exploitation [67, 68]. These local operators change the solution until reaching the desired stopping criteria, T_0. This can be set to a maximum number of iterations, a maximum running time, a minimum temperature level, a solution quality threshold, or when the method converges as it no longer finds better solutions. We have applied the minimum temperature level as a stopping criterion. Because when the minimum temperature is almost reached, the SA should focus on intensification rather than diversification and converge to a near global optimum. In this research, the goal is to decrease the penalties associated with violating soft constraints within the objective function. The method described is the same for all instances evaluated in this research. The basic algorithm of SA can be found in Algorithm 1. The input parameters considered are the start temperature T_{start}, stopping criteria T_0, the number of iterations for each temperature denoted by MCL, and the factor that decreases the temperature after the number of iterations is reached α.

The starting temperature T_{start} is based on the objective value of the initial solution to provide an instance-based value, as proposed by Ropke and Pisinger [70]. In their approach, they determine T_{start} such that a solution that is $w\%$ worse than the initial solution is accepted with a probability of 0.5. The only parameter that has to be set is w, which is denoted as the start temperature control parameter. In this research, w is set equal to 0.8, indicating that a solution that is 1.8 worse than the initial solution is accepted. This result in Equation 4.7, as used in Guericke [66].

$$e^{-\frac{(1.0+w\%)x_0-x_0}{T_{start}}} = 0.5 \Leftrightarrow T_{start} = \frac{(1.0+w\%)x_0 - x_0}{-ln(0.5)} \tag{4.7}$$

Algorithm 1: Simulated Annealing [Henderson et al. [68], Michalewicz and Fogel [69]]

Input : Initial Solution x_0, start temperature T_{start}, stopping criteria T_0, Markov chain length MCL, decrease factor α

Output: Best solution x^*

$T \leftarrow T_{T\,start}$, $m \leftarrow 0$, $x \leftarrow x_0$, $x^* \leftarrow x_0$;

while $T > T_0$ **do**

  foreach m *in* MCL **do**

    $x_n \leftarrow$ select a random neighbour solution from $N(x)$;

    if $objective(x_n) < objective(x)$ **then**

      if $objective(x_n) < objective(x^*)$ **then**

        $x^* \leftarrow x_n$

      end

      $x \leftarrow x_n$

    end

    else if $random[0,1) \leq e^{\frac{objective(x)-objective(x_n)}{T}}$ **then**

      $x \leftarrow x_n$

    end

  end

  $T \leftarrow \alpha{*}\mathrm{T}$

end

return x^*

4.2.2 The initial feasible solution

To generate an initial feasible solution, all hard constraints should be met. In this research, we use a constructive heuristic to generate the initial solution; the pseudocode can be found in Algorithm 2. The constructive heuristic follows the following steps.

First, the priority levels of the shifts are determined, which divides the shifts into different sets. Based on the interviews with the care organizations, these priorities are determined and represent the shift sequence. An example can be found in Table 4.1. For each set of shifts, the shifts are sorted in descending order based on the shift duration. As in practice, first, the long shifts are assigned. Hereafter, the days are sorted based on the highest demand. This is because the dates with the highest demand are more likely to result in undercoverage. By sorting the days, we try to prevent this and assign nurses more efficiently. Then, while not all shifts of the set have been checked if they can be assigned, search for a shift with a date corresponding to the highest demand. If no shift has been found, we go to the next date with the highest demand. Otherwise, the selected shift is assigned to one of the available nurses. A nurse is available when assigning the shift, and none of the hard constraints described in is violated. In addition, we determine if the nurse has enough contractual hours left if the shift is assigned. If all requirements are met, the shift is assigned to the nurse. The demand of the corresponding date is updated, and the dates are again sorted based on the remaining demand. This procedure is repeated until all shifts in the set and priority levels have been checked.

Algorithm 2: Constructive heuristic for the initial solution

```
Input  : The set of shift with priority level that must be assigned
Output: Initial feasible solution
foreach Priority level do
    Get the set of shifts with the corresponding priority level;
    Sort the set of shifts based on the shift duration;
    while not all shifts checked do
        sort the days based on the highest demand
        foreach day in the planning horizon do
            date ← the first date with the highest demand;
            shift ← find a shift with the same date;
            if no shift is found then
                go to the next date with the highest demand;
            else
                set the status of the shift as checked;
                check if the shift can be assigned to a nurse Appendix B Algorithm 3;
                if Assigned is True then
                    update the list with assigned shifts for the date;
                    break;
                else
                    update the list with not assigned shifts for the date;
                end
            end
        end
    end
end
return Initial solution
```

Table 4.1: Example priority levels of the shifts for the constructive heuristic based on the shift sequence explained in Section 2.2.2

Priority level	**Shift type**
1	Night shifts level 3
2	Night shifts
3	Day shifts level 3
4	Weekend shifts
5	Remaining shifts

4.2.3 From tactical to operational schedule

The initial solution is used as input in the SA algorithm to optimize the tactical schedule. Hereafter, to generate the operational schedule, the planned absenteeism is removed from the optimized tactical schedule. The procedure to include the planned absenteeism from the year planning is visualized in Figure 4.3. For each nurse and assigned tactical shift, we determine if there is an overlap with the planned absenteeism. If so, we remove the shift from the nurses' tactical schedule and update the remaining minutes to assign, the number of assigned shifts and free days.

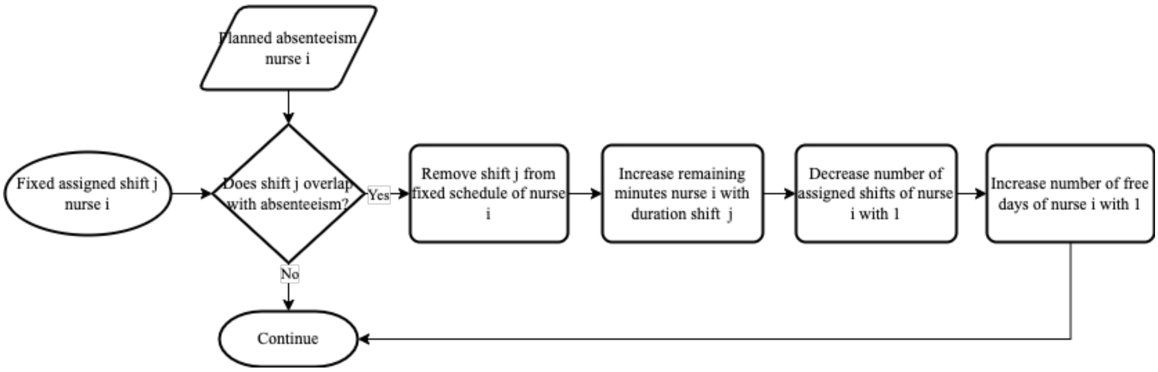


Figure 4.3: Procedure planned absenteeism and tactical assigned shifts

4.2.4 Neighbourhood structures

To find a good-quality solution, the structures of the local operators are important [57]. To search for a feasible solution, four neighbourhood structures are implemented. The structures applied are *Remove*, *Add*, *SwapSameDay*, *SwapDifferentDay* and can be found in Figure 4.4. The structures will be briefly explained. At the beginning of the procedure of each neighbourhood structure, it has been checked that either there are nurses that have shifts assigned or that there are days with unassigned shifts. To apply the two swap operators, a minimum of two nurses should have shifts assigned on either the same or different days. For the *Remove* operator, there must be one nurse that has a shift assigned, and for *Add*, there must be one shift that is not assigned.

Neighbourhood structure: Remove

The neighbourhood structure *Remove* first checks if there are nurses with assigned shifts. If so, a random nurse n and a random shift s are selected. The day d is determined from shift s. Then, shift s is removed from the list of assigned shifts of nurse n. Also, it is removed from the list of assigned shifts with the corresponding date d and appended to the list of not assigned shifts with date d. As no constraint restricts the possibility to sign-off a shift, no constraints have to be checked. The structure is shown in Figure 4.4A.

Neighbourhood structure: Add

Within the neighbourhood structure *Add*, we try to assign an extra shift to a random nurse. When there are shifts to assign, a random day d and a random shift s are selected. To check if a nurse can be assigned to this random shift, we create a list that stores the nurses' id n_{id} in random order. Next, we loop over the list with nurses N and check for each nurse n if the shift can be assigned based on the HC1-HC4. If the shift s can be assigned, we add the shift s to the list of the nurse n, and the procedure will then terminate. Opposite to the neighbourhood Remove, we add the shift s to the list with the assigned shifts of the corresponding day d and remove it from the list that holds the not assigned shifts with day d. This procedure that attempts to assign an extra shift is repeated until a nurse n has been found that can be assigned the shift or the last nurse in the random list also cannot be assigned. The structure is shown in Figure 4.4.B.

Neighbourhood structure: SwapSameDay

Two nurses, n_1 and n_2, are randomly selected. They work shifts s_1 and s_2 respectively on the same day d. For shift s_1 we check if nurse n_2 can be assigned to the shifts without violating HC2-HC4 and would not result in overtime. We do not check HC1 as the nurse has still shifted s_2 assigned, which would result in an infeasible solution. However, we do check HC2 as the time of the shift might differ and can conflict with the absenteeism of nurse n_2. We do the same for shift s_2 and nurse n_1. If the neighbourhood solution is feasible, the shifts are exchanged. Hereafter, the list of assigned shifts for the nurses is updated. In this case, we do not have to adapt the list with assigned and unassigned shifts for the day, as the amount of assigned shifts remains the same. The structure is shown in Figure 4.4C.

Neighbourhood structure: SwapDifferentDay

Randomly select two nurses n_1 and n_2 and two shifts s_1 and s_2 with different days d_1 and d_2. In addition to the procedure of neighbourhood structure *SwapSameDay*, HC1 is also verified to have a feasible solution. If so, the two shifts s_1 and s_2 are exchanged with the nurses. The structure can be found in Figure 4.4D.

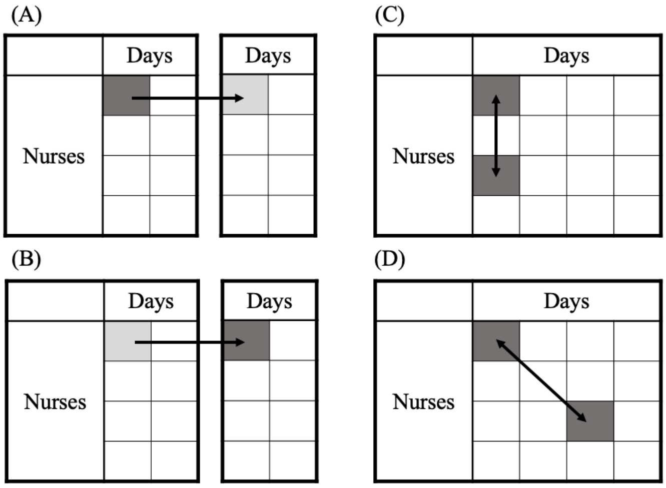


Figure 4.4: Neighbourhood structures applied in the simulated annealing algorithm. **A** Neighborhood structure Remove: sign a random shift off from a random nurse. **B** Neighborhood structure Add: assign a random shift to a random nurse. **C** Neighbourhood structure SwapSameDay: swap two random shifts on the same day with two random nurses. **D** Neighbourhood structure SwapDifferentDay: swap two random shifts with different days with two random nurses

4.2.5 Probability operators

In each iteration of the SA algorithm, an operator is chosen with a certain probability. This probability depends on and is adjusted based on the number of successes. This procedure is similar to the approach of Adaptive Large Neighbourhood Search (ALNS), which is introduced by Ropke and Pisinger [70]. They use a roulette wheel selection principle to choose the neighborhood heuristic for the iteration. Weights are assigned to each heuristic and influence the probability of being chosen. In their research, they keep track of the number of successes of the heuristic to determine the weight.

We also use this principle to determine the next neighborhood structure in the SA algorithm. The method changes the probability distribution in each iteration based on the success rates of the operators. In the beginning, the probabilities are uniformly distributed. During the procedure, the successful operators are given higher probabilities. Within the method, a smoothing factor and a default probability are used in order to ensure that during the procedure still, each operator is selected.

4.3 Quantification of the performance

To quantify the performance of the method, we introduce the following quantitative KPIs: the number of free days, the number of night shifts, the number of working weekends, the amount of over or undertime, the number of requested fixed and additional free days, total undercoverage, missing hours QL3, and the number of law violations. An overview can be found in Table 4.2, where the KPIs are divided based on the nurses' and the care organization's points of view.

According to the interviews with the care organizations, nurses pay more attention to a fair distribution of shifts, assigned requests, and distributed hours. The management of the organization or the planner focuses on the undercoverage per day in terms of demand and the coverage of QL3. Also, depending on the scheduling phase, the number of law violations needs to be minimized, which includes the RT, FRO, CWD, CNS, EOW, 2W, and the additional SFS.

Table 4.2: Overview of the quantitative measurements for personnel and organization point of view

KPI nurses	KPI organisation
Number of night shifts	Undercoverage week
Number working weekends	Undercoverage weekend
Distribution +/- hours	Undercoverage night
Requested fixed free day	Underqualified shifts
Requested additional free day	Missing hours QL3
	Number of law violations

4.4 Assumptions for the method

To use a mathematical model to solve the NSP, some simplifications and assumptions are made. The following assumptions are considered in this research:

For the shift assignment in the tactical and operational schedule, we consider a four-week planning horizon, i.e., 28 days. We assume that there is no previous or upcoming scheduling period for this problem. So, we do not consider the working days in the previous planning horizon when assigning the shifts to the nurses. Also, we do not take the first and last week into account when generating the tactical schedule.

Considering the QL3 coverage in HC3 and the QLs in SC12, we make the following assumptions to determine if a nurse has the proper QLs. Each nurse and shift is associated with a set of hour types. These are defined based on the level of care they may deliver and are required, e.g., basic care and specialized care. Nurses with a higher QL are assigned more specialized hour types or a larger amount of hour types. For example, based on the data analysis of one of the case studies, it is seen that hour types associated with shifts requiring QL3 have a list length of five. The other shifts have a length of three, all with the same or no required hour types. Hence, to determine if a nurse is over- or underqualified, we compare the length of the shifts' required hour types with the length of hour types assigned to the nurse. For HC3, only nurses with the minimum length can be assigned to the QL3 shifts. For SC12, a nurse that is underqualified can be assigned to the shift, but a penalty is encountered.

To calculate SC11, we assumed that the start of the day is at 7:00 AM. This is because the earliest shift starts at 7:00 AM, and the latest shift ends at 7:15 AM the following day. If we verify for each day that the time range from 7:00 AM on the current day to 7:00 AM on the following day, is covered by QL3 shifts, then we can be certain that there is a 24-hour QL3 coverage.

For simplifications, we do not implement the exceptions that can be made on the law legislation in the operational schedule as stated in Section 2.3.

4.5 Normalization of the penalties

Normalization is used to compare the penalties on the same scale for the TRCs and ORCs. The normalized value is calculated based on the worst-case value dependent on the constraint, which is used as the maximum value in the min-max normalization. This results in a normalized penalty in the interval [0,1], where 1 indicates the worst scenario. However, some worst-case scenarios are unlikely to occur. Therefore, assumptions are made to determine the worst-case and, thus, the maximum values of the constraints. An overview of the maximum values can be found in Table 4.3. The penalties for the RM and the RND are already within a range of 0 and 1. The penalties for the other constraints are in seconds.

For the TRC, the following assumptions are made regarding RT, FRO, CWD, CNS, EOW, 2W, and SFS.

The maximum value for RT is equal to the required RT between shifts, which is eleven hours, multiplied by the total maximum CWD. As we aim to minimize the maximum allowed CWD, it is unlikely that nurses are encountered a penalty of missing RT, which can occur due to more than five days. However, as a maximum RT of 55 hours is unlikely, we have set the maximum RT equal to the RT times three days. Additionally, the maximum value of the FRO is set equal to the difference between the earliest and latest starting time since the worst-case scenario is that a nurse gets a FRO penalty due to a night shift followed by a morning shift multiplied two. The maximum value of the CWD and CNS is determined by multiplying the maximum allowed days or shifts by two, as it is unlikely that these rules are violated for this amount of time. Hereafter, we multiplied by the rest hours that should be assigned after working the number of consecutive days or shifts.

As shown in Appendix A Figure A.3.B, the worse case for the weekend shifts is that a nurse gets shifts assigned each weekend. For penalty 2W, the maximum penalty is four shifts times the maximum shift duration and is used as the maximum value. For EOW, at most, three weekends can be assigned. Therefore, the maximum value for this constraint is three times the maximum shift duration.

Normalizing the values SC7, the forbidden pattern, the following maximum value is used. We multiply the number of days in a week by the number of hours in this week. As the SFS penalty is determined by the time between the two shifts that violate the pattern, the gap is larger or equal to a day. However, as the other soft constraints limit number of shifts assigned to a nurse, it is unlikely that this pattern will occur over the total planning horizon, which would be half of the length of the planning horizon. Therefore, we assume that the maximum value is equal to half of the total hours within the planning horizon, i.e., fourteen days in hours.

For the organizational constraints, we normalize the values for all constraints. For the total coverage during the week, weekend, night, and QL3, we determine the total demand in minutes for each instance. The worst case is that none of the shifts are assigned, resulting in a maximum value equal to the total shift duration. The worse case for assigning underqualified nurses on a day is that all shifts are assigned to these nurses, so the total demand in minutes times two is used as the maximum value for normalizing this penalty.

Table 4.3: Overview of the maximum values of the time-related soft and organizational soft constraints. RT: rest time, FRO: forward rotating order, CWD: consecutive working days, CNS: consecutive night shifts, EOW: every other weekend, 2W: maximum 2 weekends, RM: remaining minutes, SFS: forbidden pattern, RND: difference ratio night day.

TRC	**Max**
RT	118,800 sec.
FRO	54,540 sec.
CWD	1,296,000 sec.
CNS	993,600 sec.
EOW	1,530 min.
2W	2,040 min.
RM	1
SFS	1,209,600 sec.
RND	1

Table 4.4: The maximum values used for normalizing the organizational penalties.

ORC	**Max.**
Weekday	Weekday demand in minutes
Weekend	Weekend demand in minutes
Night	Night shifts in minutes
QL3	QL3 demand in minutes
UQ	Max. demand day_x in minutes * 2

4.6 Conclusion

This chapter describes the proposed method to solve the scheduling problem for residential care organizations. First, the four hard and eleven soft constraints used for constructing the tactical schedule and the additional hard and soft constraints for constructing the operational schedule are briefly explained. The objective function is given, which aims to minimize the total sum of penalties while accounting for fairness by including it as an objective that minimizes the deviation of individual penalties. Next, the two-stage scheduling approach is explained. This consists of a constructive heuristic to find an initial solution for the tactical schedule, followed by a SA algorithm that optimizes the initial solution. Hereafter, the absenteeism from the year planning is removed from the optimized tactical schedule, and the remainder is used as input for the SA algorithm to construct the operational schedule. Within the SA algorithm, four operators are implemented to find neighbour solutions. Furthermore, the assumptions regarding the method are described, regarding that there are no previous or upcoming schedules, QL3 identification and determining the QL3 coverage, and the exceptions on law legislations. At last, the assumptions for normalizing the penalties are provided, to be able to compare the penalties on the same scale.

Chapter 5

Case studies

The goal of the NSP is to assign shifts to nurses to satisfy the demand, while individual preferences and law legislation are taken into account. As mentioned, this research discusses the static version of the NSP, where all information is known at the beginning of the planning horizon. To assess the performance of the proposed method, we use data from three case studies. These are different size residential care organizations in the Netherlands. Within these organizations, the shifts are assigned manually by the planner. Besides this being a time-consuming task, it is also aimed at finding a feasible schedules without focussing on optimality [57]. This chapter describes the key elements of the different case studies used in this research. First, we will discuss the assumptions regarding the data. Hereafter follows the current performance of the manually-created schedules.

5.1 Data analysis and Assumptions

Before executing the method, data analysis is performed. Data on the nurses, historic tactical schedule, and expertise are collected using the software of Nedap.

For the nurses, this consists of the *ID*, *Age*, *Contract hours*, *Qualified hour types*, and *Absenteeism*. The absenteeism in the data set consists of different absenteeism types, e.g., *Not Available*, *Sick*, *Holiday*, *Education*, and *Meetings*. Within the data set, there are nurses without contract hours; these are assumed to be intra-organizational flex nurses who are not assigned to shifts in the tactical schedule. Depending on the age, a nurse is allowed to work a night shift. If the data set does not contain an age, it is assumed that this nurse has the correct age to be assigned to a night shift. Furthermore, it is assumed that Not Available represents the agreed fixed free days for the tactical schedule. The remaining types of absenteeism are included when generating the operational schedule, as these are part of the year's planning.

The following data is collected from the shifts: *Original assignment*, *Shift type*, *Qualifcationl level*, *Start and end time*, and *Night shift (NS)*. In the historical schedule, a shift is assigned to one or multiple nurses. We assume that a shift can be assigned to one nurse for simplification. The shift type is used to indicate the number of times the shift reoccurs in the planning horizon of the historical schedule, e.g., shift types are required every Monday, or a shift type occurs once in the total planning horizon. At last, the QLs are associated with a value and hour types, with the last representing the minimum required QL of a shift. Based on the value, a distinction is made between QL3 shifts and other shifts.

The software allows for specifying the start and end times of shifts for each day of the week but not for alternating weeks or one-time shifts. In practice, there are situations where a shift is created for, e.g., onboarding shifts that remain unassigned 90% in the planning horizon or two shifts that alternate depending on the week. These shifts are present in the data but do not have an assignment. We remove the shifts that do not have an assignment in the historical schedule to prevent the algorithm from assigning these unnecessary shifts. Additionally, it can be the case that a shift is assigned in the historical operational schedule but is not assigned in the historical tactical schedule, e.g., a last-minute shift that is added. Therefore, a threshold is introduced for the number of times a shift type must occur in the historical data to be allocated in the tactical schedule. Shifts that do not meet this threshold are scheduled in the operational schedule.

Furthermore, data analysis showed that each case study classifies the required 24-hour qualification level differently. For simplification, we further refer to QL3 as the qualification level that needs to be present 24 hours for all case studies. Within the method, the qualification level and hour types of each case study are analyzed to identify those that correspond with QL3. Additionally, we determine if the QL3 shifts within the dataset are sufficient to cover the whole day. If this is not the case, we determine the missing hours and correct this when calculating the QL3 coverage constraint.

5.2 Description case studies

The case studies used in this research each represent a Dutch residential care organization that used the software application of Nedap to generate nurse schedules. The schedules are generated for a 4-week, i.e., 28 days, planning horizon. In the tactical schedule, we only consider nurses assigned with contract hours in the data set. Whereas in the operational schedule, we also include intra-organizational flex nurses. The case studies are categorized based on the total number of shifts and the number of regular nurses, which results in a small, medium, and large case study. Within Table 5.1, the following is included: the demand in the number of shifts and minutes for the total demand, including the QL3 shifts and the separate QL3 shifts; the number of available nurses and the minutes where the tactical schedule only includes nurses with contract hours and the operational also the flex nurses and the number of QL3 nurses; and the ratio of available regular nurses in minutes and the demand in minutes, where for the tactical schedule 80% of the available minutes is considered. This ratio indicates the flexibility to meet the demand. A large ratio implies more flexibility in assigning shifts, and a ratio below zero indicates the the number of nurses is insufficient, making it impossible to meet all demands.

Based on the interviews, we adapt the priority sequence of shifts for the constructive heuristic for each case study. An overview of the total number of shifts and nurses is provided in Table 5.1.

Small case study

The first case study is a relatively small organization compared to the two other case studies. In total, 125 shifts must be assigned in the tactical schedule, where 66 are QL3 shifts. There are sixteen nurses to assign, with eleven QL3 nurses, making the distribution of QL3 shifts more flexible than the large case study. For the operational schedule, two more nurses are available who have both QL3. In the original tactical and operational schedules, six and seven nurses are not assigned to shifts, respectively. There are three types of absenteeism to consider, where the tactical schedule only considers the Not available. The ratio of available minutes for total demand and QL3 is almost twice or three times as large, indicating high flexibility in assigning the shifts. The earliest shift starts at 7:00 AM, and the latest ends at 11:00 PM. Hence, in this case, there are also no night shifts within the data set. Therefore, none of the days can be fully covered by QL3. The demand is equal for the tactical and operational schedules. This case study has the following shift sequence: the weekend shifts with QL3, the remaining weekend shifts, weekly QL3 shifts, and the remaining shifts.

Medium case study

As seen in Table 5.1, the medium case study has 252 shifts to assign for the tactical schedule and 296 shifts in the operational schedule. In total, there are 24 full-time nurses available in the tactical schedule, with fourteen of them having QL3. For the operational schedule, there are 72 nurses available, and a total of 40 have QL3. The amount of contract hours is equal for the tactical and operational schedules, as there are only additional flex nurses. Within the original tactical and operational schedules, seven and ten nurses are not scheduled, respectively. The ratio is smaller compared to the small case study. However, there are almost twice as many minutes available than the demand, indicating that there is flexibility in the shift assignments. The data analysis showed that the earliest shift starts at 7:00 AM and the latest ends at 7:15 AM, meaning there are night shifts within this data set. Further, two days cannot be covered 24 hours as there are not enough QL3 shifts. There are eight absenteeism types for the tactical and operational schedules, where only Not Available is included in the tactical schedule. Within the operational schedule, all eight types are considered when assigning shifts. The sequence of scheduling is the QL3 weekly night shifts, remaining weekly night shifts, QL3 weekend shifts, remaining weekend shifts, and remaining shifts, respectively.

Large case study

The large case study has 424 shifts that must be assigned in the tactical and operational schedules. For the tactical and operational schedules, 39 and 43 nurses are available, respectively. However, as seen in Table 5.1, only two nurses have the required QL3 level, making the distribution of QL3 shifts less flexible compared to the other two case studies. This must be considered when evaluating the penalty for the QL3 coverage. This is also reflected in the ratio of the tactical schedule. As this is below one, it is impossible to meet the requirements of 24-hour QL3 coverage. The total number of contract hours remains the same for the tactical and operational schedules, as the additional nurses are flex nurses. Within the original tactical schedule, seven nurses, and in the operational schedule, eleven nurses are not scheduled. The remaining ratios are close or equal to one, indicating that there is only little flexibility in assigning the shifts. The earliest starting time of the shifts is 7:00 AM, and the latest end time is 11:00 PM, meaning that no night shifts need to be scheduled. Absenteeism types considered in this case study are Not Available, Meeting, and Education, where only the first is considered in the tactical schedule, as the others differ per period. The demand remains the same for both schedules. The sequence of shift scheduling is as follows: first, the QL3 is scheduled for the week, followed by another defined QL. Hereafter, the weekends and the remaining shifts are assigned.

Table 5.1: Data on the case studies for the tactical and operational schedule. The demand in minutes also includes the required minutes associated with QL3. The ratio is available minutes divided by the demand in minutes. For the tactical schedule, 80% of the minutes is taken into account and 100% of the contract hours are in the operational schedule.

	Small		Medium		Large	
	Tactical	Operational	Tactical	Operational	Tactical	Operational
Shifts						
Total number of shifts	125	125	252	296	424	424
Total minutes needed	39,960	39,960	114,900	136,290	171,360	171,360
Total number of QL3 shifts	66	66	153	166	32	32
Total minutes QL3 shifts	23,700	23,700	73,260	79,110	15,720	15,720
Nurses						
Total number of nurses	17	19	24	72	39	43
Total available minutes nurses	99,120	99,120	272,160	272,160	216,740	216,740
Total QL3 nurses	12	14	14	40	2	2
Total available minutes N3	69,480	69,480	178,080	178,080	17,280	17,280
Ratio						
Available minutes/demand shifts	1.98	2.48	1.89	2.00	1.01	1.26
Available QL3/demand QL3	2.36	2.95	1.94	2.25	0.88	1.10

5.3 Current performance

For each case study, the current performance of the manual schedule is evaluated using the proposed method and objective function. This gives insight into, e.g., the number of law violations, if there is 24-hour QL3 coverage, and if there are UQ shifts. Figure 5.1 provides a visualization of an original schedule where we have used the large case study as an example. The outcomes of the performance can be found in Table 5.2. It shows the number of TRC and ORC violations within the generated schedule from the practice of the three case studies. For the small and large case studies, we removed the penalties CNS, as there are no night shifts within the data set. The values represent the number of violations for all nurses, e.g., the medium case study has one RT violation in the tactical schedule. As seen in Table 5.2, only the medium case study is missing QL3 coverage. Furthermore, according to the proposed method, each case study has assigned shifts to nurses that do not meet the coverage requirements associated with the shift resulting in violations for UQ. Also, as can be seen, the large case study has a high number of SFS violations, as can also be seen in the example of the original schedule in Figure 5.1.

Table 5.2: Soft constraint violations for the manually generated tactical and operational schedule for the three case studies. RT: rest time, FRO: forward rotating order, CWD: consecutive working days, CNS: consecutive night shifts, EOW: every other weekend, 2W: maximum two weekends, SFS: on-off-on pattern.

	Small		Medium		Large	
	Tactical	Operational	Tactical	Operational	Tactical	Operational
TRC						
RT	13	15	1	7	2	2
FRO	9	11	3	10	2	2
CWD	0	0	2	8	0	0
CNS	-	-	5	5	-	-
EOW	0	0	5	11	0	0
2W	0	3	7	11	0	0
SFS	16	16	19	29	98	98
ORC						
Missing QL3 hours	0	0	100.25	0	0	0
Week	18	0	68	0	0	0
Weekend	3	0	28	0	0	0
Night	-	-	5	0	-	-
UQ shifts	0	0	1	15	8	8

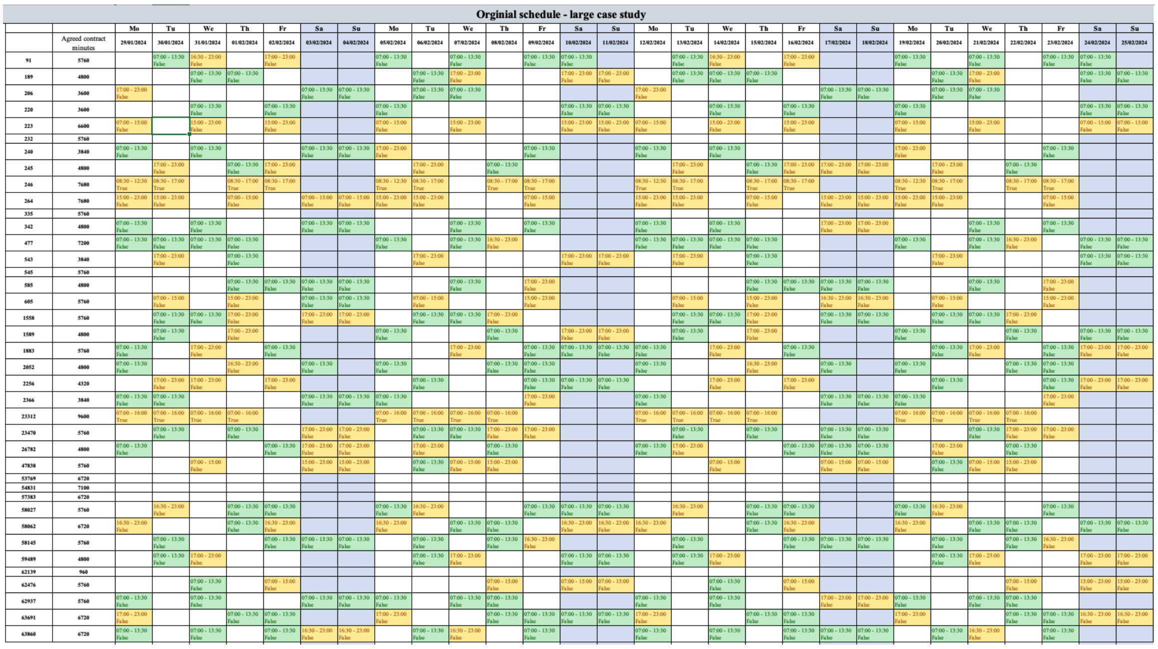


Figure 5.1: Example of the original schedule of the large case study. In green are the morning shifts from 7:00 AM until 1:30 PM, and in orange are the shifts from 1:30 PM until 11:00 PM.

5.4 Conclusion

This section provided a brief description of the three case studies that are used the evaluate the performance of the proposed method. First, the data analysis is performed, and the assumptions regarding the shifts and QLs are explained. It appeared that only the medium case study has night shifts that need to be assigned. Also, the ratio of available minutes and the demand in minutes is determined. It resulted that the large case study does not have a sufficient number of nurses to cover all QL3 shifts in the tactical schedule when scheduling for 80% of the contract hours. Shift assignments in the other two case studies are more flexible, as the ratio is close or higher than two, indicating that there are twice as minutes available than the required hours. Finally, the number of TRC and ORC violations of the current performance are described. The number of violations differs per organization, which indicates that the organization has different priorities regarding fulfilling the TRCs and ORCs. It must be pointed out that none of the organization's operational schedule has undercoverage during the week, weekend, or night if possible, and no missing QL3 hours. However, UQ shifts are assigned in both schedules for all three case studies.

Chapter 6

Model Performance

In this chapter, we present the experiments conducted in this research. First, we will present the parameter tuning process for the SA algorithm. The best parameters are used in further experiments to validate the model's performance. Next, we compare the performance of the proposed method with that of current practice for all three case studies. Hereafter, we analyze the flexibility parameter for the proposed method. Followed by a sensitivity analysis of the w_{TRCs} for all three case studies and a sensitivity analysis by relaxing the weight for the flexibility using only the medium case study. Based on these experiments, we want to identify which constraints should be considered when automating the scheduling process and the weights that should be implemented when optimizing the tactical and operational schedules. The proposed SA algorithm is implemented using Python language on a Mac OS Ventura (64-bit) operating system, intel Core i9 2.3 GHz CPU and 32 GB of RAM. The three case studies from Chapter 5 are used to evaluate the proposed SA algorithm.

The parameter values for the SA algorithm are selected based on scheduling 80% of the contract hours in the tactical schedule. In accordance with the preferences of the interviewed planners and the first few runs, we have set the weight for the TRC and ORC equal to those in Table 6.2 to find the best parameter values. These parameter values are used in the remaining experiments: to determine the trade-off between the penalty for the nurses and the organization; in the flexibility analysis in Section 6.4; and the sensitivity analysis in Section 6.5. We determine the trade-off in order to define the focus of the objective function on what should be optimized regarding the tactical and operational schedule for the three case studies by increasing w_{nurses} and decreasing $w_{organization}$. Subsequently, these settings are used to examine the optimal flexibility of the tactical schedule in order to find a valid operational schedule.

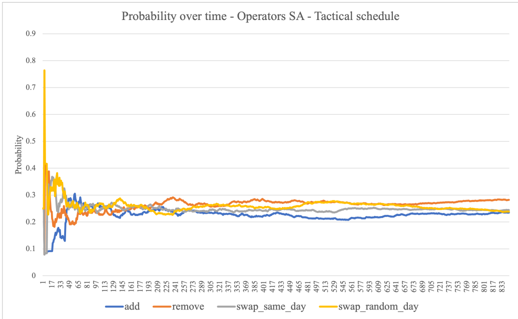
6.1 Parameter tuning simulated annealing

Since the parameter settings of the SA affect the results, we will test the proposed SA for different parameter values for each case study. We test the following parameter values: $\alpha = 0.80$, 0.90 and 0,99; $MCL = 1$, 100, and 1000; $T_0 = 0.01$, 0.001 and 0.0001. The outcome of the SA algorithm may vary in each experimental run, so each combination is repeated five times to account for randomness. This is done for each of the five experimental runs conducted in this research. For the parameter tuning phase, the maximum running time is set to five minutes, and we use an equal distribution of w_{nurses} and $w_{organization}$. The best results are in Table 6.1. The overall outcome and graphs of the SA can be found in Appendix C.1. The same best parameter settings are further used in the SA algorithm for optimizing the operational schedule.

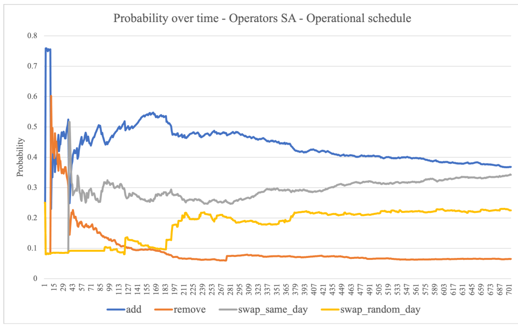
Table 6.1: Best parameter values for the different case studies after parameter tuning for the three case studies.

	Parameters			
Case study	T_{start}	T_0	α	MCL
Small	0.1169	0.0001	0.8	100
Medium	0.3199	0.0001	0.99	1
Large	0.6765	0.0001	0.99	1

To find neighbour solutions, four operators are used. As explained in Section 4.2.5, an ALNS approach is used where the number of successes determines the probability of an operator being chosen. The outcomes for the SA using the medium case study can be found in Figure 6.1. As can be seen, at the beginning of the procedure, the operators *SwapRandomDay* and *SwapSameDay* are chosen more often as they lead to better solutions for the tactical schedule. Ultimately, the probability of those two operators and *Remove* and *Add* converge to the same probability. On the contrary, in the operational schedule, the operator *Add* has a higher probability at the beginning and end of the procedure, as seen in Figure 6.1b. Where *Remove* has the lowest chance of being chosen at the end of the procedure. From the two *Swap* operators, swapping on the same day is preferred above swapping on random days. This can be explained due to the fact that in the operational schedule, the goal is to meet the coverage requirements. Since shifts are removed that conflict with absenteeism, there are more open shifts that need to be assigned. If the operator Add is used, the objective value will improve as there is less undercoverage, and visa versa for selecting the operator Remove. Additionally, the SwapSameDay is preferred over SwapRandomDay, as the flexibility parameter allows swapping shifts on the same day and will therefore result in a better objective value than when choosing the latter.



(a) Probability operators for tactical schedule.



(b) Probability operators for operational schedule.

Figure 6.1: Probability operators for both tactical and operational schedule for the medium case study

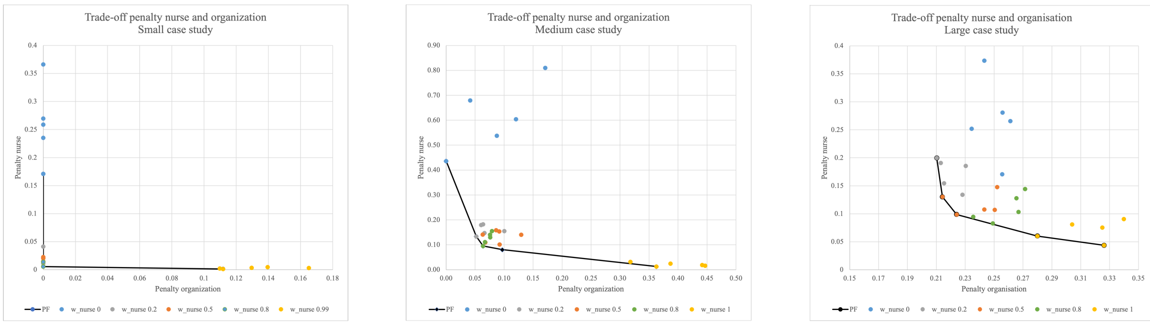
6.2 Trade-off penalty nurses and organization

Before executing the algorithm, the appropriate weights for the objective value in Equation 4.1 need to be determined. Therefore, we analyze the trade-off between the penalty encountered by the nurses, p_{nurses}, and the organization, $p_{organization}$. These will be determined based on the outcomes of the tactical schedule for each case study using the best values of α, MCL, and T_0. To evaluate the trade-off between the two penalties, different weights are assigned to w_{nurses} and $w_{organization}$. The algorithm constructing the tactical schedule is again executed five times to account for randomness. In each run, a new tactical schedule is constructed, and the values for p_{nurses} and $p_{organization}$ are determined. The results for the trade-off are presented in Figure 6.2, which shows the results of the five different experiments and the Pareto frontier.

According to Ngatchou et al. [71], the Pareto frontier shows the set of acceptable trade-off solutions where the most desirable solution is selected from the Pareto set. A solution consists to the Pareto set if there is no other solution that can improve at least one of the objectives without degrading the other [71, 72]. The concept of Pareto dominance and Pareto optimality is used to compare solutions. The latter is used if and only if there does not exist another solution that dominates it. The set of all optimal solutions is called the Pareto optimal set and forms the frontier, as shown in Figure 6.2.

As mentioned in Section 2.2.1, the goal of the tactical schedule is to create a high-quality individual schedule that meets the nurse's preferences. In contrast, the operational schedule aims to minimize organizational violations and thus open shifts by assigning them to either team members or extra-organizational flex employees while considering the nurses' preferences. Based on these objectives, the experimental results, and the Pareto frontier, we have determined the appropriate weights for w_{nurses} for each case study's tactical and operational schedule. The weight for $w_{organization}$ equals $1 - w_{nurses}$. The weights for the other constraints are set to the values specified in Table 6.2.

The resulting values for w_{nurses} for the tactical and operational schedule are as follows: 0.95 and 0.8 for the small case study, 0.95 and 0.6 for the medium case study, and 0.85 and 0.4 for the large case study. A brief explanation is provided for each case study according to the results in Figure 6.2. The differences between the three case studies can be explained by the difference in the number of available QL3 nurses, regular nurses, and the demand for shifts, which will also be pointed out in the following.



(a) Trade-off small case study (b) Trade-off medium case study (c) Trade-off large case study

Figure 6.2: Trade-off between p_{nurses} and $p_{organization}$ for different runs using the best parameters for the SA from Section 6.1 for the three case studies.

Small case study

As can be seen in Figure 6.2a, all w_{nurse} smaller than 0.99 results in a $p_{organizational}$ of 0 for the small case study. In addition, the individual experimental results of all weights, except for w_{nurses} of 0, lie close to the Pareto frontier. This can be due to the high ratio of available minutes compared to the total needed minutes and the fact that many nurses possess a QL3. This allows for an efficient shift assignment with minimal TRC and ORC violations. We choose w_{nurses} of 0.95 and 0.8 for the tactical and operational schedule, respectively, to minimize the penalty for both the nurse and the organization, with a higher focus on the nurses in the tactical schedule.

Medium case study

Figure 6.2b shows the medium case study's experimental results and the Pareto frontier. The results for weights 0.2, 0.5, and 0.8 all lie close to each other, whereas the results w_{nurse} equal to 0 deviate a lot. Choosing a value of 1 for w_{nurse} results in minimal values for p_{nurses} but in higher values for $p_{organization}$. As the experimental results of w_{nurses} equal to 0.8, lies close to the Pareto frontier and result in relatively small p_{nurses}, we choose a w_{nurses} of 0.85 for the tactical schedule as we want to minimize the amount of TRC violations. For the operational schedule, we choose a w_{nurses} of 0.6, as choosing a weight closer to 0 is less efficient than the other weights, as the distance of the dominated solution to the frontier is larger than those of the other experiments.

Large case study

As seen in Figure 6.2c, the Pareto frontier of the large case study consists of points that result from the experiments when choosing a w_{nurses} of 0.2, 0.5, 0.8, and 1. The experimental results with w_{nurses} equal to 0 results in feasible but not efficient results, which are referred to as dominated solutions. Choosing a weight of w_{nurses} equal to 0 results in an inferior solution compared to the other weights. It must be noted that the x-axis starts at 0.15, as none of the experiments resulted in a solution without ORC violations. This can be the case as only two QL3 nurses are available or due to the assumptions made regarding the hour types. Additionally, the large organization has the smallest ratio of available minutes and total minutes needed compared to the other two case studies. This makes it more challenging to assign the shifts efficiently, making it more difficult to cover all the required shifts while meeting minimizing TRCs violations, resulting in undercoverage.

Table 6.2: Weights used in the proposed method for the tactical and operational schedule in all three case studies.

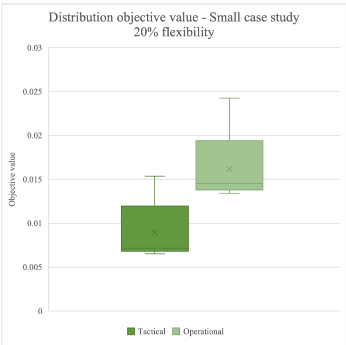
		Schedule	
Constraints	Weight	Tactical	Operational
TRC	w_{RT}	25	
	w_{FRO}	25	
	w_{CWD}	10	
	w_{CNS}	1	
	w_{RM}	10	
	w_{EOW}	20	
	w_{2W}	20	
	w_{SFS}	7	
	w_{RND}	1	
ORC	w_{week}	1	2
	$w_{weekend}$	2	2
	w_{night}	1	2
	w_{UQ}	5	1
	w_{QL3}	7	3

6.3 Experimental results current performance

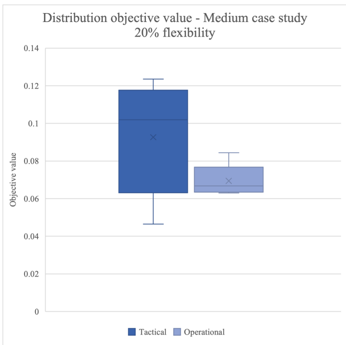
This section describes the outcomes for the three case studies using the parameter values found in Section 6.1 and Table 6.1. We compare the outcomes of the proposed method with the current practice to evaluate the performance of the method. A description of the case studies and the number of TRC and ORC violations in the historical schedule are provided in Section 5.2 and Table 5.2.

We present the number of TRC and ORC violations for both the tactical and operational schedules after running the experiments five times with a flexibility parameter of 0.2, indicating that 80% of the tactical schedule should remain. As each SA solution is random, we present the best and worst solutions for each case study and compare them with the current performance. The best and worst solutions are chosen based on the operational schedule's objective value and give insight into the variation of the results generated by the method. Figure 6.3 visualized the distribution of objective value for both the tactical and operational schedule for the three case studies. Overall, the tactical schedule's objective value variates more than the operational schedule's values, except for the small case study as seen in Figure 6.3a, which is due to the high variation in the number of TRC and ORC violations between the best and worst results.

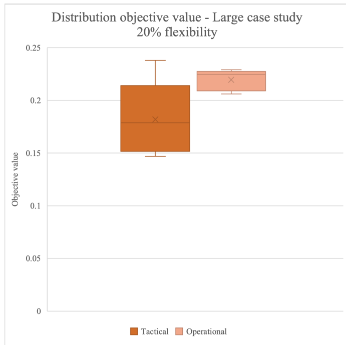
The results of the best and worst solutions are presented in Table 6.3. It includes the objective value, the number of violations, and the percentual change compared to the current performance. Additionally, Table 6.4 shows the difference in the number of violations between the operational and tactical schedules expressed by a ratio, which indicates if there has been an increase or decrease in the number of violations when constructing the operational schedule for the current performance and the results of the method. The individual experimental results are in Appendix D. To examine the KPIs for the nurses, we determine the ratio of remaining hours of the nurses, the hours assigned to intra-organizational flex nurses, and additionally for the medium case study distribution of the night shifts, which are visualized using boxplots. The KPI for the requested additional free day is left out of scope as this is not included in the data.



(a) Distribution objective value small case study



(b) Distribution objective value medium case study



(c) Distribution objective value large case study

Figure 6.3: Distribution of the objective value for the tactical and operational schedule of five experiments across three case studies.

Table 6.3: Experimental results of the number of TRC and ORC violations in both tactical and operational schedules for the current performance and the best and worst solutions from the proposed method using a flexibility parameter 0.8. The number of violations is represented along with the difference in percentage between the solution found by the method and the current performance. The TRCs are expressed in the total number of violations. The ORC for the week, weekend, night, and UQ are in the number of shifts, and the QL3 is the number of missing hours during the planning period.

					TRC																ORC												
			Objective		RT		FRO		CWD		CNS		EOW		2W		SFS		Total TRC		Week		Weekend		Night		UQ		QL3 hours		Total ORC	
Case study	Schedule	Solution	#	% (x100)	#	% (x100)	#	% (x100)	#	% (x100)	#	% (x100)	#	% (x100)	#	% (x100)	#	% (x100)			#	% (x100)	#	% (x100)	#	% (x100)	#	% (x100)	#	% (x100)		
Small	Tactical	Current	0.6125		13.00		9.00		0.00		-		0.00		0.00		16.00		38.00		11.00		3.00		-		0.00		20.00		34.00	
		Best	0.0071	-0.99	0.00	-1.00	0.00	-1.00	0.00	0.00	-	-	0.00	0.00	0.00	0.00	7.00	-0.56	7.00	-0.82	2.00	-0.82	0.00	-1.00	-	-	6.00	6.00	0.00	-1.00	8.00	-0.76
		Worst	0.0086	-0.99	0.00	-1.00	1.00	-0.89	0.00	0.00	-	-	0.00	0.00	0.00	0.00	7.00	-0.56	8.00	-0.79	8.00	-0.27	3.00	0.00	-	-	6.00	6.00	20.00	0.00	37.00	0.09
	Operational	Current	0.5986		15.00		11.00		0.00		-		0.00		0.00		16.00		42.00		0.00		0.00		-		0.00		0.00		0.00	
		Best	0.0134	-0.98	0.00	-1.00	0.00	-1.00	0.00	0.00	-	-	0.00	0.00	0.00	0.00	10.00	-0.38	10.00	-0.76	0.00	0.00	0.00	0.00	-	-	6.00	6.00	0.00	0.00	6.00	0.00
		Worst	0.0243	-0.96	0.00	-1.00	1.00	-0.91	0.00	0.00	-	-	0.00	0.00	0.00	0.00	34.00	1.13	35.00	-0.17	3.00	3.00	3.00	3.00	-	-	6.00	6.00	11.00	11.00	23.00	23.00
Medium	Tactical	Current	0.0905		1.00		3.00		2.00		5.00		5.00		7.00		19.00		37.00		68.00		28.00		5.00		1.00		100.00		202.00	
		Best	0.1112	0.23	13.00	12.00	12.00	3.00	1.00	-0.50	0.00	-1.00	16.00	2.20	5.00	-0.29	29.00	0.53	76.00	1.05	30.00	-0.56	16.00	-0.43	1.00	-0.80	7.00	6.00	17.00	-0.83	71.00	-0.65
		Worst	0.1019	0.13	6.00	5.00	6.00	1.00	0.00	-1.00	0.00	-1.00	9.00	0.80	4.00	-0.43	26.00	0.37	51.00	0.38	60.00	-0.12	28.00	0.00	8.00	0.60	8.00	7.00	119.00	0.19	223.00	0.10
	Operational	Current	0.1791		7.00		10.00		8.00		5.00		11.00		11.00		29.00		76.00		0.00		0.00		0.00		15.00		0.00		15.00	
		Best	0.0630	-0.65	7.00	0.00	7.00	-0.30	1.00	-0.88	0.00	-1.00	17.00	0.55	8.00	-0.27	35.00	0.21	75.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	8.00	-0.47	0.00	0.00	8.00	-0.47
		Worst	0.0845	-0.53	13.00	0.86	13.00	0.30	0.00	-1.00	0.00	-1.00	17.00	0.55	10.00	-0.09	35.00	0.21	88.00	0.16	0.00	0.00	0.00	0.00	0.00	0.00	7.00	-0.53	0.00	0.00	7.00	-0.53
Large	Tactical	Current	0.0498		2.00		2.00		0.00		-		0.00		0.00		98.00		102.00		0.00		0.00		-		8.00		0.00		8.00	
		Best	0.1468	1.95	1.00	-0.50	5.00	1.50	0.00	0.00	-	-	19.00	19.00	2.00	2.00	53.00	-0.46	80.00	-0.22	97.00	97.00	28.00	28.00	-	-	64.00	7.00	16.00	16.00	205.00	24.63
		Worst	0.1787	2.59	13.00	5.50	17.00	7.50	0.00	0.00	-	-	18.00	18.00	3.00	3.00	64.00	-0.35	115.00	0.13	79.00	79.00	36.00	36.00	-	-	68.00	7.50	18.50	18.50	201.50	24.19
	Operational	Current	0.0487		2.00		2.00		0.00		-		0.00		0.00		98.00		102.00		0.00		0.00		-		8.00		0.00		8.00	
		Best	0.2060	3.23	1.00	-0.50	5.00	1.50	0.00	0.00	-	-	19.00	19.00	1.00	1.00	53.00	-0.46	79.00	-0.23	97.00	97.00	28.00	28.00	-	-	64.00	7.00	16.00	16.00	205.00	24.63
		Worst	0.2291	3.70	13.00	5.50	17.00	7.50	0.00	0.00	-	-	18.00	18.00	0.00	0.00	65.00	-0.34	113.00	0.11	78.00	78.00	36.00	36.00	-	-	68.00	7.50	19.00	19.00	201.00	24.13

Table 6.4: The table shows the ratio increase or decrease of TRC and ORC violations in the operational schedule compared to the tactical schedule for the current schedule and best and worst solutions. Green indicates a decrease, representing an improvement in the number of violations, while red indicates an increase.

			TRC							ORC				
Case study	Solution	Objective	RT	FRO	CWD	CNS	EOW	2W	SFS	Week	Weekend	Night	UQ	QL3 hours
Small	Current	-0.02	0.15	0.22	0.00	-	0.00	0.00	0.00	-1.00	-1.00	-	0.00	-1.00
	Best	0.89	0.00	0.00	0.00	-	0.00	0.00	0.43	-1.00	0.00	-	0.00	0.00
	worst	1.83	0.00	0.00	0.00	-	0.00	0.00	3.86	-0.63	0.00	-	0.00	-0.45
Medium	Current	0.98	6.00	2.33	3.00	0.00	1.20	0.57	0.53	-1.00	-1.00	-1.00	14.00	-1.00
	Best	-0.43	-0.46	-0.42	0.00	0.00	0.06	0.60	0.21	-1.00	-1.00	-1.00	0.14	-1.00
	worst	-0.17	1.17	1.17	0.00	0.00	0.89	1.50	0.35	-1.00	-1.00	-1.00	-0.13	-1.00
Large	Current	-0.02	0.00	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00	-	0.00	0.00
	Best	0.40	0.00	0.00	0.00	-	0.00	-0.50	0.00	0.00	0.00	-	0.00	0.00
	worst	0.28	0.00	0.00	0.00	-	0.00	-1.00	0.02	-0.01	0.00	-	0.00	0.03

Small case study

The result for the small case study can be found in Tables 6.4 and 6.3 and Figure 6.4. As there are no NS, we do not consider the CNS violations, undercoverage, and the distribution of the NS.

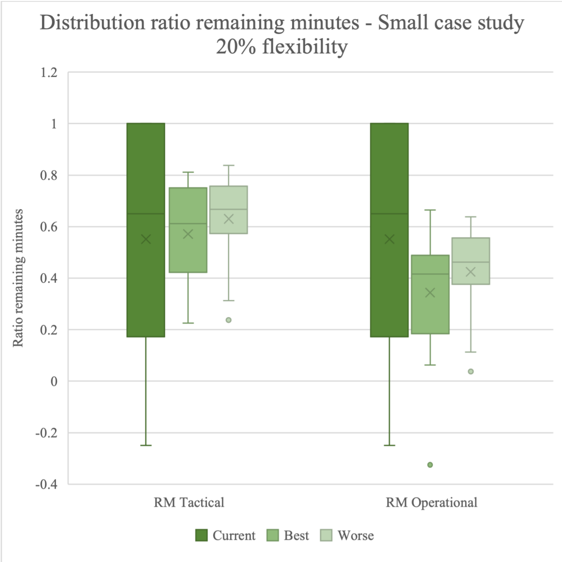
As seen in Table 6.4, the operational objective value of the current schedule decreases by 2% compared to the current tactical schedule. This results from a 100% decrease in ORC violations but comes at the cost of a 15% and 22% increase in RT and FRO violations. On the other hand, the operational objective value of the best and worst solution found by the method increases, resulting from an increase in SFS violations, not resolving the UQ shifts, and in the case of the worst solution, not meeting all coverage requirements. The best solution has resolved all undercoverage in the operational schedule but comes at the cost of a 43% increase in TRC violations.

However, the best and worst solution found by the method outperforms the current performance in terms of the objective value and the number of TRC violations. Both solutions have decreased the objective value by 99% regarding the tactical schedule, and the best solution has decreased the objective value by 98% in the operational schedule. The lower objective value results from the decrease in TRC violations, where the method has decreased the total number of TRC violations by 82% and 79% in the tactical schedule and by 76% and 17% in the operational schedule, regarding the best and worst solution, respectively. Only the worst performance results in an increase of SFS violations compared to the current performance. Looking at the ORC violations in Table 6.3, the method assigns six times more UQ shifts in both tactical and operational schedule compared to the current performance. This is possibly due to the weights assigned to the constraint, as it does result in an 82% decrease for the best solution and a 27% decrease for the worst solution of undercoverage during the week in the tactical schedule. In the operational schedule, the UQ shifts may not be rescheduled due to the allowed 0.2 flexibility and the goal to meet the demand despite the fact it can result in UQ shifts. The worst solution has improved the number of weekly shifts assigned by 27% in the tactical schedule. However, it does result in undercoverage and missing QL3 hours in the operational schedule.

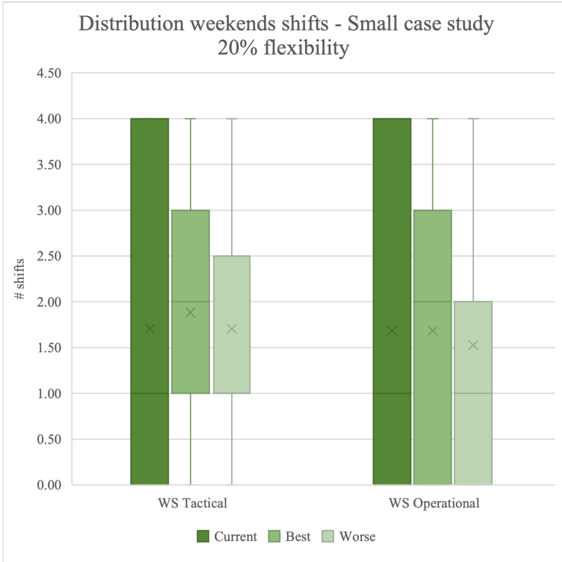
Comparing the results in Figure 6.4, there appears to be more variability in the distribution of RM in the current practice than in the proposed method. Indicating that the proposed method distributed the contract hours more fairly. The current performance distributed the RM similarly in the tactical and operational schedule. Additionally, there are two things to point out. First, in both schedules, the current practice assigns overtime to the nurses, indicated by the negative ratio RM on the y-axis. In contrast, the proposed method only assigns overtime in the operational schedule in the best solution to one nurse, as indicated by the outlier. The worst solution does not assign overtime, which explains the undercoverage. Second, the proposed method assigns shifts to all nurses in the operational schedule. In contrast, the current schedule also has nurses without shifts assigned, as seen in Figure 6.4a. All solutions have more variability towards the lower ratios of RM. Yet, the worst solution is centered around higher values for RM, which is another factor that contributes to the undercoverage.

Figures 6.4b, 6.4c, and 6.4d visualizes the distribution of the weekend shifts (WS). As can be seen, almost 50% of the nurses in practice get no weekend shifts assigned, indicated by the boxplot in Figure 6.4b and the lines in Figure 6.4c and 6.4d. When assigned, most nurses get four WS assigned in both tactical and operational current performance. The distribution of the best and worst solutions is comparable. However, the worst solution often assigns more single WS, explaining the undercoverage during the weekend. The best solution assigns two WS more often, also indicated by the cumulative percentiles on the second y-axis. The highest variation in weekend shifts assignment occurs in the current schedule, indicating a more fair WS distribution by the proposed method.

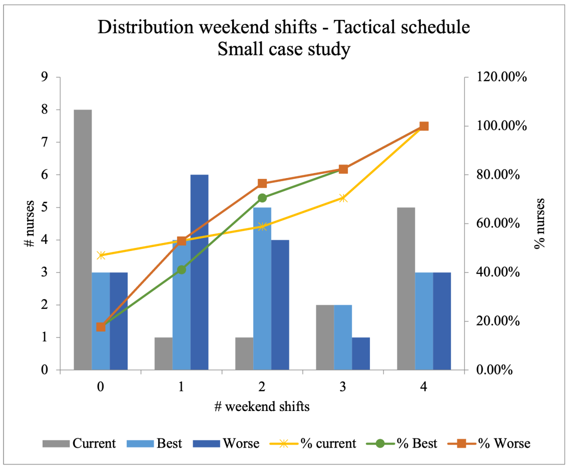
The small case study has only two intra-organizational flex nurses. The current practice has assigned 4,260 and 0 minutes to the two nurses, whereas the best and worst solutions have assigned 132 and 12, and 300 and 0 minutes, respectively. The method prioritizes assigning minutes to regular nurses compared to the current performance.



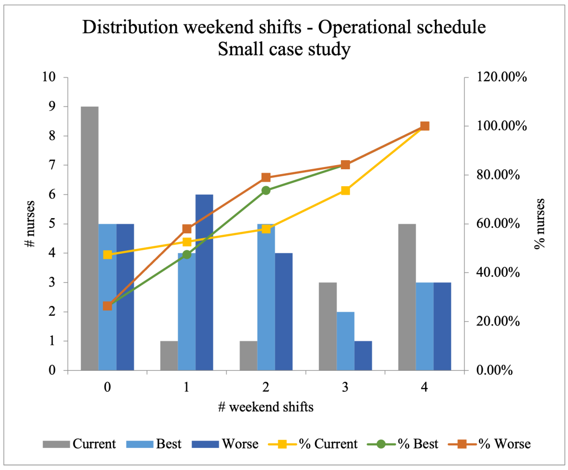
(a) Distribution of the remaining minutes of the regular nurses for the tactical and operational schedule of the small case study. The negative ratio on the y-axis indicate min-hours or overtime.



(b) Distribution of the number of weekend shifts assigned of the small case study.



(c) Distribution of weekend shifts in the tactical schedule.



(d) Distribution of weekend shifts in the operational schedule.

Figure 6.4: Small case study: Results of the manual schedule and proposed method for the KPIs of the nurses with the distribution of remaining minutes of the regualr nurses and the distribution of the weekend shifts.

Medium case study

First, it must be noted that all three operational schedules have a large decrease in ORC violations, being 93%, 89%, and 97% for the current, best, and worst schedules, respectively. Within the operational schedule, more nurses are available, which results in less or no undercoverage. Comparing the difference between the tactical and operational schedule in Table 6.4, only the best and worst solutions generated by the method have optimized the objective value, as indicated by the decrease of 43% and 17%. This is mainly due to the decrease in ORC violations in both solutions and the 46% in RT and 42% reduction in FRO violations in the best solution. As can be seen, the 93% reduction of ORC violations comes at the cost of a 105% increase in the total number of TRC violations and an additional 14 UQ shifts in the current performance. It is shown that the decrease in ORC violations comes at the cost of an increase in either or both TRC violations or UQ shifts. We will discuss the performance of the best and worst solutions compared to the current schedule.

As seen in Table 6.3, the current performance of the medium case study has a better objective value for the tactical schedule compared to the best and worst solution, as both increase the objective value with 23% and 13% respectively. This results from the increase in RT, FRO, EOW, and SFS violations which contribute to the total increase of 53% and 37% increase in total TRC violations. Nevertheless, the best solution has covered almost 50% more of the uncovered week and weekend shifts and 80% more of the missing QL3 hours of the current schedule. The worst solution only performs better in weekly shifts, covering 12% more compared to the current tactical schedule. Despite that, the solutions perform Worst regarding the number of UQ shifts, which is six and seven as high compared to the current performance. This can be due to the same reason stated for the small case study, as we have assigned certain weights to the UQ constraint that influences the optimization method.

Regarding the operational schedule, the best and worst solutions have decreased the objective value of the current performance by 65% and 63%. It must be pointed out that none of the three solutions has remaining shifts or uncovered QL3 in the operational schedule. The following reasons can explain the decrease in the objective value. First, there is a 47% and 53% decrease in the number of UQ shifts for the best and worst case, respectively. Additionally, the two solutions have 88% and 100% less CWD, 100% less CNS, and 27% and 9% less 2W violations. Moreover, the best solution has improved the number of FRO violations by 30%. The current practice outperforms the proposed method regarding the number of EOW and SFS violations, as the method has violated the EOW by 55% and the SFS by 21% compared to the current performance.

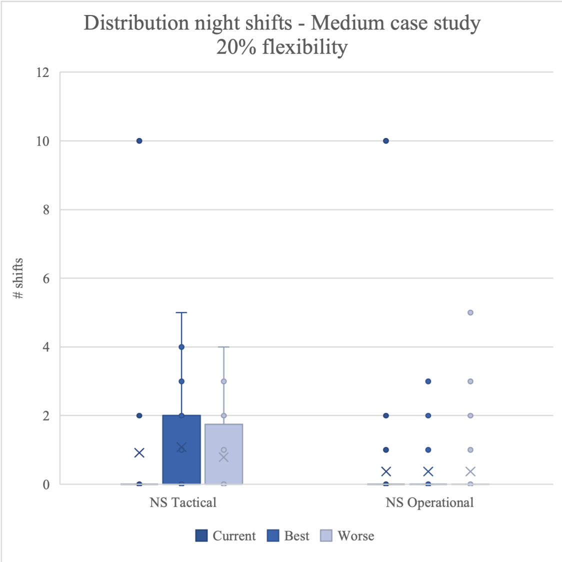
Figures 6.5 and 6.6 give an overview of the KPIs for the nurses in the medium case study. The current performance has assigned almost 90% of the nurses no NS in both the tactical and operational schedule, as it assigns two shifts to two nurses and ten shifts to another nurse, as indicated by the outliers in Figure 6.5a and can be seen in Figures 6.6a and 6.6b. The two solutions found by the method distribute the NS more similarly, as seen in Figure 6.6a and 6.6b, where almost 60% get no night shift assigned in the tactical schedule and 80% in the operational schedule. This increase in the operational schedule can be explained as there are more nurses available than NS that must be assigned. Figure 6.6a shows that the worst solution does not assign more than four night shifts, explaining the undercoverage in the tactical schedule. In the operational schedule, the best solution assigns not more than three NS, whereas the worst solution assigns five night shifts to two nurses.

This indicates that the best solution distributes the night shifts more evenly among the nurses, compared to the current performance and worst solution.

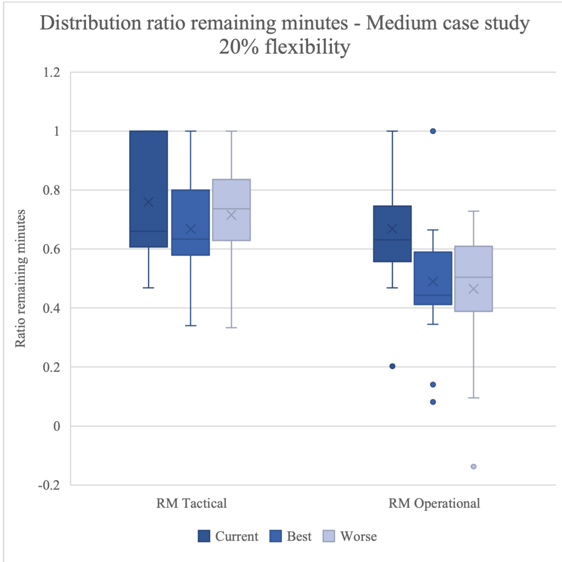
Figure 6.5b displays the distribution of the RM of the nurses. In the tactical schedule, the current performance has more nurses not assigned to any shifts, whereas the method did not assign shifts to two and one nurse in the best and worst solution. Indicating that the proposed method distributes the RM in the tactical schedule more fairly, even if there is a higher variation of RM. This can occur as there are nurses with only a few contract hours, letting the ratio decrease quickly when one shift is assigned. The distribution of RM in the operational schedule is more comparable for the three solutions. Despite the fact that there are two nurses not assigned in the best solution, it obtains the smallest variation and thus a more fair distribution of RM. Likewise, the proposed method has assigned more shifts to regular nurses compared to the current performance. It must be pointed out that only the worst solution has been assigned overtime, indicated by the negative outlier. To fulfill the coverage requirement of the operational schedule, the current practice has assigned more hours to flex nurses, as shown in Figure 6.5c. However, the small boxplot indicates that a majority is assigned equal hours. The large number of shifts assigned to flex nurses can explain the increase of 14 UQ shifts in the operational schedule. The assigned flex hours in the proposed method are similar for the best and worst solutions. Combining the outcomes of the ratio for RM in Figure 6.5b, it appears that the method assigns more hours to regular nurses than flex nurses.

Lastly, Figures 6.5d, 6.6c, and 6.6d visualize the distribution of the WS. Almost 40% of the nurses in the current tactical schedule are assigned no WS, whereas the best and worst solutions assign two and one WS the most, respectively. Both have an outlier in the tactical schedule by assigning six and five shifts to one nurse, as seen in Figures 6.5d and 6.6c. Additionally, only a few nurses are assigned no WS in the best solution, which explains the 43% decrease in undercoverage of weekend shifts in the tactical schedule, but also explains the 220% increase in EOW violations.

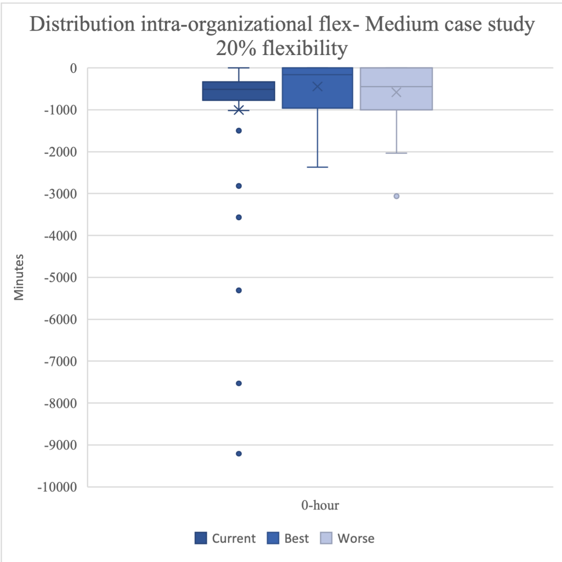
The distribution of WS is more comparable in the operational schedule for the three solutions, where more than 50% of nurses are assigned no WS, which results from the fact that there are more nurses than weekend shifts.



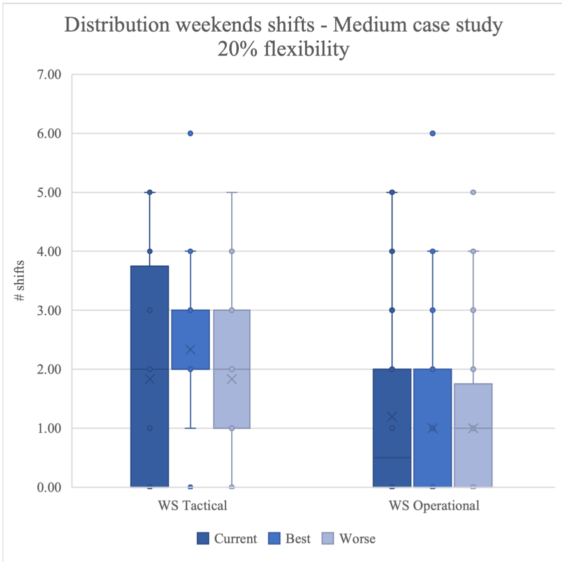
(a) Distribution of the night shifts within the tactical and operational schedule.



(b) Distribution of the remaining minutes of the full-time nurses for the tactical and operational schedule.

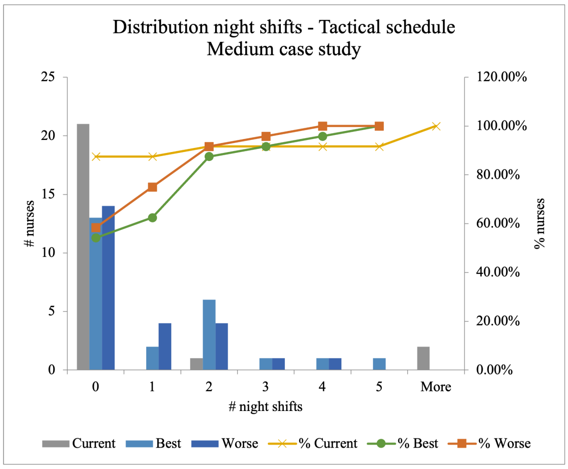


(c) Distribution of the hours assigned to the nurses with 0-hour contracts.

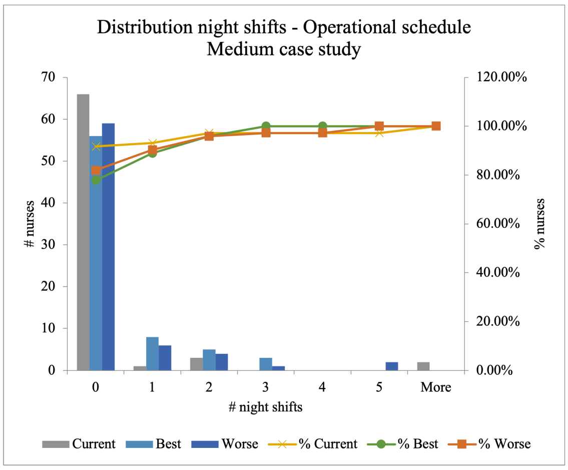


(d) Distribution of the number of weekend shifts assigned.

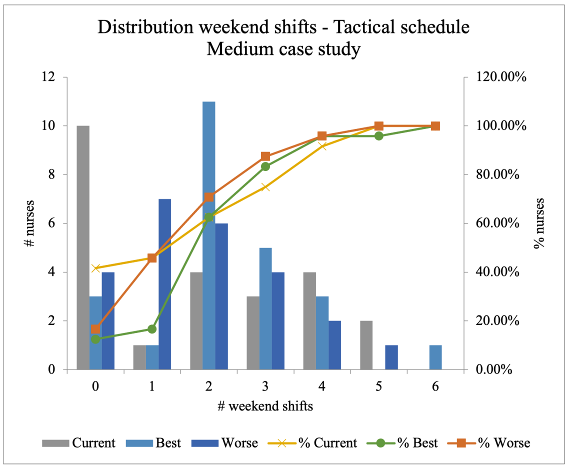
Figure 6.5: Medium case study: Result of the current performance and proposed method for different flexibility parameters. Including the KPIs for the nurses with the distribution of night shifts, remaining minutes of the full-time nurses, the minutes assigned to employees with 0-hour contracts, and the distribution of the weekend shifts.



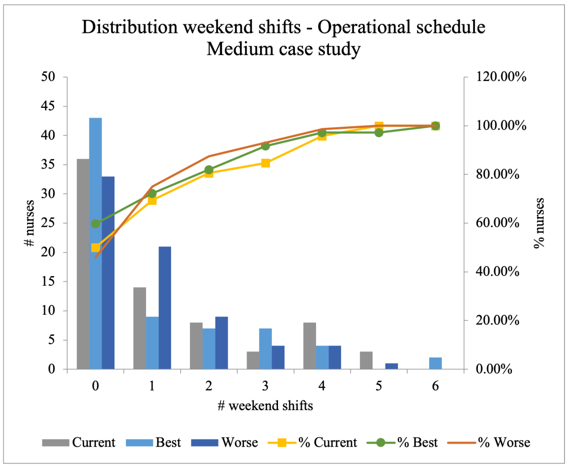
(a) Distribution night shifts in the tactical schedule.



(b) Distribution night shifts in the operational schedule.



(c) Distribution of the weekend shifts in the tactical schedule.



(d) Distribution of the weekend shifts in the operational schedule.

Figure 6.6: Medium case study: Distribution of the night and weekend shifts of the tactical and operational schedule for the current, best, and worst solution. It displays both the count and percentage of nurses, as well as the number of shifts they have been assigned.

Large case study

The results for the TRC and ORC for the large case study are shown in Tables 6.3 and Table 6.4 and for the nurses' KPI in Figure 6.7. As there are no NS, again, we do not consider the CNS violations, undercoverage, and the distribution of the NS.

As seen in Table 6.3, the objective value of the current performance is improved by 2%. Since there is no change in the number of TRC and ORC violations, this results from changing the weights assigned to w_{nurses}. Both best and worst solutions resulted in a 40% and 28% increased objective value in the operational schedule. As seen in Tables 6.4 and 6.3, only the number of ORC violations has changed in the worst solution, as there is one additional shift assigned but also results in a 3%

increase of missing rest hours. However, as seen in Table 6.3, the best solution outperforms the worst solution in the number of undercoverage during the weekend and missing QL3 hours.

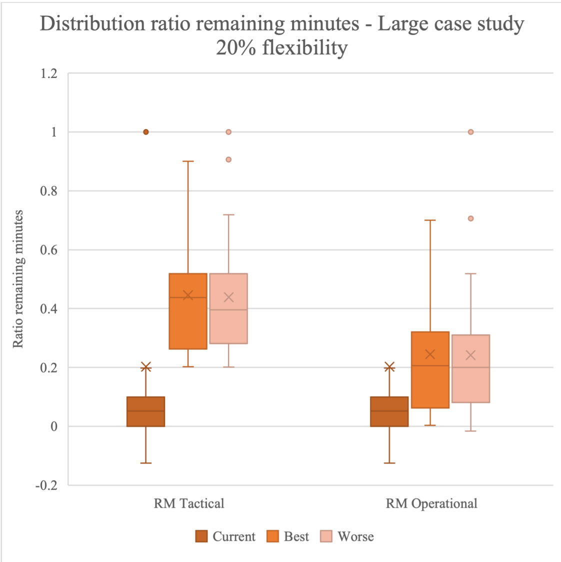
Looking at Table 6.3, the current schedule outperforms the method in terms of different TRCs and ORCs violations, which result in a large difference in objective value. Nevertheless, the best solution has decreased the total amount of TRC by 22% in the tactical schedule and by 23% in the operational schedules. The method performs better in terms of SFS violations, where it decreases the number of violations by 46% in the best solution and 35% in the worst solution. The variation in the number of RT violations is large for the best and worst solution, as the difference between the current and the worst solution is almost ten times as large as the current and the best solution. It is noteworthy that none of the solutions result in CWD violations in both schedules. However, the proposed method does not result in a schedule that meets all coverage requirements, as there is still a large proportion of undercoverage and UQ shifts. We will explain these observations by further evaluating the results regarding the nurses' KPIs.

Figure 6.7 shows the distribution of RM and WS. As there are only four intra-organizational flex nurses, we did not report on this subgroup in a figure. Yet, none of the solutions has assigned hours to these flex nurses.

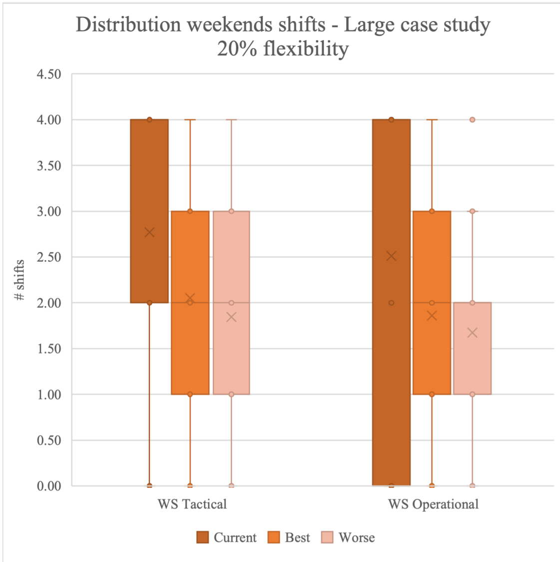
Figure 6.7a shows the distribution of RM. The first thing to point out is that the current schedule has assigned overtime both in the tactical and operational schedules. Contrarily, the method prohibits assigning overtime in the tactical schedule, and only one nurse is assigned overtime in the worst operational schedule. This results in the large difference between the current and method performance ORC violations. Also, as seen in 6.7a, there is more variation in the distribution of RM in the proposed method, indicating a more unfair distribution. However, the best tactical solution has assigned shifts to all nurses, whereas the current and worst solution does not assign shifts to all nurses. Moreover, as mentioned in Section 5.2, the ratio of available minutes and demand in minutes is almost equal to 1 as we attempt to assign at most 80% of the contract hours. This makes it difficult or impossible to assign all shifts in the tactical schedule while meeting the other requirements and minimizing the penalties for the TRC. This is also shown in Figure 6.2a, that visualizes the trade-off, as none of the experiments resulted in a solution without ORC violations. Also, as mentioned in Section 6.1, the parameter settings of the SA affect the results. As we only ran the method during the parameter tuning phase for five minutes to find the best parameter settings, it is possible that other parameter settings would result in better solutions as the algorithm could escape the local optima.

At last, we examine the distribution of WS, which is visualized in Figures 6.7b, 6.7c, and 6.7d. As shown in Figure 6.7c and 6.7c, the current practice only assigns none or even numbers of WS. Most nurses are assigned four WS in the current tactical and operational schedule, resulting in fully covered weekends. The distribution of the best and worst tactical solutions is almost equal.

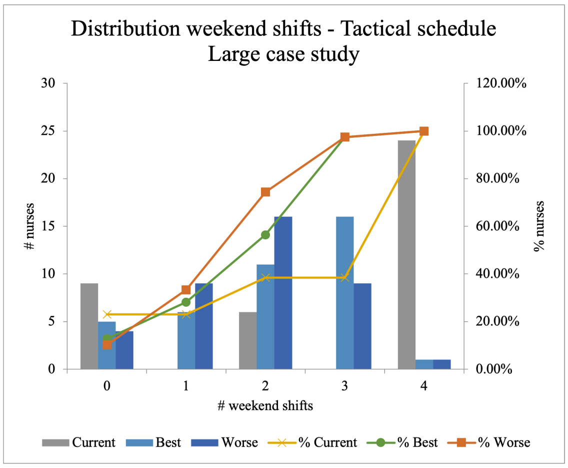
However, the best solution assigns three WS more often, resulting in less undercoverage of WS than the worst solution.



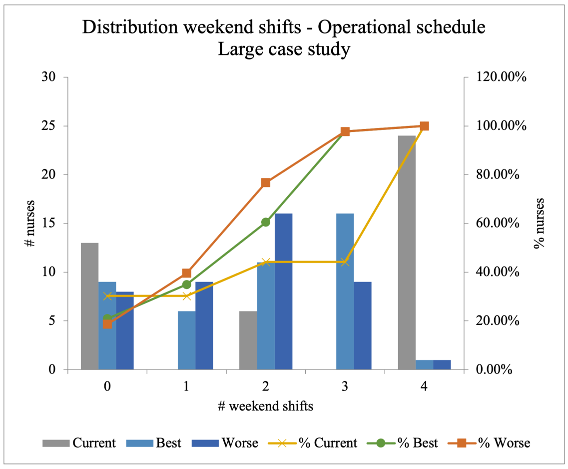
(a) Distribution of the remaining minutes of the full-time nurses for the tactical and operational schedule of case study 2.



(b) Distribution of the number of weekend shifts assigned.



(c) Distribution of weekend shift in the tactical schedule.



(d) Distribution of weekend shift in the operational schedule.

Figure 6.7: Case study 2: Results of the manual schedule and proposed method for the KPIs of the nurses with the distribution of remaining minutes of the full-time nurses and the distribution of the weekend shifts.

Overall outcome

Based on the results of the three case studies, the proposed method resulted in better objective values in the operational schedule than the current schedule for the small and medium case study, indicated by the 53%-98% decrease in Table 6.4. In both cases, the best operational schedule did not have undercoverage or missing QL3 hours. Additionally, based on the best solutions, the proposed method can decrease the number of RT, FRO, CWD, CNS, 2W, and SFS violations and decrease the total number of operational TRC violations by 76%, 1%, and 22% in the best solutions for the three case studies respectively. Additionally, the proposed method also decreases the total number of TRC in the tactical schedule for the best solutions of the small and medium case study by 82% and 22%, respectively. Regarding the latter, this comes at the cost of not meeting all the coverage requirements. It is case-dependent in which TRCs are decreased. Noteworthy is that the proposed method did not violate the CWD constraint in each case study, except for the best solution of the medium case study, which is improved by 50% compared to the current solution. The proposed method must improve in optimizing the number of EOW and UQ violations, as the current performance results in fewer EOW violations for the medium and large case studies and fewer UQ shifts for all three case studies, except for the operational schedule of the medium case study.

Based on the results from the KPIs, the proposed method can distribute the RM more fairly in the small and medium case study, where it assigns shifts to almost all nurses. Also, it does not assign overtime and more than 80% of the contract hours in the tactical schedule of all three case studies. Furthermore, the proposed method prioritizes assigning shifts to regular nurses rather than intra-organizational flex nurses, despite the fact that it can result in undercoverage. Depending on the case study, the WS in the proposed method and current practice are distributed more similarly. Most nurses are assigned no WS in the operational schedule of the medium case studies in the proposed method and current solution, which is due to the fact that there are more nurses than weekend shifts. The distribution of WS differs the most in the large case study, as in practice, only none or even weekend shifts are assigned, resulting in no EOW or 2W violations. Nevertheless, the distribution of weekend shifts would have been more fairly if all nurses had gotten the same amount of weekend shifts assigned rather than four or more and others none.

When evaluating the best solutions, the proposed method can improve both the tactical and operational schedules for the small case study regarding both TRC and ORC violations and the operational schedule for the medium case study regarding CWD, CNS, 2W, and total ORC violations. However, as the worst solution did not always result in better or equal solutions and the variation between the number of TRC and ORC violations differs, it must be kept in mind that the SA generated random solutions. This was also shown in Figure 6.3, where the distribution of the objective value is visualized, resulting in a larger variation in the tactical schedule compared to the operational schedule for the medium and large case studies.

6.4 Flexibility analysis

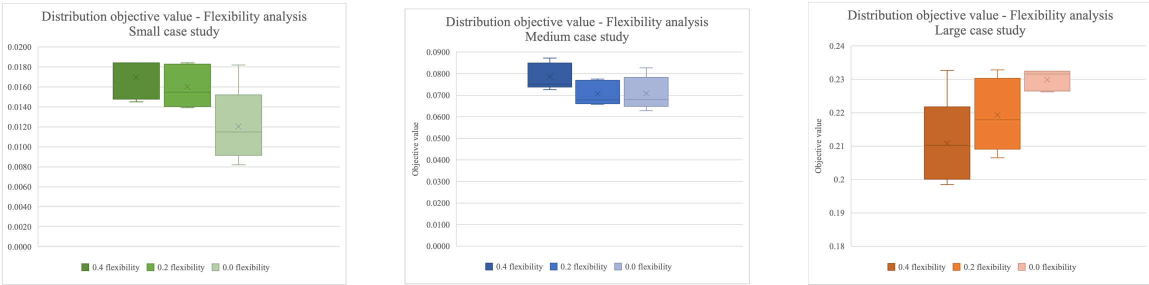
We present an experimental evaluation of the algorithm by changing the flexibility parameter. The goal is to analyze how the parameter influences the outcome and what percentage of the tactical schedule should remain unchanged to get a high-quality fair operational schedule, using the KPIs from Section 4.3. As mentioned in Section 4.2.3, to construct the operational schedule, we use the tactical schedule as input, which is optimized for the allowed flexibility.

Experiments are conducted using flexibility parameters of 0.4, 0.2, and 0.0, representing the percentage of tactical shift assignments that are allowed to be rescheduled in the operational schedule. The other parameter values remain the same as discussed in Section 6.1 and 6.3, which include the parameters for the SA and the weights for the objective. Table 6.5 shows the best objective values from five experimental runs, the fairness measure, the number of iterations, and the running time for each case study. The distribution of the operational objective value can be found in Figure 6.8. We evaluate fairness by the distribution of RM, WS, and the fairness measure.

We examine the effect of the flexibility parameter by determining the gap between the number of TRC and ORC in the tactical and operational schedules, denoted by $\Delta O - T$, and the percentual change. Tables 6.6, 6.7, and 6.8 show the best and worst results for TRC and ORC violations after running the method five times for each case study. In Figures 6.10, 6.11, and 6.9, the experimental results for the nurses' KPIs are shown. The same tactical schedule is used as input for the experiments to make an equal comparison of the influence of the flexibility parameter, which is generated by running the SA algorithm once.

Table 6.5: The results of the flexibility analysis, which include the objective value, the fairness objective, the number of iterations, and the running time for the best solution for the three case studies and flexibility parameters, bold values indicate the optimum performance for each case study.

Case study	Schedule	Flexibility	Objective	Fairness	Iterations	Run Time(s)
Small	Tactical		0.0074	0.0048	3002	270.69
hline	Operational	0.4	0.0145	0.0121	2501	282.98
		0.2	0.0139	0.0116	2501	263.33
		0	**0.0082**	0.0050	2501	266.88
Medium	Tactical		0.0921	0.1053	847	137.27
hline	Operational	0.4	0.0726	0.0812	690	172.78
		0.2	0.0658	0.0667	690	196.77
		0	**0.0628**	0.0694	690	189.50
Large	Tactical		0.2102	0.0460	787	227.68
hline	Operational	0.4	**0.1985**	0.1071	787	160.82
		0.2	0.2065	0.1550	787	190.80
		0	0.2262	0.0732	787	166.53



(a) Objective value small case study

(b) Objective value medium case study

(c) Objective value large case study

Figure 6.8: Distribution of the operational objective value of the five experiments for different flexibility parameters for the three case studies.

Small case study

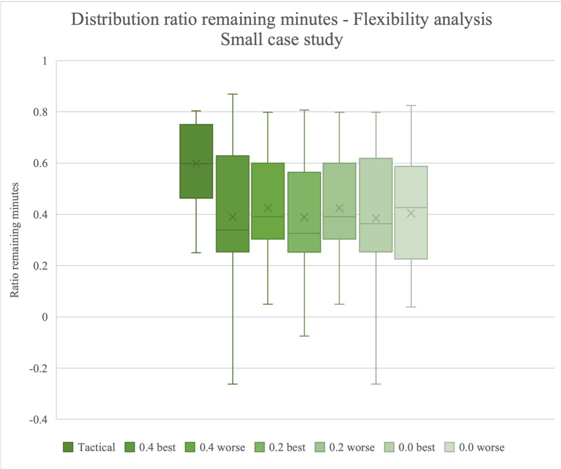
Table 6.6 and Figure 6.9 show the results of the experiments for the small case study. As shown in Table 6.6, all best solutions result in an operational schedule without undercoverage during the week, weekend, and has no missing QL3 hours. As can be seen, all best solutions and both solutions using flexibility of 0.0, result in a total ORC decrease of 85% or 90%. Additionally, all best solutions result in a decrease of UQ shifts, where the 0.4 flexibility parameter has the largest decrease of 50% of UQ shifts. However, the decrease in ORC violations and meeting the coverage requirements comes at the cost of a large increase in the number of TRC violations, where the smallest increase is 0.75 using a flexibility of 0.0 and the largest using a flexibility of 0.2, which increases 7.75 times compared with the tactical violations. This increase in TRC violations is mainly due to the large increase of SFS violations, with a maximum increase of 7.75 times the number of tactical violations or a small increase of violations of other TRCs. It must be noted that none of the parameters results in an increase of RT violations and only a single violation of the FRO and CWD constraints.

The combination of the smallest increase of TRC violations and meeting all the demands resulting in a decrease in ORC violations results in the best objective value for flexibility using 0.0. It also results in the smallest fairness outcome, which indicates that there is a fair distribution of TRC penalties among the nurses. However, as indicated by Figure 6.8a, using a flexibility of 0.0 results in a higher variation in the objective value. This is due to the variation of SFS violations in the best and worst solutions, as the decrease in ORC violations is equal for both solutions. The worse solutions of the 0.2 and 0.4 parameters are identical and do not resolve any undercoverage of ORC violations and, thus, not in a change of TRC violations. This can be the case as, due to the flexibility parameter, the search space of the SA algorithm is larger compared to a 0.0 flexibility. It, therefore, does not result in another best solution as it can reassign more shifts which results in a worse solution.

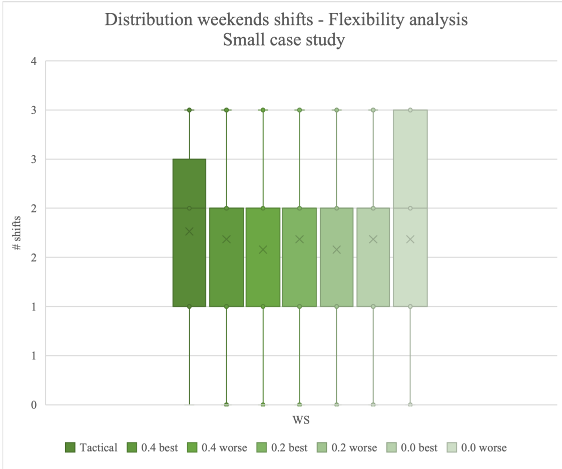
For the RM we compare the results of the solutions that covered all shifts. Comparing the results in Figure 6.9a, it results that none of the nurses is assigned no shifts. However, there is more variation in RM for the three best solutions, with the most comparable and highest variation with 0.4 and 0.0 flexibility and the least in 0.2 or worse 0.0 flexibility in the worst solution, indicating a more fair distribution of RM using the latter two. As can be seen, using a 0.4 and 0.0 flexibility are the only two solutions that have assigned overtime to meet the coverage requirements resulting in a higher variation of RM. All operational solutions result in a higher variation in RM compared to the tactical schedule, likely due to the processed absenteeism and other nurses that have to fill these shifts. Furthermore, only the worst solution using a 0.0 flexibility parameter has assigned hours to flex nurses, resulting in a smaller variation of RM in the worst solution. Looking at Figures 6.9b and 6.9c, the WS assignments are similar for the three flexibility parameters. When allowing 0.0 flexibility, more nurses get a higher amount of WS assigned than 0.4 flexibility, but the difference is small.

Table 6.6: Small case study: Best and worst results for the flexibility analysis after five runs for flexibility parameters of 0.0, 0.2, and 0.4.

			TRC violations							ORC violations				
Schedule	Flexibility	Solution	RT	FRO	CWD	EOW	2W	SFS	Total TRC	Week	Weekend	UQ	QL3 hours	Total ORC
Tactical			0	0	0	0	0	4	4	7	2	4	7	20
Operational	0.4	Best	0	0	1	2	0	13	16	0	0	2	0	2
		Worst	0	0	0	0	0	4	4	7	2	4	7	20
	0.2	Best	0	1	0	0	0	34	35	0	0	3	0	3
		Worst	0	0	0	0	0	4	4	7	2	4	7	20
	0	Best	0	0	0	0	0	7	7	0	0	3	0	3
		Worst	0	0	0	2	0	11	13	0	0	3	0	3
$-T$	0.4	Best	0	0	1	2	0	9	12	-7	-2	-2	-7	-18
		Worst	0	0	0	0	0	0	0	0	0	0	0	0
	0.2	Best	0	1	0	0	0	30	31	-7	-2	-1	-7	-17
		Worst	0	0	0	0	0	0	0	0	0	0	0	0
	0	Best	0	0	0	0	0	3	3	-7	-2	-1	-7	-17
		Worst	0	0	0	2	0	7	9	-7	-2	-1	-7	-17
% (x100)	0.4	Best	0	0	1	2	0	2.25	3	-1	-1	-0.5	-1	-0.9
		Worst	0	0	0	0	0	0	0	0	0	0	0	0
	0.2	Best	0	1	0	0	0	7.5	7.75	-1	-1	-0.25	-1	-0.85
		Worst	0	0	0	0	0	0	0	0	0	0	0	0
	0	Best	0	0	0	0	0	0.75	0.75	-1	-1	-0.25	-1	-0.85
		Worst	0	0	0	2	0	1.75	2.25	-1	-1	-0.25	-1	-0.85



(a) Distribution of the remaining minutes of the full-time nurses for the tactical and operational schedule.



(b) Distribution of the number of weekend shifts assigned in a boxplot.

(c) Distribution of the number of weekend shifts assigned in a histogram.

Figure 6.9: Small case study: Experimental result of the proposed method for different flexibility parameters. Including the KPIs for the nurses with the distribution of remaining minutes of the full-time nurse and the distribution of the weekend shifts.

Medium case study

The best objective value is obtained using a 0.0 flexibility, which results in the highest decrease in total ORC violations of 91% and a small 20% increase in total TRC violations. As shown in Table 6.7, the best solutions of 0.4 and 0.0 flexibility result in a total decrease of 90% and 91% of ORC violations and thereby resulting in no ORC violations, except for an 11% increase of UQ shifts with a 0.4 flexibility. The best solution of 0.2 flexibility results in undercoverage during the week, weekend, and night but a smaller increase of 17% in the total number of TRC violations. Compared to the worst 0.2 solution, the best solution performs better regarding TRC violations. At the same time, it only decreases the uncovered week and weekend shifts by 70% and 84%, respectively, and increases the UQ shifts by 11%. The fairness measure is comparable for the three best solutions. However, when meeting all coverage requirements, the 0.0 flexibility results in the smallest fairness measure.

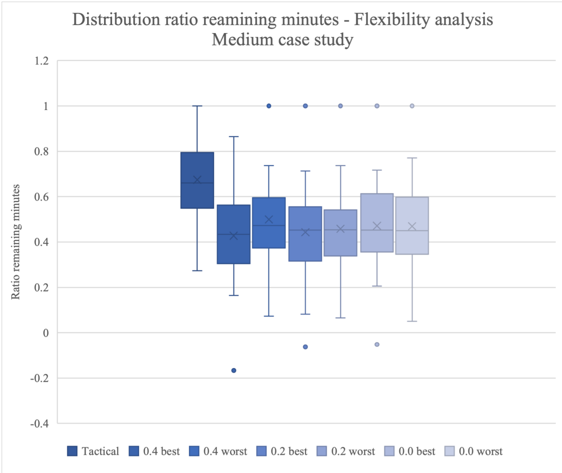
We compare the results of the best solution of 0.4 and 0.0 and the worst of 0.2 regarding the TRC violations, as these solutions result in a fully covered schedule. Allowing no flexibility results in the smallest increase of 40% RT and 30% EOW violations. When allowing 0.4 flexibility, there is an 8% decrease in SFS violations, the largest increase in 2W of 200%, and a 50% increase in FRO violations. The worst solution using a 0.2 flexibility results in a 22% decrease in UQ shifts, while meeting all coverage requirements. Allowing some flexibility provides opportunities to reschedule some UQ shifts. However, comes at the cost of an increase of RT, FRO, EOW, and 2W violations. Comparing the other two worst solutions, the 0.0 flexibility does result in a full schedule. Still, this results in the highest increase of total TRC violations, equal to 44%, and the highest increase of UQ shifts, equal to 33%. Therefore, this flexibility parameter results in the highest objective value variation, as seen in Figure 6.8b. As the goal is to fulfill most demand, this comes at the cost of more UQ shifts and an increase in TRC violations.

Figure 6.10 shows the results of the distribution of the nurses' KPIs. As shown in Figure 6.10a, the distribution of RM is comparable for the three flexibility parameters. In contrast with the other solutions, the best solution using the flexibility of 0.4 has assigned shifts to all nurses. The smallest variation is obtained by 0.2 worst and best 0.0 solutions, indicated by the smaller boxplot and short whiskers, contributing to a more fair distribution of RM. Also, it must be pointed out that all three best solutions have assigned overtime, indicated by the negative outliers. Additionally, we evaluate the number of minutes assigned to the flex nurses in Figure 6.10b. The best 0.4, 0.0 and worst 0.2 flexibility are comparable in the number and distribution of hours assigned to flex nurses, resulting in no undercoverage as shown in Table 6.7. On the other hand, the best 0.2 solution has assigned fewer shifts to flex nurses, and the 0.4 worst solution has assigned no hours to flex nurses, explaining the small decrease and the increase in undercoverage, respectively.

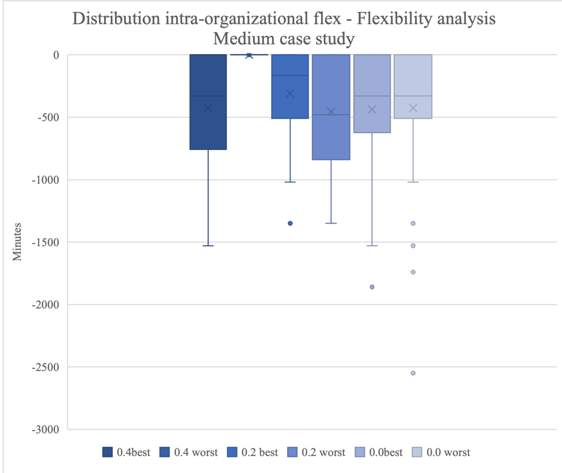
Next, we evaluate the distribution of NS and WS with Figures 6.10c, 6.10d, 6.10e, and 6.10f. As can be seen, almost 80% of the nurses get no NS and WS assigned, which can be explained by the increase in available nurses in the operational schedule as mentioned in Section 6.3. Only the worst 0.2 flexibility has assigned five-night shifts to one nurse. At the same time, the three solutions that did not result in undercoverage are the only ones that have assigned up to four night shifts. None of the operational schedules has resolved the EOW and 2W violations, as nurses are still assigned with more than four WS in the operational schedule, which directly implies that there has been a violation of the two constraints, as one can work at most two weekend shifts in a weekend. The best solution of 0.2 has assigned most nurses to a single weekend shift, which explains the undercoverage. Allowing flexibility of 0.0 results in the least amount of outliers, indicating a more fair WS distribution.

Table 6.7: Medium case study: Results for the flexibility analysis after five runs for flexibility parameters of 0.0, 0.2, and 0.4.

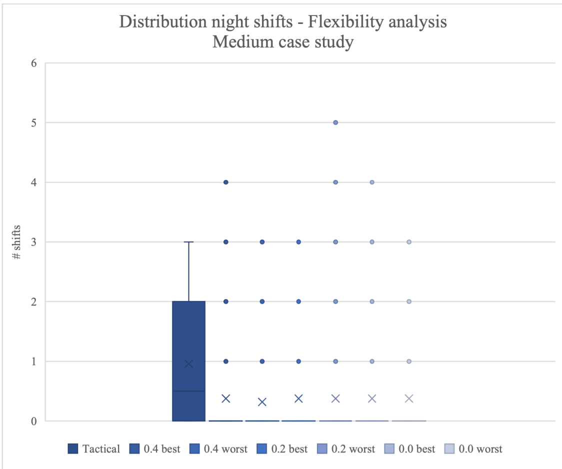
			TRC violations								ORC violations					
Schedule	Flexibility	Solution	RT	FRO	CWD	CNS	EOW	2W	SFS	Total TRC	Week	Weekend	Night	UQ	QL3 hours	Total ORC
Tactical			5	6	0	0	10	4	39	64	33	19	4	9	32	97
Operational	0.4	Best	8	9	0	0	16	12	36	81	0	0	0	10	0	10
		Worst	5	5	0	0	10	4	41	65	37	20	4	9	32	102
	0.2	Best	8	8	0	0	12	5	42	75	10	3	0	10	0	23
		Worst	10	11	0	0	15	9	39	84	0	0	0	7	0	7
	0	Best	7	10	0	0	13	4	43	77	0	0	0	9	0	9
		Worst	12	14	1	0	14	6	45	92	0	0	0	12	0	12
O-T	0.4	Best	3	3	0	0	6	8	-3	17	-33	-19	-4	1	-32	-87
		Worst	0	-1	0	0	0	0	2	1	4	1	0	0	0	5
	0.2	Best	3	2	0	0	2	1	3	11	-23	-16	-4	1	-32	-74
		Worst	5	5	0	0	5	5	0	20	-33	-19	-4	-2	-32	-90
	0	Best	2	4	0	0	3	0	4	13	-33	-19	-4	0	-32	-88
		Worst	7	8	1	0	4	2	6	28	-33	-19	-4	3	-32	-85
% (x100)	0.4	Best	0.60	0.50	0.00	0.00	0.60	2.00	-0.08	0.27	-1.00	-1.00	-1.00	0.11	-1.00	-0.90
		Worst	0.00	-0.17	0.00	0.00	0.00	0.00	0.05	0.02	0.12	0.05	0.00	0.00	0.00	0.05
	0.2	Best	0.60	0.33	0.00	0.00	0.20	0.25	0.08	0.17	-0.70	-0.84	-1.00	0.11	-1.00	-0.76
		Worst	1.00	0.83	0.00	0.00	0.50	1.25	0.00	0.31	-1.00	-1.00	-1.00	-0.22	-1.00	-0.93
	0	Best	0.40	0.67	0.00	0.00	0.30	0.00	0.10	0.20	-1.00	-1.00	-1.00	0.00	-1.00	-0.91
		Worst	1.40	1.33	1.00	0.00	0.40	0.50	0.15	0.44	-1.00	-1.00	-1.00	0.33	-1.00	-0.88



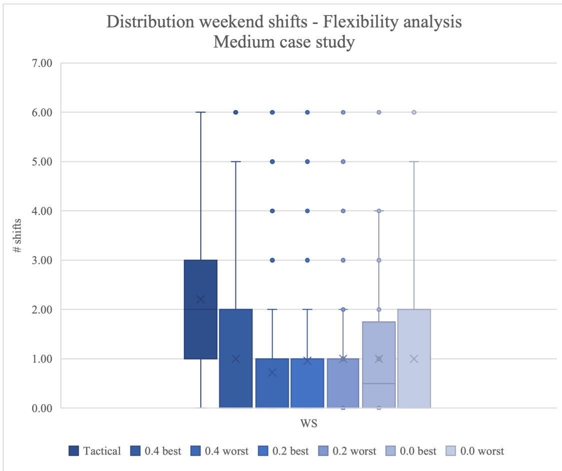
(a) Distribution of the remaining minutes of the full-time nurses for the tactical and operational schedule of case study 1.



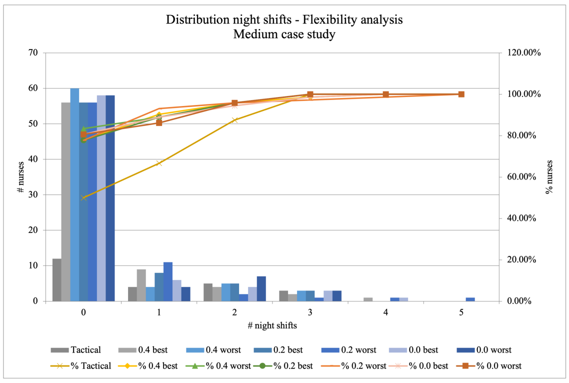
(b) Distribution of the hours assigned to the nurses with 0-hour contracts.



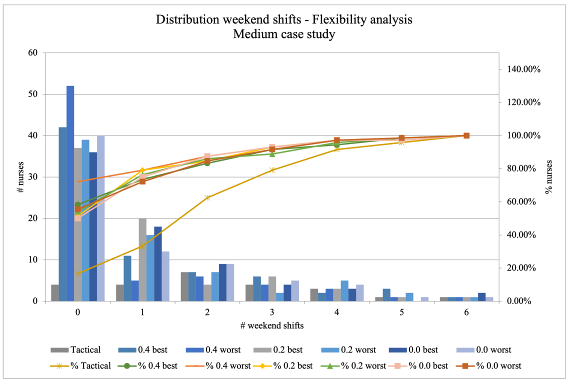
(c) Distribution of the night shifts within the tactical and operational schedule for case study 1.



(d) Distribution of the number of weekend shifts assigned.



(e) Distribution of the night shifts of the medium case study



(f) Distribution of the weekend shifts of the medium case study.

Figure 6.10: Medium case study: Experimental result of the proposed method for different flexibility parameters. Including the KPIs for the nurses with the distribution of night shifts, remaining minutes of the full-time nurses, the minutes assigned to employees with 0-hour contracts, and the distribution of the weekend shifts.

Large case study

The results of the large case study are shown in Table 6.8 and 6.11. It must be pointed out that none of the operational schedules meets the coverage requirements, but all the best solutions result in a decrease in total ORC violations. The results have shown that applying flexibility of 0.2 results in the largest decrease of total ORC violations, which equals a 62% decrease, but comes at the cost of an increase in total TRC violations of 157% and the largest increase of UQ shifts by 81%. This increase in UQ shifts results that the objective value of 0.2 being a little higher than that of the 0.4 flexibility, as the 0.4 flexibility has only an increase of 72% UQ shifts. All three best solutions that result in a decrease in ORC violations result in a decrease of 175%, 157%, and 25% of the total TRC violations for the 0.4, 0.2, and 0.0 flexibility parameters, respectively. In contrast with the other two case studies, the best objective value is obtained using flexibility of 0.4. However, the smallest fairness value and the least varied objective value is obtained using a flexibility of 0.0 yet has the lowest increase of TRC violations.

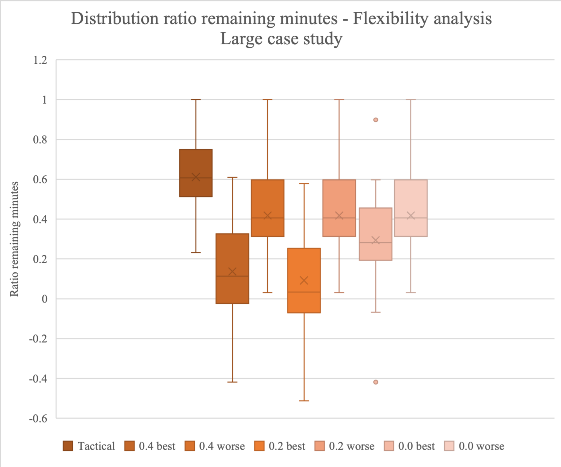
The 0.4 flexibility results in the highest increase in EOW, 2W, and SFS violations being 2.88, 7, and 1.15 times the amount of tactical TRC violations, respectively. It would be expected that this increase in EOW and 2W would resolve more undercoverage during the weekend. Nevertheless, the smaller increase using the flexibility of 0.2 results in a higher decrease in undercoverage of weekend shifts, being 98%, compared to the 0.4 flexibility, which has an 80% decrease. Additionally, the single 2W violation in the tactical schedule is resolved using the 0.2 and 0.0 flexibility. The best 0.2 flexibility results in the highest increase in RT and FRO violations being 4.00 and 2.33 times the amount of tactical TRC violations, respectively. The smallest increase of TRC violations is obtained when allowing no flexibility, but it also results in the smallest decrease in ORC violations with 44%.

Furthermore, it must be noted that the three best solutions have assigned shifts to all nurses and have assigned overtime, as shown in Figure 6.11a. The smallest variation in RM is obtained in the best solution using a flexibility of 0.0, indicating a more fair distribution of RM and contributing to the lowest fairness measure. Moreover, the best 0.2 and 0.4, and worst 0.4 solutions are the only three that have assigned hours to flex nurses. Nevertheless, the 0.4 worst solution results in undercoverage, which can be explained by the results in Figure 6.11a. These two best solutions had a significantly lower RM distribution than the 0.4 worst solution, with a significance level 0.05.

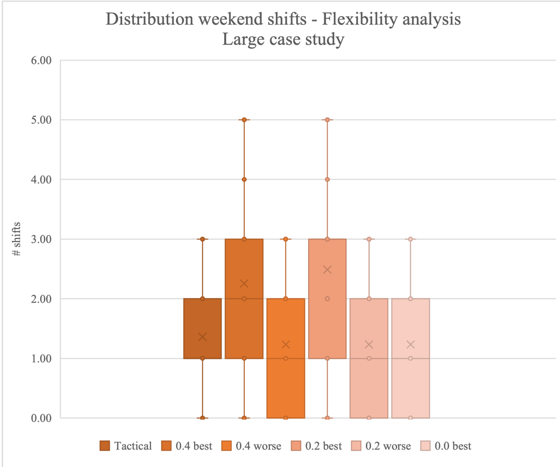
As seen in Figure 6.11c, the percentage of nurses that get none or a single WS assigned increases with the decrease of the flexibility parameter. Also, the three worst solutions follow the same WS distribution and increase in percentage. The two best 0.4 and 0.2 solutions assign a higher number of weekend shifts to most of the nurses, where more nurses are assigned three or four weekend shifts when allowing 0.2 flexibility, and more are assigned one or two when allowing for 0.4 flexibility. This explains the higher decrease in uncovered WS in when allowing a 0.2 flexibility.

Table 6.8: Large case study: Results for the flexibility analysis after five runs for flexibility parameters of 0.0, 0.2, and 0.4.

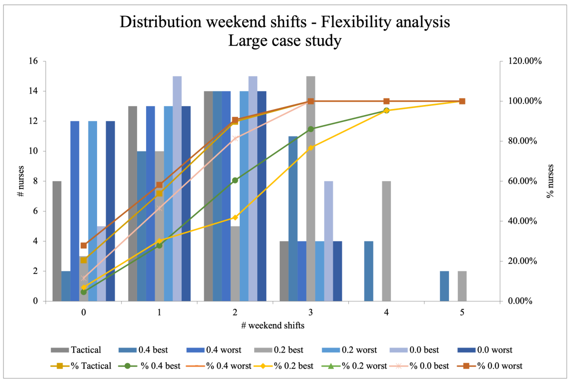
			TRC violations							ORC violations				
Schedule	Flexibility	Solution	RT	FRO	CWD	EOW	2W	SFS	Total TRC	Week	Weekend	UQ	QL3 hours	Total ORC
Tactical			3	6	0	8	1	33	51	151	55	43	37	286
Operational	0.4	Best	13	17	0	31	8	71	140	39	11	74	10	134
		Worst	3	6	0	8	1	32	50	152	55	43	37	287
	0.2	Best	15	20	1	28	0	67	131	25	1	78	6	110
		Worst	3	6	0	8	1	33	51	152	55	43	37	287
	0	Best	6	10	0	13	1	34	64	39	39	60	22	160
		Worst	3	6	0	8	0	34	51	152	55	43	37	287
$\Delta O - T$	0.4	Best	10	11	0	23	7	38	89	-112	-44	31	-27	-152
		Worst	0	0	0	0	0	-1	-1	1	0	0	0	1
	0.2	Best	12	14	1	20	-1	34	80	-126	-54	35	-31	-176
		Worst	0	0	0	0	0	0	0	1	0	0	0	1
	0	Best	0	0	0	0	-1	1	0	1	0	0	0	1
		Worst	0	0	0	0	-1	1	0	1	0	0	0	1
% (x100)	0.4	Best	3.33	1.83	0.00	2.88	7.00	1.15	1.75	-0.74	-0.80	0.72	-0.73	-0.53
		Worst	0.00	0.00	2.00	0.00	0.00	-0.03	-0.02	0.01	0.00	0.00	0.00	0.00
	0.2	Best	4.00	2.33	0.00	2.50	-1.00	1.03	1.57	-0.83	-0.98	0.81	-0.84	-0.62
		Worst	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	0	Best	1.00	0.67	0.00	0.63	0.00	0.03	0.25	-0.74	-0.29	0.40	-0.41	-0.44
		Worst	0.00	0.00	0.00	0.00	-1.00	0.03	0.00	0.01	0.00	0.00	0.00	0.00



(a) Distribution of the remaining minutes of the full-time nurses for the tactical and operational schedule of case study 1.



(b) Distribution of the number of weekend shifts assigned.



(c) Distribution of the number of weekend shifts assigned.

Figure 6.11: Large case study: Experimental result of the proposed method for different flexibility parameters. Including the KPIs for the nurses with the distribution of remaining minutes of the full-time nurses, the minutes assigned to employees with 0-hour contracts, and the distribution of the weekend shifts.

Overall outcome

Comparing the results of the three case studies in Table 6.5, it shows that a flexibility parameter of 0.0 results in the lowest objective value the small and medium case studies. Nonetheless, as seen in Figure 6.3, this parameter also results in a larger varied objective value but does result in a fully covered operational schedule in both best and worst solutions. This indicated that this variation occurs due to the number of TRC violations. In contrast, a flexibility parameter of 0.4 results in the smallest objective value for the large case study. However, there is more variability between the objective values when allowing a 0.4 flexibility for the large case study, as seen in Figure 6.8c. For all three case studies, it resulted that the decrease in ORC violations is at the cost of an increase of TRC violations. The largest decrease in ORC violations results for both the medium and large case study when using a flexibilty of 0.2 being 93% and 62%, respectively, and for the small case study when using a flexibility of 0.4 being 90%. The smallest increase in TRC violations, while resolving all undercoverage in the small and medium case study, is obtained using a flexibility of 0.0. These equals an increase of 25% and 20%, respectively, compared to the tactical schedule violations. However, it resulted in the smallest or same decrease of UQ shifts for the small case study and no or the largest increase in UQ shifts for the best and worst results of the medium case study. Also, the 0.0 flexibility parameter results in the smallest increase of total TRC violations for the large case study yet did not result in the largest decrease in total ORC violations. It appears that resolving more undercoverage comes at the cost of more UQ shifts. The latter can be explained as the goal is to meet the demand despite the fact that shifts are assigned to underqualified nurses. All three best solutions increase the number of SFS and EOW violations when decreasing the undercoverage. The SFS can be increased because shifts assigned that overlap with absenteeism are removed from the tactical schedule, which can result in a higher amount of unfavorable patterns. The increase in EOW violations can result because the priority was to fulfill most weekend shifts, and fewer nurses might be available during the weekend.

The distribution of RM varied more for the small case study than the medium case study, which can be due to more absenteeism in the first case. However, the distribution of RM is similar for the three flexibility parameters in the small case study. Yet, using a flexibility parameter of 0.0 in the medium and large case study resulted in a more fair distribution of RM indicated by the smaller variation. This can be the case as the method focuses on assigning the open shifts rather than rescheduling other shifts, thereby assigning shifts to nurses that have more RM. The distribution of RM can differ between the organizations due to the number of available nurses, the demand in minutes, and the absenteeism that must be taken into account. When only considering the solutions that resulted in a fully covered operational schedule, a flexibility of 0.0 resulted in the fairest schedule for the small and medium case study, indicated by the fairness measure in Table 6.5. Also, the lowest fairness measure is obtained when using a 0.0 flexibility in the large case study due to a more fair distribution of RM. At the same time, it resulted in the smallest increase of TRC violations meaning that there are fewer TRC penalties to distribute among the nures. Therefore, the best flexibility parameter to generate a fair schedule should be further evaluated when the schedule is fully covered.

To conclude, the decrease in ORC violations comes at the cost of an increase in TRC violations. It must be noted that all three flexibility parameters can result in a fully covered schedule for the small and medium case studies. However, this comes at the cost of an increase in TRC violations and UQ shifts. Therefore, choosing the appropriate parameter depends on the priorities of the organization, which should be reflected by the weights chosen.

6.5 Sensitivity analysis

The outcome of the final schedule depends on the preferences and requirements of the nurses, planner, and organization. To test the performance of the method with different priorities for TRC, we perform a sensitivity analysis on the three case studies. First, we will conduct a sensitivity analysis on the weights of the TRC. Subsequently, we perform a sensitivity analysis on the weight of the flexibility parameter for the medium case study. We will only perform the latter analysis on the medium case study as we assume the effect will be equal for the other two case studies. In contrast, the priorities chosen for the TRC are case-dependent and will therefore be conducted on each case study. Each experiment is run five times using the best parameters found for the SA from Section 6.1 and flexibility parameter of 0.2 on the tactical schedule. As mentioned, we ran the experiments five times to account for the SA algorithm's randomness and evaluate the method's robustness based on the best and worst solutions. From the experimental results, the best and worst solutions are selected based on the objective value of the operational schedule and presented in Tables 6.9, 6.10, and 6.11.

6.5.1 Sensitivity analysis weights TRCs

The input parameters that are included in the sensitivity analysis for the weights of the TRCs are w_{RT}, w_{FRO}, w_{RM}, and w_{EOW}. The parameters selected will alternately be assigned a weight equal to 10, while the other weights are set equal to 1. We look at the effect on the number of TRC and ORC violations for both schedules. A color scale is used to indicate the best and worst results regarding the number of TRC and ORC violations. Subsequently, we have determined the percentual difference between the best and worst-case solutions to identify more stable solutions.

Small case study

The best and worst solution of the sensitivity analysis on the w_{TRC} of the small case study can be found in Table 6.9. The first thing to point out is that by increasing w_{EOW}, the variation in ORC violations in both tactical and operational schedules also increases, and it results in the only operational schedule that is not fully covered in the worst-case. It also results in two times more FRO violations compared to the implemented weights. Similarly, as increasing w_{RM}, increasing w_{EOW} results in more UQ shifts in comparison when increasing w_{RT} or w_{FRO}. For w_{RM}, this results from the fact that the priority is given to assigning as many minutes as possible and will be chosen over fewer UQ shifts. More UQ shifts occur when increasing EOW, as the nurses that have the appropriate QLs might already work the weekend before or after. Then the priority is given to meeting the demand instead of fewer UQ shifts. It must be pointed out that increasing the w_{EOW}

results in no EOW and 2W violations in both solutions and schedules, similar to the performance of the implemented weights.

The only RT and most FRO violations in the tactical schedule, EOW violations in both schedules and CWD violations in the operational schedule are encountered when increasing the w_{RM}, as the priority is assigning the most minutes regardless of violating other TRC. However, this does result in minimum ORC violations, except for the UQ shifts, as mentioned above. Increasing w_{FRO} results in no RT, FRO, and CWD violations in the best and worst solutions of both schedules, whereas increasing w_{RT} results in one additional FRO violation and a more varied increase in EOW and SFS violations between the best and worst solution. Noteworthy is that both the best and worst operational solutions when increasing w_{RT} and w_{FRO} result in no undercoverage during the week, weekend, missing QL3 hours, and a minimum amount of UQ shifts and TRC violations. Compared to the implemented weights increasing w_{RT}, w_{FRO}, and w_{RM} more stable outcomes are generated, as the gap between the total number of TRC violations for best and worst solutions is smaller. The smallest percentual difference regarding the total TRC violations for the operational schedule is obtained when increasing w_{FRO} and w_{RM}, being 27% and 24% respectively, and for the total number of ORC violations by increasing w_{RT} being 0%. However, the smallest percentual difference for the TRC violations for the tactical schedule, being 13%, is obtained using the implemented weights.

Table 6.9: Small case study: Best and worst results of the sensitivity analysis on the weights of the TRC. The color scale indicates in green the best performance and red the worst performance per TRC and ORC. The percentual difference is the difference between the best and worst solution.

				TRC								ORC					
Weights	Schedule	Solution	Objective	RT	FRO	CWD	EOW	2W	SFS	Total TRC	% Difference	Week	Weekend	UQ	QL3 hours	Total ORC	% Difference
Implemented weights	Tactical	Best	0.0071	0	0	0	0	0	7	7	0.13	2	0	6	0	8	1.29
		Worst	0.0086	0	1	0	0	0	7	8		8	3	6	20	37	
	Operational	Best	0.0134	0	0	0	0	0	10	10	1.11	0	0	6	0	6	1.17
		Worst	0.0243	0	1	0	0	0	34	35		3	3	6	11	23	
RT	Tactical	Best	**0.0058**	0	1	0	1	0	9	11	0.17	5	0	2	15	22	0.00
		Worst	0.0081	0	1	0	3	0	9	13		7	2	3	10	22	
	Operational	Best	0.0094	0	2	0	2	0	7	11	0.67	0	0	2	0	2	0.00
		Worst	0.0188	0	0	0	7	1	14	22		0	0	2	0	2	
FRO	Tactical	Best	0.0091	0	0	0	2	0	10	12	0.09	1	0	0	5	6	0.40
		Worst	0.0115	0	0	0	2	0	9	11		0	0	4	0	4	
	Operational	Best	0.0095	0	0	0	2	0	11	13	0.27	0	0	0	0	0	2.00
		Worst	0.0164	0	0	0	3	1	13	17		0	0	4	0	4	
RM	Tactical	Best	0.0116	1	3	0	4	0	12	20	0.00	0	0	6	0	6	0.40
		Worst	0.0159	2	5	0	4	0	9	20		1	0	8	0	9	
	Operational	Best	0.0151	0	0	2	4	0	5	11	0.24	0	0	4	0	4	0.22
		Worst	0.0162	0	2	0	6	0	6	14		0	0	5	0	5	
EOW	Tactical	Best	0.0081	0	2	0	0	0	11	13	0.36	1	1	5	0	7	1.51
		Worst	0.0325	0	1	0	0	0	8	9		10	6	6	28	50	
	Operational	Best	**0.0069**	0	0	1	0	0	34	35	1.11	0	0	4	0	4	1.69
		Worst	0.0325	0	2	0	0	0	8	10		9	6	6	26	47	

Medium case study

Table 6.10 shows the best and worst solutions for the four experiments for the medium case study. First, it must be pointed out that none of the experiments resulted in violations of CNS and only a small number of violations of the CWD. Additionally, all operational schedules are fully covered and only have a few UQ shifts. Nevertheless, there is a high amount of other TRC violations in all solutions, and none result in zero violations.

Increasing the w_{RT} for the medium case study results in more RT and FRO violations in both the tactical and operational schedule compared to increasing w_{FRO}. In addition, increasing the latter results in the least ORC violations in both the best and worst tactical schedules. However, it does result in more or a similar amount of UQ shifts compared to the other experiments and more EOW, 2W, and SFS violations. The least amount of EOW and 2W occur when increasing w_{RM} and w_{EOW}, but it also results in more undercoverage during the weekend in the tactical schedule compared to when increasing w_{FRO}. None of the tactical schedules, except the best solution when increasing w_{RM}, have fully covered the QL3 hours because of the weights assigned to w_{nurses} and $w_{organization}$. The exception can be explained by the same reasoning as for the small case study, as increasing w_{RM} tries to minimize the RM and, in combination with the highest ORC weight assigned to w_{QL3}, the method prioritizes covering most of the QL3 hours. As most nurses will be assigned to shifts when increasing w_{RM}, the minimization of RM also results in the least 2W violations in both tactical and operational schedules in the best and worst solution and, therefore, in the smallest objective value of the operational schedule. It must be noted that most SFS violations occur when increasing w_{FRO}, which occurs due to the fact that if two shifts are assigned successively, the FRO constraint must be met. This constraint is immediately met if a nurse gets assigned a day off between two shifts, increasing the number of SFS. The most stable operational schedule regarding the total number of TRC violations is obtained when increasing the w_{RM} or w_{EOW}, as the percentual difference is 8%. The latter also results in the smallest percentual difference for TRC violations in the tactical schedule, with a difference of 20%.

Table 6.10: Medium case study: Best and worst results of the sensitivity analysis on the weights of the TRC. The color scale indicates in green the best performance and red the worst performance per TRC and ORC. The percentual difference is the difference between the best and worst solution.

				TRC									ORC						
Weights	Schedule	Solution	Objective	RT	FRO	CWD	CNS	EOW	2W	SFS	Total TRC	% Difference	Week	Weekend	Night	UQ	QL3 hours	Total ORC	% Difference
Implemented weights	Tactical	Best	0.1112	13	12	1	0	16	5	29	76	0.39	30	16	1	7	17	71	1.03
		Worst	0.0109	6	6	0	0	9	4	26	51		60	28	8	8	119	223	
	Operational	Best	0.0630	7	7	1	0	17	8	35	75	0.16	0	0	0	8	0	8	0.13
		Worst	0.0845	13	13	0	0	17	10	35	88		0	0	0	7	0	7	
RT	Tactical	Best	0.1289	3	9	0	0	13	4	32	61	0.54	22	15	1	10	17	65	0.67
		Worst	0.0504	13	18	1	0	22	14	38	106		38	20	7	8	57	130	
	Operational	Best	0.0434	6	8	0	0	17	9	42	82	0.16	0	0	0	9	0	9	0.20
		Worst	0.0529	10	14	0	0	24	14	34	96		0	0	0	10	1	11	
FRO	Tactical	Best	0.0658	1	4	0	0	15	6	48	74	0.30	16	7	2	9	17	51	0.08
		Worst	0.1222	11	10	2	0	25	14	38	100		18	5	1	12	11	47	
	Operational	Best	0.0426	4	7	0	0	18	8	49	86	0.12	0	0	0	3	0	3	1.25
		Worst	0.0720	9	8	1	0	25	13	41	97		0	0	0	13	0	13	
RM	Tactical	Best	0.0931	8	12	0	0	11	2	41	74	0.35	21	14	0	5	0	40	1.23
		Worst	0.0943	10	14	0	0	8	2	18	52		49	23	9	9	79	169	
	Operational	Best	0.0400	10	15	0	0	11	1	46	83	0.08	0	0	0	6	0	6	0.50
		Worst	0.0682	13	15	3	0	15	4	40	90		0	0	0	10	0	10	
EOW	Tactical	Best	0.1527	14	16	0	0	8	3	26	67	0.20	33	18	4	8	32	95	0.09
		Worst	0.1047	7	11	1	0	5	2	29	55		24	15	5	0	43	87	
	Operational	Best	0.0825	13	16	0	0	14	3	35	81	0.08	0	0	0	8	0	8	0.32
		Worst	0.0938	14	19	1	0	11	7	36	88		0	0	0	11	0	11	

Large case study

The results of the sensitivity analysis for the large case study are shown in Table 6.11. First, it must be pointed out that all schedules result in none or little CWD violations. As shown, none of the operational solutions result in a schedule that meets all coverage requirements. Nevertheless, all best operational solutions of the sensitivity analysis result in a minimum of 73% decrease in weekly undercoverage compared to the implemented weights. In addition, increasing w_{RM} and w_{EOW}

results in less undercoverage on the weekend and QL3 hours in the operational schedule. This comes at the cost of more UQ shifts and an increase in RT and FRO violations, as we want to minimize the RM regardless of other violations. Increasing w_{EOW} results in less weekend undercoverage and fewer EOW violations in the best solution as the method assigns more nurses to WS. However, there is a 182% and 44% percentual difference between the best and worst values for the number of EOW violations and undercoverage in the weekend.

Again, increasing w_{FRO} results in fewer RT and FRO violations compared to increasing w_{RT} in the best solution for both tactical and operational schedules. Increasing these two weights result in a higher amount of 2W violations. With the increased w_{FRO}, the method prioritizes minimizing the FRO, resulting in fewer RT violations and, in combination with the small amount of CWD violations, potentially causing the same nurses to work the weekends, leading to more EOW and 2W violations. For this case study, the smallest percentual difference in the total number of TRC violations is obtained when increasing w_{RT}, equal to 8% for the operational schedule. However, it results in the highest percentual difference for the total number of ORC violations, being 58%.

Table 6.11: Large case study: Best and worst results of the sensitivity analysis on the weights of the TRC. The color scale indicates in green the best performance and red the worst performance per TRC and ORC. The percentual difference is the difference between the best and worst solution.

				TRC								ORC					
Weights	Schedule	Solution	Objective	RT	FRO	CWD	EOW	2W	SFS	Total TRC	% Difference	Week	Weekend	UQ	QL3 hours	Total ORC	% Difference
Implemented weights	Tactical	Best	0.1468	1	5	0	19	2	53	80	0.25	97	28	64	16	205	0.09
		Worst	0.1787	13	17	0	18	2	53	103		79	28	64	16	187	
	Operational	Best	0.2060	1	5	0	19	1	53	79	0.35	97	28	64	16	205	0.02
		Worst	0.2291	13	17	0	18	0	65	113		78	36	68	19	201	
RT	Tactical	Best	0.1365	8	13	1	26	10	45	103	0.23	99	28	54	15	196	0.17
		Worst	0.1655	7	11	0	21	4	39	82		117	31	69	16	233	
	Operational	Best	0.1899	8	13	1	25	10	24	81	0.08	26	26	54	19	125	0.58
		Worst	0.2284	6	10	0	25	4	43	88		112	27	71	16	226	
FRO	Tactical	Best	0.0072	5	6	0	28	8	51	98	0.35	92	20	61	13	186	0.17
		Worst	0.0070	12	15	2	30	10	71	140		70	13	74	0	157	
	Operational	Best	0.1995	5	6	0	30	10	34	85	0.47	18	18	62	13	111	0.40
		Worst	0.2450	11	14	0	30	10	72	137		70	13	74	10	167	
RM	Tactical	Best	0.0234	15	19	0	25	6	67	132	0.14	51	14	80	10	155	0.17
		Worst	0.0151	15	18	0	18	7	57	115		70	22	82	9	183	
	Operational	Best	0.1919	14	18	0	25	6	34	97	0.15	14	14	79	10	117	0.43
		Worst	0.2116	15	18	0	18	5	57	113		70	22	81	9	182	
EOW	Tactical	Best	0.0065	3	6	0	8	1	33	51	0.77	151	55	43	37	286	0.44
		Worst	0.0069	15	18	0	18	7	57	115		70	22	82	9	183	
	Operational	Best	0.2212	15	20	1	28	0	67	131	0.15	25	1	78	6	110	0.49
		Worst	0.2596	15	18	0	18	5	57	113		70	22	81	9	182	

Overall outcome

The results of the sensitivity analysis for all three case studies have shown that increasing w_{FRO} results in both a decrease of FRO violations as RT violations, whereas increasing w_{RT} results only in a decrease of RT violations. Furthermore, increasing w_{RM} results in less undercoverage in both tactical and operational schedules compared with the other experiments. However, this comes at the cost of an increase in RT and FRO violations and a decrease in 2W violations for all three case studies, and an increase in UQ shifts for the small and large case studies. Increasing w_{EOW} resulted, as expected, in fewer EOW violations but did not necessarily result in fewer 2W violations.

6.5.2 Sensitivity analysis weights flexibility parameter

The final sensitivity analysis is performed by relaxing the weights of the flexibility parameter, which increases the chance of accepting a worse solution in the SA algorithm or a solution with violations of allowed percentual flexibility. This is done for the medium case study, using a flexibility parameter of 0.2 and changing the weight to 1,000, 10, and 1. We look at the increase and decrease in the number of TRC and ORC violations, denoted by $\Delta O - T$, and the change in percentage. Table 6.12 shows the best and worst results of the sensitivity analysis for three different weights. Again, the same tactical schedule is used as input for the operational schedule. Figure 6.12 shows the distribution of the operational objective function.

As seen in Figure 6.12, a weight equal to 10 results in the least varied objective value compared to a weight of 1 or 1000. The lower the weight, the higher the chance of accepting a worse solution in the SA algorithm. In the case of a weight equal to 1, the solution space becomes larger and can result in better but worse solutions, as seen in Table 6.12. Additionally, by relaxing the weight to 1, there is a 33% and 17% decrease in FRO violations and no increase in RT violations for the best and worst solutions. However, there is a higher variation between the best and worst solution regarding ORC violations, where the latter results in a not fully covered schedule. To have a more stable outcome, a higher weight should be implemented.

The results of the other two experiments are comparable, and both result in operational schedules without undercoverage and only a single or no increase in the number of UQ shifts. By relaxing the weight to 10, the change of accepting worse solutions is larger than when a weight of 1000 is used. However, the best solution of implementing a weight of 10 results in 0.46 EOW and 2.5 more 2W violations, whereas the worst solution is comparable with the best solution of implementing a weight of 1,000 regarding RT, FRO, and SFS violations. In these two solutions, the number of violations increases by 200%, 83%, and almost 10% compared to the tactical schedule. Only the weight of 10 resulted in a 22% and 8% decrease in SFS violations in both the best and worst solutions.

More experiments should be conducted to find the best weight for the flexibility parameter. The findings indicate that increasing the weight leads to more stable outcomes, as evidenced by the minimal difference in percentual increase or decrease in Table 6.12.

Table 6.12: The best and worst results for the sensitivity analysis on the weight of the flexibility parameter for the medium case study after 5 runs. The color scale indicates the highest and lowest decrease or increase in percentage.

				TRC violations								ORC violations					
Schedule	Weight	Solution	Objective value	RT	FRO	CWD	CNS	EOW	2W	SFS	Total TRC	Week	Weekend	Night	UQ	QL3 hours	Total ORC
Tactical			0.0833	3	6	0	0	13	2	36	60	57	27	5	1	50	140
Operational	1000	Best	0.060	9	11	0	0	17	5	32	74	0	0	0	2	0	2
		Worst	0.082	7	10	2	0	16	5	40	80	0	0	0	2	0	2
	10	Best	0.060	6	8	0	0	19	7	28	68	0	0	0	1	0	1
		Worst	0.070	9	11	2	0	14	5	33	74	0	0	0	2	0	2
	1	Best	0.056	3	4	4	0	18	6	38	73	0	0	0	2	0	2
		Worst	0.092	3	5	0	0	13	5	34	60	61	27	5	1	50	144
$\Delta O - T$	1000	Best	-0.0238	6	5	0	0	4	3	-4	14	-57	-27	-5	1	-50	-138
		Worst	-0.0015	4	4	2	0	3	3	4	20	-57	-27	-5	1	-50	-138
	10	Best	-0.0236	3	2	0	0	6	5	-8	8	-57	-27	-5	0	-50	-139
		Worst	-0.0136	6	5	2	0	1	3	-3	14	-57	-27	-5	1	-50	-138
	1	Best	-0.027	0	-2	4	0	5	4	2	13	-57	-27	-5	1	-50	-138
		Worst	0.0082	0	-1	0	0	0	3	-2	0	4	0	0	0	0	4
% (x100)	1000	Best	-0.29	2.00	0.83	0.00	0.00	0.31	1.50	-0.11	0.23	-1.00	-1.00	-1.00	1.00	-1.00	-0.99
		Worst	-0.02	1.33	0.67	2.00	0.00	0.23	1.50	0.11	0.33	-1.00	-1.00	-1.00	1.00	-1.00	-0.99
	10	Best	-0.28	1.00	0.33	0.00	0.00	0.46	2.50	-0.22	0.13	-1.00	-1.00	-1.00	0.00	-1.00	-0.99
		Worst	-0.16	2.00	0.83	2.00	0.00	0.08	1.50	-0.08	0.23	-1.00	-1.00	-1.00	1.00	-1.00	-0.99
	1	Best	-0.32	0.00	-0.33	4.00	0.00	0.38	2.00	0.06	0.22	-1.00	-1.00	-1.00	1.00	-1.00	-0.99
		Worst	0.10	0.00	-0.17	0.00	0.00	0.00	1.50	-0.06	0.00	0.07	0.00	0.00	0.00	0.00	0.03

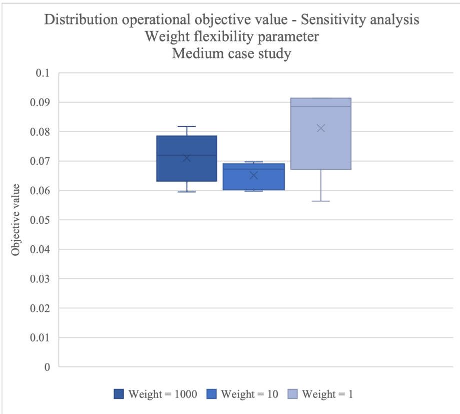


Figure 6.12: Distribution of the objective value after five runs for the sensitivity analysis on three different weights for the medium case study.

6.6 Conclusion

To evaluate the performance of the proposed method and answer the fourth research question, we compared the results with the current performance from practice using three case studies. This was evaluated, allowing a flexibility parameter of 0.2, indicating that 80% of the tactical schedule should remain in the operational schedule. When applying the proposed method for the small and medium case study, the objective values for the tactical and operational decreases by 96%-99% due to a 38-56% reduction in TRC violations for the small case study. The operational objective for the medium case study reduces by 53%-65% due to a 47%-53% decrease in ORC violations and 88%-100% decrease in CWD, and 100% decrease in CNS violations. The results showed that the proposed method can decrease the amount of RT, FRO, and SFS violations in both tactical and operational schedules for the small and large case studies. However, the results are case-dependent, as the decrease in the number of TRC violations differs per case study as these depend on the current performance and priorities of the organization. The method does not outperform the performance of the large case study as there is undercoverage and a high number of UQ shifts in the operational schedule.

Hereafter, a flexibility analysis is performed by changing the flexibility parameter to 0.4 and 0.0. This provides an answer to the question *Is including flexibility a good way to incorporate fairness in the operational schedule, and how does it impact the outcome of the schedule?* For each case study, a random tactical schedule is constructed and is used as input to evaluate the difference in performance using the different parameters. The results showed that allowing no flexibility using a flexibility parameter of 0.0 resulted in the best operational objective value for the small and medium case study. Also, it constructs the fairest schedules when evaluating the fairness measure for the solutions that result in a fully covered schedule. In contrast, the best objective value for the large case study was obtained with a parameter, but none of the solutions resulted in a fully covered schedule. All results showed that a decrease in undercoverage comes at the cost of TRC violations, where the total number of TRC violations increased by 75%, 20%, and 175% for the best solutions of the small, medium, and large case studies, respectively.

A sensitivity analysis was performed on the following four input parameters: w_{RT}, w_{FRO}, w_{RM}, and w_{EOW} to answer the sixth research question. This was to evaluate the effect of the weights chosen on the outcome of the schedules and the robustness of the schedule. All three case studies hold that increasing w_{FRO} resulted in both a decrease in FRO violations and a decrease in RT violations. Increasing w_{RM} resulted in more coverage in the tactical schedule and fewer 2W violations. However, this comes at the cost of an increase in other TRC violations for all case studies and a higher amount of UQ shifts in the small and large case studies. At last, we performed a sensitivity analysis on the weights of the flexibility parameter for the medium case study. The results showed that by increasing the weight, more stable outcomes are generated.

All these results have shown that the results are case-dependent. The weights assigned to the constraints affect the outcome of the schedule, and there is a trade-off between the number of ORC and TRC violations for the different schedules. Therefore, the weights and flexibility parameters should be chosen based on the organization's priorities.

Chapter 7

Conclusion and Discussion

The chapter gives the conclusion to the overall research question based on the results of this research. Additionally, the practical and theoretical contributions of this research are described. Followed by the limitations of the study, regarding the lack of data and the simplifications and assumptions made for the method. We provide recommendations for Nedap on how they can implement the findings of this research in their software and what should be considered when extending the method. Finally, we provide opportunities for future research.

7.1 Conclusion

Due to the lack of nursing staff and an increase in demand for care due to the ageing population, care organizations must use their resources efficiently and address the organizational support of nurses to enhance employee satisfaction.

Nowadays, Nedap facilitates residential care organizations with their software to generate nurses' schedules. However, often these are still made manually and later implemented in the software and do not aim at optimality. Therefore, this research is conducted to support Dutch residential care organizations in constructing nurses' schedules to enhance job satisfaction. These schedules should meet the needs of the nurses and organizational requirements while complying with law legislation.

The objective of this research is: *Develop a nurse scheduling method that supports Dutch residential care organizations in constructing fairness-enhanced tactical and operational schedules to sustain a better work-life balance and increase employee satisfaction.* The results have shown that the proposed method can generate tactical and operational schedules within a short amount of time. Additionally, it distributed the penalties due to violations of TRC fairly among the nurses. While it does not necessarily outperforms the current practice in terms of TRC and ORC violations and does not always result in a valid schedule, it does create opportunities to support planners in constructing these schedules in a shorter time. Moreover, it provides more insight into the effect of the flexibility parameter used on the tactical schedule to construct operational schedules.

We have conducted multiple experiments to evaluate the performance of the method and the effect of the flexibility parameter. The results from Chapter 6 have shown that there is a trade-off between meeting demand or providing opportunities for nurses to have a better work-life balance. When using a flexibility parameter of 0.0, which means that no tactical reassignments are allowed, we provide nurses with the most predictable and stable schedule. When allowing flexibility more organizational requirements reflected by the ORCs are met, but comes at the cost of an increase in total TRC violations. Depending on the goal of the organization and the preferences of the nurses, the preferred parameters must be implemented.

First, the performance was evaluated based on the current performance of three different size case studies. The method outperformed two out of three case studies, resulting in fully covered operational schedules and reduced the number of RT, FRO, CWD, CNS, 2W and SFS violations. The total number of TRC violations can be reduced by 76%, 1%, and 22% in the operational schedule for the small, medium and large case study respectively. Additionally, the proposed method distributes the RM more fairly as there is less variation in the distributed minutes. This arises from that the method prioritizes to assign minutes to regular nurses instead of flex nurses, and thereby assigning shifts to all regular nurses as opposed to the current performance. Yet, the current performance of all three case studies outperforms or results in an equal number of EOW violations. To conclude, the method can improve the results of both schedule types for a small size organization where the total number of TRC can be reduced by 17%-76% in the operational schedule in the worst and best case. Also, the operational schedule of medium size organizations can be improved with a reduction of ORC violations of 47%-53%. This outperforms the method of Kiermaier et al. [29], which resulted in a reduction of 10% of undercoverage. However, the method did not result in a valid schedule for a large size organization with a smaller ratio of available minutes and the demand in minutes.

Hereafter, a flexibility analysis was performed to evaluate the effect of the flexibility parameter and to determined if allowing flexibility can result in more fair schedules. A parameter of 0.0 resulted in the valid and fairest schedules for the small and medium case study. However, there was a higher variation in the objective function, indicating less stable outcomes. Using a flexibility parameter of 0.0 also resulted in a more fair distribution of RM, indicated by the less varied distribution. A flexibility of 0.4 resulted in the best objective value for the large case study. Yet, allowing flexibility of 0.2 for the medium and large case study resulted in the largest decrease of total ORC violations, being 93% and 62%, respectively. The largest decrease in total ORC violations for the small case study was encountered when implementing a flexibility of 0.4, resulting in a 90% decrease. For all three case studies, the smallest increase in total TRC violations was realized using a 0.0 flexibility, 25%, 20%, and 75%, respectively. At the same time, it also resulted in a valid schedule for the small and medium case study. Furthermore, the results showed that meeting demand comes at the cost of an increase in UQ shifts and SFS and EOW violations.

At last, a sensitivity analysis was performed on four different weights of TRCs and the weight of the flexibility analysis. It was shown that increasing w_{FRO} minimizes both RT and FRO violations, whereas increasing w_{RT} increased the number of FRO violations. Moreover, increasing w_{RM} results in less undercoverage, but as in the operational schedule, this comes at the cost of an increase of other TRC violations.

Concluding, this research developed a method that constructs tactical and operational schedules, which provides opportunities to enhance employee satisfaction through an improved work-life balance and fairness-enhanced scheduling. While it does not outperform the current performance of all case studies, it does show improvement in the number of or both TRC and ORC violations. By validating the performance in practice, the proposed method can be improved and meet the unique preferences of individual care organizations.

7.1.1 Scientific contribution

To the best of our knowledge, we are the first to develop a method that uses the tactical schedule to construct the operational schedule, while preserving a certain percentage of the tactical schedule denoted as the flexibility parameter and accounting for fairness. The developed method used a constructive heuristic to find an initial tactical schedule, which is optimized using SA. Hereafter, after removing the planned absenteeism, the same SA procedure is used to optimize the tactical schedule for the allowed flexibility to generate the operational schedule. The experiments conducted in the flexibility analysis have shown the effect of the flexibility parameter on the outcome of the operational schedule. Where two out of the three cases result in the lowest objective value and fairest schedule when implementing a flexibility parameter of 0.0 and the smallest increase of total TRC violations. However, the outcomes are sensitive to the values of the parameter tuning and the weights assigned to the SC, as shown by the results of the sensitivity analysis. The number of TRC violations depends on the weights implemented, where increasing or decreasing a weight influences all TRC and ORC violations. This insight provides opportunities for further research, which will be discussed in 7.4.

7.1.2 Practical contribution

In practice, planners are not concerned with the objective value of the solution but are only interested in the schedule itself. Therefore, we present the number of TRC and ORC violations encountered for the schedules. Furthermore, our proposed method can support planners in generating tactical and operational schedules in a shorter time, as indicated by Table 6.5. To give insight into the outcome of the schedule, we do not only present the schedule itself but also include the outcomes of the KPIs defined in Section 4.2 to give a better overview to assess the quality of the schedule. Additionally, this research gives insight into the percentage of flexibility that can be allowed on the tactical schedule to generate operational schedules. Based on the preferences of the organization, nurses can be assured that a certain part of their fixed assignments will remain preserved, resulting in opportunities for a better work-life balance. However, it must be noted that the outcomes are case-dependent and the optimal flexibility should be evaluated for each case, as indicated by the varying optimal parameter settings as discussed in Section 6.4. Subsequently, the weights should be adapted based on the priorities of the organizations, nurses, and planners.

7.2 Discussion

We will reflect on the results stated in the conclusion and relate them to the literature. Also, we will discuss the limitations to this research, as due to time restrictions and a lack of data, not all requirements and agreements could be taken into account.

7.2.1 Discussion results

Hadwan and Ayob [48, 49], Lavygina et al. [52], Jafari and Salmasi [57] have developed different methods to solve the NSP using SA and tested the performance using real-world cases. Their results showed that their developed methods could generate valid schedules in a short amount of time and meet additional requirements. However, none of these studies have used a fairness measure as an objective. Also, these studies focused on generating one schedule that meets the periodic demand. In this research, we have included fairness as an objective and implemented a two-stage scheduling approach. Therefore, it can be the case that the method did not find valid schedules for the large case study in this research and did not outperform the current practice in terms of TRC violations. This is similar to the results of Lin et al. [59], where the proposed method did not perform better than the manual schedule in terms of consistency due to the fact that constraints conflict which each other and are hard to satisfy simultaneously. They have shown that the algorithm does not violate the hard constraints independent of the number of nurses is sufficient, which also holds for the proposed method in this study. Despite the fact that not all coverage requirements are met, the hard constraints are not violated. For example, the large case study had only two QL3 nurses available, which resulted in more undercoverage of QL3 shifts, but none of the shifts were assigned to underqualified nurses. This would suggest that the HC will not be violated if the method generates a schedule for an organization with insufficient staffing. However, the schedule constructed will not be valid. Regardless, the number of TRC violations will increase as the method tries to assign as many shifts as possible. This is indicated by the results of the flexibility analysis of the large case study, where the number of TRC violations increases by 25%-175% as the number of ORC violations decreases by 44%-62% compared to the tactical schedule, depending on the flexibility allowed. Additionally, the use of fairness as an objective can also result in a non-valid schedule, which can also be augmented by the results of Kletzander and Musliu [63], as they showed that several studies that have implemented fairness as an objective did not all result in valid schedules.

Another factor that can influence the outcomes of the results is the parameter values for the SA algorithm and the weights assigned. To determine the best parameter values, we have conducted a parameter tuning procedure by running the algorithm for a maximum of five minutes to construct the tactical schedule five times while scheduling 80% of the contract hours, as discussed in Section 6.1. Subsequently, these values are also used in the SA algorithm when optimizing the operational schedule. Since it is unknown whether these best values are also optimal for the operational schedule, different results could have been found if we would tune the parameters specifically for the operational schedule.

As shown in Appendix C.1.3, the graphs for the parameter tuning for the large case study do not follow the true shape of a SA graph, with a clear distinction in the diversification and intensification. This influences the quality of the optimal tactical and operational schedule and the chance to find near-optimal solutions. Also, as shown in Figure 6.2a, none of the weights result in a schedule without an $p_{organization}$ equal to zero. These results can be attributed to various factors and explanations. First, the large case study has only two QL3 nurses available. Compared to the other case studies, this is relatively low. Consequently, this can result in more permanent undercoverage of QL3 hours and UQ shifts. This can be due to the assumptions made regarding the hour types and could have been misclassified, resulting in exceptionally high values of UQ in the current performance in Table 6.3. At last, we have implemented four operators to find a neighbour solution. SwapRandomDay and SwapSameDay both have a maximum number of attempts to find a neighbour solution to prevent the algorithm from getting stuck on finding a solution. However, if this maximum is reached too often, it prevents the algorithm from escaping the local optima, as the neighbour solution will be restored. In addition, to tune the parameters, we have set the running time to a maximum of 5 minutes. Combining these two facts can result in a cut-off during the parameter-tuning phase and limit the algorithm to finding the best parameter values.

The weights assigned to the constraints affect the outcome of the schedules, as also shown by the sensitivity analysis conducted in Section 6.5. These weights represent the priorities defined by the organizations, planners, and nurses. In the proposed method, the weights assigned to w_{TRC} and w_{ORC} are equal for all three case studies. This implies no differentiation in relative importance or priority assigned to these constraints, regardless of the organization's specific goals. The weights are chosen according to the overall outcome of the interviews. Nonetheless, in practice, the weights depend on the organizations' goals and the preferences of the planner and nurses. As indicated by the results of the current performance in Table 5.2 Section 5.3, the organizations have different priorities regarding violations of TRCs, e.g., the large case study has numerous violations in the current schedule, indicating that this constraint is not considered as important.

Additionally, as shown in the sensitivity analysis, increasing w_{FRO} results in different violations than when increasing w_{RM}. The w_{TRC} are set equal for the tactical and operational schedule. However, the goal of these two differs, which can be reflected by the weights assigned in the method. If an organization would minimize the number of RT and FRO violations in the tactical schedule, one should increase w_{FRO}. Whereas the goal of the operational schedule is to meet the demand regardless of some TRC violations, one can choose to increase w_{RM}.

Currently, this research tried to enhance fairness by treating all nurses equally and distributing the penalties among the nurses, which aims at group fairness. To better reflect the nurses' priorities, individual weights can be taken into account. This also ensures individual fairness, as individual aspects are included, and the weights reflect the individual perception of fairness [45]. Another way to determine the nurses' preferences is using self-scheduling as done by Ouelhadj et al. [43], Smet et al. [53], Tsaia and Leeb [55], Lin et al. [59]. However, instead of letting nurses make a schedule each period, we will propose that nurses can make an individual schedule each year or half a year. These individual schedules and occurring patterns can be considered as preferences. This would also give insight if a nurse would prefer to work more consecutive shifts, does not prioritize the FRO constraints as important, or would like to work two weekends in a row instead of EOW.

Moreover, we only considered short-term fairness, assuming there is no previous or upcoming period. However, to guarantee long-term fairness, previous schedules should be considered when generating operational schedules [28]. It is important to keep track of and accumulate the number of additional requests that have been granted overtime to guarantee long-term fairness, as done by Wolbeck and Kliewer [17]. Here, long-term fairness is assured as the granted additional requests are accumulated with the satisfaction of the previous period.
Furthermore, the results of Chapter 6 have shown that the distribution of WS and NS, in the case of the medium case study, is evenly distributed over all nurses, as an even portion gets the same amount of WS assigned. However, this does not reflect the nurses' KPI of a fair WS distribution. It would be fairer if every nurse would work the weekend instead of 25% of the nurses getting four WS assigned. When applying another objective function or adding an additional constraint, a more fair distribution of WS and NS can be achieved, e.g., $min - max$ objective or a constraint that considers the ratio of weekend shifts and week shifts. This can be explored in future research. Moreover, the proposed method outperforms or has an equal performance as the current practice regarding EOW violations. Among others, Hadwan and Ayob [48, 49], Lavygina et al. [52], Ceschia et al. [60] have defined shift patterns beforehand that can be assigned instead of assigning single shifts as in our approach. When using shift patterns, the number of FRO and SFS can also be reduced as

there is no variability in which single shifts are assigned. By minimizing the SFS violations, more stable schedules can be generated [48, 49]. As these shifts are assigned in predetermined patterns, this can also help to distribute the WS and NS more fairly, thereby reducing the number of EOW and 2W violations. Another way to distribute the WS and NS more fairly is by considering the approach of Hadwan and Ayob [49]. They have used a semi-cyclic shift approach where only the night shifts are allocated cyclically. This reduces the number of shift patterns, and the fair allocation of night shifts becomes more manageable. However, in the method developed, the assumption is made that all shift types are considered when constructing the tactical schedule. Future research can be conducted to explore the effects of the distribution of WS and NS when only considering these shifts when constructing the tactical schedule.

At last, we did not include the maximum allowed working time per week or period of seven days defined in the law legislation, which is considered by Lavygina et al. [52]. This can be implemented to distribute the workload more evenly within a week, contributing to fairer schedules. This might also reduce the number of EOW and 2W violations. For example, when a nurse works one weekend, this constraint, combined with minimizing SFS, would prevent a nurse from working two weekends in a row as the maximum allowed working time in seven days can already be reached.

7.2.2 Limitations to this research

The method presented in this research is subject to several limitations, mostly due to a lack of data and time restrictions also, when implementing the constraints, several simplifications and assumptions had to be made as discussed in Sections 4.4 and 5.1.

7.2.3 Lack of data

To generate nurse schedules, it is crucial to know the staffing demand, coverage requirements, and contractual agreements. To implement the proposed method, historical schedules from residential care organizations are utilized as input data to identify this information. The data is gathered using the software of Nedap. Nowadays, Nedap does not facilitate automated scheduling, therefore, the data gathered is not complete, and assumptions have to be made during the implementation process. Moreover, given that each organization uses the Nedap software differently, there are variations in how data is stored. We will discuss the data that is missing related to the nurses and shift requirements, which lead to the assumptions made in Section 5.1.

First, individual contract agreements are not included in the data set. Consequently, we could not take those into account, including the number of additional consecutive working days, working only night shifts, and the sequence or number of weekends a nurse can work. Therefore, the assumption is that all nurses have the same contract agreements based on the law legislation. This also influences the number of violations calculated for the current practice in 5.3 and in the constructed schedules. Additionally, the constructed schedules could be valid when scheduling according to individual contract agreements.

Secondly, nurses are allowed to make additional requests, which may or may not be granted. However, these are not stored within the data and could not be taken into account as an additional constraint. The distribution of additional free days cannot be used as a fairness measure. However, it is key to enhancing employee satisfaction and guaranteeing long-term fairness [17].

Due to experience, planners know by heart which shifts need to be covered 24 hours a day and are classified as QL3 shifts. However, these are not specified in the data or stored in the software. Therefore, the QL3 shifts are now identified based on the interviews with the care organizations. In the data, hour types represent the skills required for the specific shifts. However, these differ per organization. Therefore, assumptions had to be made as mentioned in Section 5.1. The results of Section 5.3 indicates that nurses are assigned to shifts without meeting the specific requirements. However, it is unlikely that this occurs in a tactical schedule, as this would mean that a schedule permanently contains UQ shifts, thereby not delivering the right care.

Section 5.2 described the data of the case studies and pointed out that the data of the small and large case studies did not include night shifts. Based on the available data, it is unknown how the allocation of night shifts is organized as there may be a separate team that covers the night shift, if there are no night shifts at all, or if there is a separate night shift schedule that is stored in the software or on paper. If the latter two are true, the generated tactical and operational schedule could be infeasible due to violations of HC1 or result in more SC violations, e.g., more RT, FRO, and CWD violations.

7.2.4 Limitations of the method

As described in Section 4.4, multiple assumptions and simplifications are made in order to develop a method in a reasonable time that generates fair schedules for residential care organizations.

7.2.4.1 Assumptions of the method

First, it is assumed that there is no period before or after the period currently considered in the scheduling process. However, when generating a tactical schedule that is repeated over a predetermined number of periods, we should take into consideration the shifts assigned in the first and last week as done by Jafari and Salmasi [57]. Because the allowed shift type to be assigned during the beginning of the first and end of the last week impact each other, i.e. if a nurse is assigned four-day shifts in the first four days of the first week, then in the last week, the nurse can only be assigned one day shift in the weekend without violating the maximum consecutive shifts and cannot be assigned a night shift without violating the forward rotating order legislation.

The tactical schedule assumes that all nurses with contract hours assigned are scheduled. But in practice, there is a distinction between large and small contracts, where the latter is scheduled for a different percentage or not scheduled at all in the tactical schedule. First, this could influence the distribution of RM. As indicated by the results, small contracts rapidly result in lower RM values when one or two shifts are assigned. This implies that the RM are not distributed fairly, as there is more variation due to these small ratios. Secondly, this provides more flexibility in shift assignments in the operational schedule. According to the planners from practice, nurses with small contracts are less complicated to assign, e.g., the number of number plus hours that has to be compensated in the next planning periods is smaller as the contract agreements are easier met when only a few shifts are assigned, and due to fewer shift assignments also reducing the chance of TRC violations.

7.3 Recommendations

Nedap can improve its services by implementing automatic scheduling measures to benefit its customers. In this research, we have created a two-stage scheduling approach based on insights and requirements gathered from interviews with eight residential care organizations and literature. Nedap can use these results to extend its software. We suggest the following recommendations, which include the objective regarding the tactical and operational schedule; additional constraints to implement in the method; opportunities to support individual and long-term fairness; organization of the data; and validation of the method with practice.

First, from the interviews, it became clear that there is a difference between the priorities for the tactical and operational schedules. As mentioned, the goal of care organizations is to meet the required demand. This is done by constructing high-quality schedules that meet the staffing levels while meeting the preferences of the nurses. There is a distinction between the main goal of the tactical and operational schedules. With the tactical schedule, the aim is to provide nurses with a way to improve their work-life balance, as the schedules are predictable and known in advance. Second, we aim to have no law legislation violations and have a fair distribution of shifts based on the agreed contract hours. The tactical schedule represents the ideal work pattern for the nurses, with at least one fixed day off and that complies with the contract agreements. It is recommended to increase the weights regarding FRO, as this decreases the number of FRO and RT violations. Furthermore, when the individual agreements are considered, the results can be reevaluated and appropriate weights for the other constraints can be assigned. At the same time, the goal of the operational schedule is to deliver the required care within that specific period. Therefore, if the staffing levels are in order, the main goal is to ensure that all shifts are assigned to those that meet the coverage requirements. Exceptions to law legislation are acceptable but should be minimized. Therefore, the weight for the RM should be increased, but the individual agreements should be taken into account.

In this study, we aim to minimize the number of tactical reassignments when constructing the operational schedule to preserve the benefits of the tactical schedule by including a flexibility parameter. By doing this, we aim to provide a stable, predictable schedule for each nurse, promoting a better work-life balance and eventually increasing job satisfaction. Depending on the goals of the organization and the preferences of the nurse, an appropriate parameter can be chosen. However, to have the fairest schedule and the most stable schedule, we recommend for the small and medium case study to apply an 0.0 flexibility parameter. To define the parameter for the large case study, further research should be conducted as the method did not find a valid schedule and the fairness measure is influenced by the number of TRC violations.

To ensure optimal support for their customers, we recommend Nedap consider including the following additional rules within the proposed method, as not all constraints are currently accounted for or addressed.

To further meet the requirements from practice, the method should be extended by taking the additional requests for free days into account. As mentioned in Section 7.2.3, these are not considered in this research. However, in practice, care organizations aim to fulfill these requirements in order to satisfy their nurses, and it is thus crucial to meet requirements from practice. Furthermore, Nedap should ensure that the data is organized and complete. As discussed in Section 7.2.3, some essential data is missing, including individual contract agreements, additional absenteeism for the operational schedule, and the QL of the responsible shift, e.g., individual agreements should be gathered to meet the specific individual needs of the nurses and to account for individual fairness. The lack of data limits the proposed method to create high-quality schedules that meet the unique requirements of each nurse. In addition, a better understanding of the hour types and QL classification can ensure that the shift requirements are met, and shifts that do not meet the requirements are correctly identified. This can result in a more precise method to optimize the scheduling process.

Additionally, we recommend consideration of the following constraints. In practice, certain organizations consider the Friday late and night shifts as weekend shifts and try to assign those to nurses working the weekend. By including this rule, we could have a better distribution of the weekend shift and optimize scheduling in blocks. The constraints implemented in the method focused on a planning horizon of four weeks. However, there are law legislations that consider a longer planning horizon of a year. Such rules include that a nurse should have 22 weekends off in a year, should work an average of 48 hours per week in a period of 16 weeks, and work at most 35-night shifts in a period of 13 weeks. Furthermore, we did not consider the agreements working on public holidays, including if you work on both Christmas days, you are free with New Year's Eve, being free on either Easter or Pentecost weekend, and free on Ascension Day or King's Day. Also, we did not include the specified rules of working at most 38 hours between 00:00 and 06:00 AM in a consecutive period of two weeks. At last, planned holidays are processed into the operational schedule. The collective labor agreements stipulate that at least once a year, the weekend before or after a holiday of at least a week should be assigned free. This is not considered in the proposed method.

Most importantly, to evaluate the performance of the method, we recommend Nedap to validate the constructed schedules with planners and nurses from practice. Based on the validation, the method can be fine-tuned, and appropriate weights can be assigned to each constraint. It must be noted that the preferences differ per organization. Therefore, it is important to validate the method for multiple organizations.

7.4 Future research

Based on the results, it can be concluded that the proposed method can be used to support planners from residential care organizations in constructing fair tactical and operational shift schedules for their nurses. However, due to time restrictions, the lack of data, and simplifications and assumptions made, not all requirements and agreements are considered. This provides opportunities for future research.

For simplifications, this research discusses the static version of the NRP, where the capacity requirements and available nurses are known in advance. For a more realistic representation of real-world scenarios, future research can implement a dynamic version. As in practice, nurses retire or resign, and the demand can change over time. This is to investigate the effect of changing demand and capacity on the outcomes of the mode.

Furthermore, as discussed in Section 7.2.1, various weights have been implemented in the method. As shown by the sensitivity analysis, these all affect the outcomes of the schedule as they are correlated. Therefore, future research should be conducted to determine the appropriate weights for the tactical and operational schedules implemented in the model. Additionally, they should look into if these differ per organization or if an optimal selection of weights can be suggested. By validating the schedules with planners and nurses from practice, the algorithm can be adapted. Keeping track of the adjustments that would be made to the proposed schedules would allow further research to use other algorithms, such as machine learning, to analyze the data or to find patterns in order to find the appropriate weights. Also, the individual perception of fairness should be considered to assess individual fairness within the scheduling process and adjust the weights based on the preferences of individual nurses. Extending the method such that individual contract agreements can be considered, a first step in considering individual preferences and individual fairness can be achieved. When the optimal weights have been determined, the parameter tuning procedure should be conducted again for both the tactical and operational schedules in future research. Additionally, as mentioned in Section 7.2.1, we have assumed that there is no previous or upcoming period. Future research can investigate the effect by also taking into account previous schedules. When also additional requests are considered, long-term fairness can be achieved. Furthermore, this will also improve the tactical schedule as the last and first-week impact each other on which shifts are allowed to be assigned. In addition, as mentioned in Section 7.2.1, shift patterns can be explored to reduce the number of TRC violations and have more stable schedules. As well, future research can be conducted to make a distinction between shifts types that are included in the tactical schedule or only considered in the operational schedule, e.g., only assigning the NS and or WS in the tactical schedule.

Future experiments can be conducted using different initial schedules within the SA algorithm. In this research, a constructive heuristic is used to find an initial feasible solution for the tactical schedule, which can influence the end results as it defines the search space for the operators. Because historical tactical schedules are used to validate the performance of the method, these could also have been used as input in the SA algorithm. Insight would be gained into the performance and impact of the proposed constructive heuristic on the final schedule. If similar or better results than the current performance are obtained when using the constructive heuristic, we suggest future research focus on adapting the constructive heuristic or testing the proposed method using more historical schedules. Secondly, in our study, we utilize the tactical schedule as input for the SA algorithm. It is important to investigate the effects of using the tactical schedule as input for the SA algorithm in the operational scheduling process. Whereas our research focussed on organizations that use a tactical schedule, it is worthwhile to explore alternative approaches where the operational schedule is constructed independently for each period without using a tactical schedule. This also supports care organizations not interested in using a tactical schedule as a scheduling method.

Bibliography

[1] CBS. Ouderen. URL https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/leeftijd/ouderen. Accessed on 10-10-2022.

[2] Rijksinstituut voor Volksgezondheid en Milieu. Infographic impact van de vergrijzing. URL https://www.rivm.nl/infographic-impact-van-vergrijzing. Accessed on 15-11-2022.

[3] Ourega-Zoé Ejebu, Chiara Dall'Ora, and Peter Griffiths. Nurses' experiences and preferences around shift patterns: A scoping review. *PLOS ONE*, 16:e0256300, 8 2021. ISSN 1932-6203. doi: 10.1371/journal.pone.0256300. URL https://dx.plos.org/10.1371/journal.pone.0256300.

[4] NOS Nieuws. Onderzoek: tekort aan zorgpersoneel op lange termijn alleen maar groter. URL https://nos.nl/artikel/2413851-onderzoek-tekort-aan-zorgpersoneel-op-lange-termijn-alleen-maar-groter. Accessed on 21-10-2022.

[5] Sandy Heydrich, Rasmus Schroeder, and Sebastian Velten. Collaborative duty rostering in health care professions. *Operations Research for Health Care*, 27:100278, 12 2020. ISSN 22116923. doi: 10.1016/j.orhc.2020.100278. URL https://linkinghub.elsevier.com/retrieve/pii/S2211692320300588.

[6] Sociaal economische Raad. Verkenning 20/02 — juni 2020 zorg voor de toekomst - over de toekomstbestendigheid van de zorg. 6 2020. URL https://www.ser.nl/-/media/ser/downloads/adviezen/2020/zorg-voor-de-toekomst.pdf.

[7] B Maenhout and M Vanhoucke. The impact of incorporating nurse-specific characteristics in a cyclical scheduling approach. *Journal of the Operational Research Society*, 60:1683–1698, 12 2009. ISSN 0160-5682. doi: 10.1057/jors.2008.131. URL https://www.tandfonline.com/doi/full/10.1057/jors.2008.131.

[8] Chong Man Ngoo, Say Leng Goh, San Nah Sze, Nasser R. Sabar, Salwani Abdullah, and Graham Kendall. A survey of the nurse rostering solution methodologies: The state-of-the-art

and emerging trends. *IEEE Access*, 10:56504–56524, 2022. ISSN 2169-3536. doi: 10.1109/ACCESS.2022.3177280. URL `https://ieeexplore.ieee.org/document/9780256/`.

[9] P. Daniel Wright and Stephen Mahar. Centralized nurse scheduling to simultaneously improve schedule cost and nurse satisfaction. *Omega*, 41:1042–1052, 12 2013. ISSN 03050483. doi: 10.1016/j.omega.2012.08.004. URL `https://linkinghub.elsevier.com/retrieve/pii/S0305048312001119`.

[10] Rebecca M. Brossoit, Tori L. Crain, Leslie B. Hammer, Soomi Lee, Todd E. Bodner, and Orfeu M. Buxton. Associations among patient care workers' schedule control, sleep, job satisfaction and turnover intentions. *Stress and Health*, 36:442–456, 10 2020. ISSN 1532-3005. doi: 10.1002/smi.2941.

[11] Catherine Laura Hainey. Implementation of a ward staff self-rostering system: improving morale and retention. *Nursing Management*, 28:22–27, 6 2021. ISSN 1354-5760. doi: 10.7748/nm.2021.e1987. URL `https://journals.rcni.com/doi/10.7748/nm.2021.e1987`.

[12] Elín Björk Bövarsdóttir, Niels-Christian Fink Bagger, Laura Elise Høffner, and Thomas J. R. Stidsen. A flexible mixed integer programming-based system for real-world nurse rostering. *Journal of Scheduling*, 25:59–88, 2 2022. ISSN 1094-6136. doi: 10.1007/s10951-021-00705-7. URL `https://link.springer.com/10.1007/s10951-021-00705-7`.

[13] E. van der Veen, J. L. Hurink, J. M. J. Schutten, and S. T. Uijland. A flexible iterative improvement heuristic to support creation of feasible shift rosters in self-rostering. *Annals of Operations Research*, 239:189–206, 4 2016. ISSN 0254-5330. doi: 10.1007/s10479-014-1540-7. URL `http://link.springer.com/10.1007/s10479-014-1540-7`.

[14] Franklin Leung, Yee-Chun Lau, Martin Law, and Shih-Kien Djeng. Artificial intelligence and end user tools to develop a nurse duty roster scheduling system. *International Journal of Nursing Sciences*, 9(3):373–377, 2022. ISSN 2352-0132. doi: https://doi.org/10.1016/j.ijnss.2022.06.013. URL `https://www.sciencedirect.com/science/article/pii/S235201322200045X`.

[15] Jorne Van den Bergh, Jeroen Beliën, Philippe De Bruecker, Erik Demeulemeester, and Liesje De Boeck. Personnel scheduling: A literature review. *European Journal of Operational Research*, 226:367–385, 5 2013. ISSN 03772217. doi: 10.1016/j.ejor.2012.11.029. URL `https://linkinghub.elsevier.com/retrieve/pii/S0377221712008776`.

[16] Edmund K. Burke, Patrick De Causmaecker, Greet Vanden Berghe, and Hendrik Van Landeghem. The state of the art of nurse rostering. *Journal of Scheduling*, 7:441–499, 11 2004. ISSN 1094-6136. doi: 10.1023/B:JOSH.0000046076.75950.0b. URL `http://link.springer.com/10.1023/B:JOSH.0000046076.75950.0b`.

[17] Lena Wolbeck and Natalia Kliewer. Maximising employee satisfaction with individual preferences: Fairness enhanced rostering. 2016.

[18] Alarith Uhde, Nadine Schlicker, Dieter P. Wallach, and Marc Hassenzahl. Fairness and decision-making in collaborative shift scheduling systems. pages 1–13. ACM, 4 2020. ISBN 9781450367080. doi: 10.1145/3313831.3376656. URL `https://dl.acm.org/doi/10.1145/3313831.3376656`.

[19] Millicent F Nelson and Richard J Tarpey. Work scheduling satisfaction and work life balance for nurses: The perception of organizational justice. *Academy of Health Care Management Journal*, 6, 2010.

[20] Ahmed Osman, Nasr Al-Hinai, and Sujan Piya. Development of automated schedule generator for nurses in emergency department. pages 1–3. IEEE, 4 2019. ISBN 978-1-5386-7684-4. doi: 10.1109/ICMSAO.2019.8880428. URL `https://ieeexplore.ieee.org/document/8880428/`.

[21] Nanne A. Dieleman, Martijn Buitink, René Bekker, and Dennis Moeke. A three-step framework for capacity planning in a nursing home context. *Health Systems*, pages 1–18, 5 2022. ISSN 2047-6965. doi: 10.1080/20476965.2022.2062461. URL `https://www.tandfonline.com/doi/full/10.1080/20476965.2022.2062461`.

[22] Y. P. Tsang, C. H. Wu, Polly P. L. Leung, W. H. Ip, and W. K. Ching. Blockchain-iot-driven nursing workforce planning for effective long-term care management in nursing homes. *Journal of Healthcare Engineering*, 2021:1–19, 11 2021. ISSN 2040-2309. doi: 10.1155/2021/9974059. URL `https://www.hindawi.com/journals/jhe/2021/9974059/`.

[23] Peter J H Hulshof, Nikky Kortbeek, Richard J Boucherie, Erwin W Hans, and Piet J M Bakker. Taxonomic classification of planning decisions in health care: a structured review of the state of the art in or/ms. *Health Systems*, 1(2):129–175, 2012. doi: 10.1057/hs.2012.18.

[24] Erwin W. Hans, Mark van Houdenhoven, and Peter J. H. Hulshof. A framework for healthcare planning and control. pages 303–320, 2012. doi: 10.1007/978-1-4614-1734-7_12. URL `http://link.springer.com/10.1007/978-1-4614-1734-7_12`.

[25] Zahraa A. Abdalkareem, Amiza Amir, Mohammed Azmi Al-Betar, Phaklen Ekhan, and Abdelaziz I. Hammouri. Healthcare scheduling in optimization context: a review. *Health and Technology*, 11:445–469, 5 2021. ISSN 21907196. doi: 10.1007/s12553-021-00547-5.

[26] Elina Rönnberg and Torbjörn Larsson. Automating the self-scheduling process of nurses in swedish healthcare: a pilot study. *Health Care Management Science*, 13:35–53, 3 2010. ISSN 1386-9620. doi: 10.1007/s10729-009-9107-x. URL `http://link.springer.com/10.1007/s10729-009-9107-x`.

[27] Renee Barrett and Annie Holme. Self-rostering can improve work-life balance and staff retention in the nhs. *British Journal of Nursing*, 27:264–265, 3 2018. ISSN 09660461. doi: 10.12968/bjon.2018.27.5.264.

[28] Lena Wolbeck. Fairness aspects in personnel scheduling. 2019. doi: 10.17169/refubium-26050.

[29] Ferdinand Kiermaier, Markus Frey, and Jonathan F. Bard. Flexible cyclic rostering in the service industry. *IIE Transactions*, 48:1139–1155, 12 2016. ISSN 0740-817X. doi: 10.1080/0740817X.2016.1200202. URL `https://www.tandfonline.com/doi/full/10.1080/0740817X.2016.1200202`.

[30] Edmund K. Burke, Patrick De Causmaecker, Sanja Petrovic, and Greet Vanden Berghe. Metaheuristics for handling time interval coverage constraints in nurse scheduling. *Applied Artificial Intelligence*, 20:743–766, 12 2006. ISSN 0883-9514. doi: 10.1080/08839510600903841. URL `https://www.tandfonline.com/doi/full/10.1080/08839510600903841`.

[31] Jhunievieve Soriano, Eugene Rex Jalao, and Iris Ann Martinez. Integrated employee scheduling with known employee demand, including breaks, overtime, and employee preferences. *Journal of Industrial Engineering and Management*, 13:451, 9 2020. ISSN 2013-0953. doi: 10.3926/jiem.3126. URL `http://www.jiem.org/index.php/jiem/article/view/3126`.

[32] ActiZ. Cao verpleeg, verzorgingshuizen, thuizorg en jeugdgezondheidszorg 2022-2022. 12 2022.

[33] Werkgroep Domus Magnus. Roosterbeleid domus magnus. 4 2022.

[34] Regiomanager. Beleid plannen & roosteren stmg. 2 2022.

[35] Attent Zorg en Behandeling. Kaderregeling roosteren en werktijden attent zorg en behandeling. 2 2019.

[36] Carinova. Carinova slim plannen & roosteren. 2022.

[37] TriviumMeulenbeltZorg. Roosterhandboek tmz. 11 2022.

[38] Manager Flexbureau. Rooster en flexbeleid warande. 1 2023.

[39] Mark Horsman. Roosterhandboek. 7 2022.

[40] Ministerie van Sociale Zaken en Werkgelegenheid. De arbeidstijdenwet. 4 2011.

[41] Jonathan F. Bard and Hadi W. Purnomo. Preference scheduling for nurses using column generation. *European Journal of Operational Research*, 164(2):510–534, 7 2005. URL `https://ideas.repec.org/a/eee/ejores/v164y2005i2p510-534.html`.

[42] D Michael Warner. Scheduling nursing personnel according to nursing preference: A mathematical programming approach. *Research*, 24:842–856. URL https://www.jstor.org/stable/169810.

[43] Djamila Ouelhadj, Simon Martin, P. Smet, E. Ozcan, and G. Vanden Berghe. Fairness in nurse rostering. January 2012. Institution: University of Portsmouth. Department: Mathematics.

[44] Huaizhou Shi, R. Venkatesha Prasad, Ertan Onur, and I.G.M.M. Niemegeers. Fairness in wireless networks: Issues, measures and challenges. *IEEE Communications Surveys Tutorials*, 16(1):5–24, 2014. doi: 10.1109/SURV.2013.050113.00015.

[45] E. Fehr and K. M. Schmidt. A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114:817–868, 8 1999. ISSN 0033-5533. doi: 10.1162/003355399556151.

[46] R. Jain, D. Chiu, and W. Hawe. A quantitative measure of fairness and discrimination for resource allocation in shared computer systems, 1998.

[47] Rostislav Burget and Hana Rudová. Teacher-oriented fairness in course timetabling. In Ender Özcan Barry McCollum Andrea Schaerf Edmund K. Burke, Luca Di Gaspero, editor, *Proceedings of the 11th International Conference on the Practice and Theory of Automated Timetabling (PATAT-2016)*, pages 33–44, Udine, Italy, 2016. PATAT. ISBN 978-0-9929984-1-7.

[48] Mohammed Hadwan and Masri Ayob. A constructive shift patterns approach with simulated annealing for nurse rostering problem. *Proceedings 2010 International Symposium on Information Technology - Visual Informatics, ITSim'10*, 1, 2010. doi: 10.1109/ITSIM.2010.5561304.

[49] Mohammed Hadwan and Masri Ayob. A semi-cyclic shift patterns approach for nurse rostering problems. *Conference on Data Mining and Optimization*, pages 184–189, 2011. ISSN 21556938. doi: 10.1109/DMO.2011.5976525.

[50] Simon Martin, Djamila Ouelhadj, Pieter Smet, Greet Vanden Berghe, and Ender¨ozcan Ender¨ Ender¨ozcan. Cooperative search for fair nurse rosters. 2013.

[51] Anna Lavygina, Kristopher Welsh, and Alan Crispin. On fairness as a rostering objective. *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pages 217–218, 7 2019. doi: 10.1145/3319619.3321913. URL https://dl.acm.org/doi/10.1145/3319619.3321913.

[52] Anna Lavygina, Kris Welsh, and Alan Crispin. *Doctor Rostering in Compliance with the New UK Junior Doctor Contract*, pages 394–408. 01 2017. ISBN 978-3-319-71146-1. doi: 10.1007/978-3-319-71147-8_27.

[53] Pieter Smet, K U Leuven, Simon Martin, and Djamila Ouelhadj. Investigation of fairness measures for nurse rostering. 2012. URL `https://www.researchgate.net/publication/255568989`.

[54] Ademir A. Constantino, Everton Luiz de Melo, Dario Landa-Silva, and Wesley Romao. A heuristic algorithm for nurse scheduling with balanced preference satisfaction. *2011 IEEE Symposium on Computational Intelligence in Scheduling (SCIS)*, pages 39–45, 4 2011. doi: 10.1109/SCIS.2011.5976549. URL `http://ieeexplore.ieee.org/document/5976549/`.

[55] Chang Chun Tsaia and Cheng Jung Leeb. Optimization of nurse scheduling problem with a two-stage mathematical programming model. *Asia Pacific Management Review*, 15, 2010. ISSN 10293132.

[56] Takayuki Osogami and Hiroshi Imai. *Classification of Various Neighborhood Operations for the Nurse Scheduling Problem*, pages 72–83. 2000. doi: 10.1007/3-540-40996-3_7.

[57] Hamed Jafari and Nasser Salmasi. Maximizing the nurses' preferences in nurse scheduling problem: mathematical modeling and a meta-heuristic algorithm. *Journal of Industrial Engineering International*, 11:439–458, 9 2015. ISSN 1735-5702. doi: 10.1007/s40092-015-0111-0. URL `http://link.springer.com/10.1007/s40092-015-0111-0`.

[58] Aykut Melih Turhan and Bilge Bilgen. A hybrid fix-and-optimize and simulated annealing approaches for nurse rostering problem. *Computers and Industrial Engineering*, 145, 7 2020. ISSN 03608352. doi: 10.1016/j.cie.2020.106531.

[59] Dingding Lin, Yue Tong, Ganggang Niu, Yongqing Xue, Xin Shi, Changrui Ren, and Zongying Zhang. Scheduling workforce for retail stores with employee preferences. pages 37–42. IEEE, 11 2015. ISBN 978-1-4673-8480-3. doi: 10.1109/SOLI.2015.7367407. URL `http://ieeexplore.ieee.org/document/7367407/`.

[60] Sara Ceschia, · Nguyen, Thi Thanh Dang, Patrick De Causmaecker, Stefaan Haspeslagh, Andrea Schaerf, S Ceschia, · A Schaerf, N Thi, Thanh Dang, · P De Causmaecker, K U Leuven, and S Haspeslagh. Second international nurse rostering competition: mathematical model, 2018.

[61] Djamila Ouelhadj, Simon Martin, Pieter Smet, Ender Ozcan, and Vanden Berghe. Fairness in nurse rostering.

[62] E.K. Burke, P. De Causmaecker, S. Petrovic, and G.V. Berghe. Fitness evaluation for nurse scheduling problems. volume 2, pages 1139–1146. IEEE. ISBN 0-7803-6657-3. doi: 10.1109/CEC.2001.934319. URL `http://ieeexplore.ieee.org/document/934319/`.

[63] Lucas Kletzander and Nysret Musliu. Solving the general employee scheduling problem. *Computers and Operations Research*, 113, 1 2020. ISSN 03050548. doi: 10.1016/j.cor.2019.104794.

[64] Christopher A Oyeleye, Grace O Oladele, Oluwaseun M Alade, Oniyide A Bello, Adeyemi I Adeyemo, and Titilayo O Adedeji. Modified genetic algorithm for solving nurse scheduling problem. *International Research Journal of Computer Science (IRJCS)*, 07:33–41, 2020. doi: 10.26562/IRJCS.2020.APCS10081. URL `www.irjcs.com`.

[65] Patrick De Causmaecker and Greet Vanden Berghe. Relaxation of coverage constraints in hospital personnel rostering. In Edmund Burke and Patrick De Causmaecker, editors, *Practice and Theory of Automated Timetabling IV*, pages 129–147, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg. ISBN 978-3-540-45157-0.

[66] Daniela Guericke. Routing and scheduling for home care services solution approaches for static and dynamic settings in steinfurt. 2016.

[67] Scott Kirkpatrick, C. Gelatt, and M. Vecchi. Optimization by simulated annealing. *Science (New York, N.Y.)*, 220:671–80, 06 1983. doi: 10.1126/science.220.4598.671.

[68] Darrall Henderson, Sheldon H. Jacobson, and Alan W. Johnson. *The Theory and Practice of Simulated Annealing*, pages 287–319. Kluwer Academic Publishers. doi: 10.1007/0-306-48056-5_10.

[69] Zbigniew Michalewicz and David B. Fogel. *Traditional Methods — Part 2*, page 120. Springer Berlin Heidelberg, 2004. doi: 10.1007/978-3-662-07807-5_5.

[70] Stefan Ropke and David Pisinger. An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. *Transportation Science*, 40:455–472, 11 2006. ISSN 0041-1655. doi: 10.1287/trsc.1050.0135.

[71] P. Ngatchou, A. Zarei, and A. El-Sharkawi. Pareto multi objective optimization. pages 84–91. IEEE, 2005. ISBN 1-59975-174-7. doi: 10.1109/ISAP.2005.1599245. URL `http://ieeexplore.ieee.org/document/1599245/`.

[72] Christopher A. Mattson and Achille Messac. Pareto frontier based concept selection under uncertainty, with visualization. *Optimization and Engineering*, 6:85–115, 3 2005. ISSN 1389-4420. doi: 10.1023/B:OPTE.0000048538.35456.45.

Appendix A

Calculation Soft Constraints

A.1 Calculations time-related soft constraints

We provide the description on how the penalties for the TRCs are calculated. This is based on the approach used in Lavygina et al. [52]. The missing rest time is penalized in SC1-SC4. SC5 is determined by penalizing the violating weekend shifts and SC7 by determining the time between the two shifts is the forbidden pattern occurs. At last, SC8 is determined by calculating the ratio of remaining hours compared to the agreed contract hours. SC1-SC7 are in seconds and converted to minutes, and SC8 is a ratio in a range of [0,1].

SC1: Rest time between shifts.

The rest time between shifts is calculated by sorting the shifts assigned to each nurse based on the start date. Hereafter, the rest time between every two consecutive shifts is calculated. If this is less than the required rest time of 11 hours [40], we calculate the missing rest hours by subtracting the required rest time from the assigned rest time. If this result in a negative value, meaning that the start time of the next shift starts before the end time of the first shift, the total missing rest hours is equal to the required rest hours. At last, for each nurse, the total missing rest hours are calculated in seconds and multiplied by the number of violations. This is in order to give weight to the number of violations instead of only the missing rest hours. The pseudocode can be found in Appendix **??**

SC2: Forward rotating order.

Shifts that are assigned consecutively should follow a forward rotating order. Again, the shifts for each nurse are sorted based on the start date. The start time of the current and consecutive shifts are determined. The difference between the start dates of the two consecutive shifts is calculated to determine if the shifts occur on consecutive days. If so, we determine if the start time of the next shift starts before the start time of the current shift. When this is true, it means that the forward rotating order constraint is violated and results in a penalty of the hours the upcomming shift has started too early. The total penalty is determined by the violation in seconds times the number of violations. An example is provided in Figure A.1. We have implemented a slack variable, as we allow the next shift to start one hour earlier than the current shift depending on the organisation. This is repeated until all assigned nurse shifts are checked, resulting in a total number of violated seconds. The pseudocode can be found in Appendix **??**.

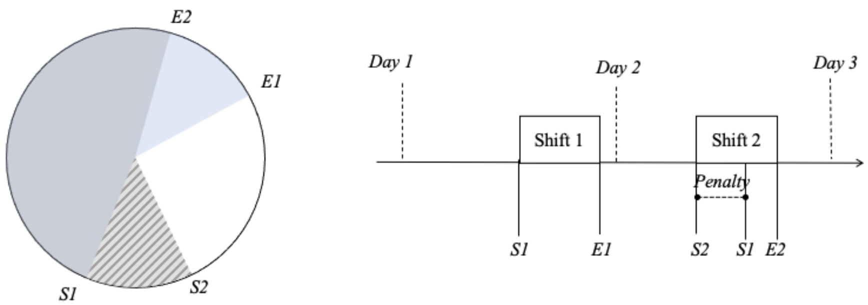


Figure A.1: Example of forward rotating order. The striped area represents the penalty resulting from the difference between the start time of shift 2 and shift 1.

SC3 and SC4: Consecutive working days and night shifts.

According to law legislation, an employee should get 36 hours of rest time after 5 consecutive working days and 48 hours of rest time after 3 or more consecutive night shifts. To determine the penalty for missing rest hours for the maximum amount of consecutive working days or night shifts, the assigned shifts for the nurse are again sorted on the start date. We determine the number of consecutive working days in the assigned shifts. These shifts are appended to a list, which is used to calculate the penalty. First, the number of consecutive shifts is determined. If this exceeds the maximum amount, the missing hours of rest time between each violated shift are determined. So for example, for the number of consecutive working days, if 8 consecutive shifts are assigned, as visualized in figure A.2, there are 3 more shifts assigned than allowed. For these 3 violated shifts, the missing rest time compared to the required rest time is calculated and added to the total missing rest time. The pseudocode can be found in Appendix **??** and **??**.

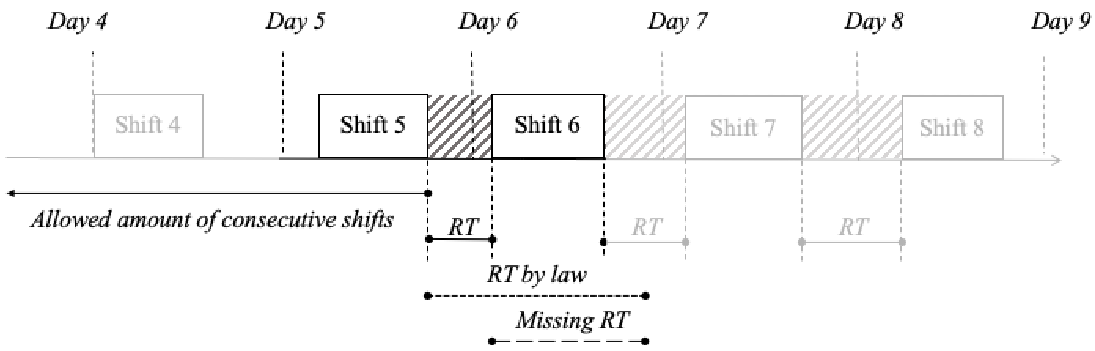


Figure A.2: Example of violation of consecutive working days. The stripes represent the total missing rest hours. The penalty for the missing rest time is determined by subtracting the assigned rest time from the required rest time by law.

SC5 and SC6: Weekend shifts.

Nurses are limited to working two weekends in a four-week period and should work every other weekend according to law legislation. The penalty for violating these rules is calculated based on the number of hours worked during the violated weekends. This results in a penalty *Max2Weekends* and *EveryOtherWeekend*, which follow the same procedure. The assigned weekend shifts are appended in ascending order to a list to determine the total number of weekend shifts and the number of working weekends. To determine if a nurse works every other weekend, the difference between the days of the shifts is calculated. The value of the penalty depends on the number of violated shifts in *Max2Weekends* and the occurrence of violating *EveryOtherWeekend*. These are multiplied by the shift duration to get a total penalty. Figure A.3 provides examples of weekend shift assignments and the corresponding penalties in number of shifts. The pseudocode weekend shifts can be found in Appendix **??**.

SC7: Forbidden patterns.

To ensure that there is consistency in the nurses' schedule, we try to prevent a nurse has the following shift assignment pattern: on-off-on. We determine the penalty as follows. For the sorted shifts we determine the difference between the consecutive shifts. If this is equal to two, it means that there is a day off assigned between the two shifts. The penalty is then determined by subtracting the start time of the next shift from the end time of the current shift. In order to prevent that small violation that occur more often are preferred above less but large violations, we multiply the violation in seconds by the number of violations. The pseudocode can be found in Appendix **??**.

Week 1		Week 2		Week 3		Week 4	
Sa	Su	Sa	Su	Sa	Su	Sa	Su
		X	X			X	X

(A) TW = 2, CW = 0, TS = 4, PTW = 0S, PCW = 0

Week 1		Week 2		Week 3		Week 4	
Sa	Su	Sa	Su	Sa	Su	Sa	Su
X	X	X	X	X	X	X	X

(B) TW = 4, CW = 3, TS = 8, PTW = 4S, PCW = 3S

Week 1		Week 2		Week 3		Week 4	
Sa	Su	Sa	Su	Sa	Su	Sa	Su
X		X	X			X	X

(C) TW = 3, CW = 1, TS = 5, PTW = 1S, PCW = 1S

Week 1		Week 2		Week 3		Week 4	
Sa	Su	Sa	Su	Sa	Su	Sa	Su
X		X	X	X		X	X

(D) TW = 4, CW = 3, TS = 6, PTW = 2S, PCW = 3S

Figure A.3: Example of weekend shift assignments. X: assigned shift; TW: total weekends, CW: consecutive weekends, PTW: penalty total weekends in number of shifts, PCW: penalty consecutive weekends in number of shifts

SC8: Remaining contract hours.

Each nurse has an agreed amount of contractual hours that he or she needs to work in the planning horizon. To guarantee that nurses work according to their contracts, we try to minimize the unassigned hours for each nurse. By determining the ratio of unassigned hours in relation to the agreed contract hours, we try to get a fair distribution of assigned hours. This will give us a ratio between 0 and 1. When the ratio is 1 none of the agreed contract hours is assigned and visa versa. To balance the workload and have a fair distribution of assigned contract hours we try to balance the ratio between the nurses. By using the Log2 function and raising it to power 10, higher penalties are assigned to larger discrepancies between the remaining and agreed hours. For the periodic schedule, we also penalize min hours to prevent the workload becomes much higher for one nurse.

For example, if a nurse has a contract of 13,600 minutes, and another nurse of 1,440 minutes. They have got the following minutes assigned, 10,000 and 1,000 minutes respectively. The ratio of missing hours compared to the agreed contract hours is then 0.26 and 0.30 respectively.

In comparison, the nurse with the large contract has assigned more of the agreed contract hours than the nurse with the smaller contract. The pseudocode can be found in Appendix **??**

A.2 Calculations organisational related soft constraints

We explain how the penalties for the ORCs are calculated. We determine the total undercoverage and UQ shifts in minutes.

SC10: Coverage constraints.

The goal of a care organisation is to deliver the right care at the right time. As it is not evident that the staffing capacity is fitting, we try to minimize the amount of uncovered shift hours per day. Within the algorithm, the missing hours per day are determined which are then summed to determine the missing hours during the whole planning horizon. Depending on the care organisation, a percentage of the shifts are reserved in advance for flex workers. Therefore, we allow a coverage constraint violation which is equal to the hours reserved for flex. The penalty will then be equal to 0. Otherwise, the penalty will be the positive difference between the missing hours and the hours covered by flex is the penalty. The pseudocode can be found in **??**.

SC11: QL3 during the day.

During the whole day, a nurse with QL3 should be present. However, again it is not evident that the staffing capacity is fitting. Therefore, we minimize the missing hours of level 3 coverage as a soft constraint. If this constraint were a hard constraint it would restrict the operators in the SA algorithm, as unassigning a QL3 shift will result in an infeasible solution and restrict the search space of the neighbourhoods.

A visualization of two QL3 coverage examples are provided in Figure A.4. In the procedure, the shifts are sorted by starting time, and the start and end times are determined. if the start time of the first shift is before or equal to the start time of the day, the end time of the shift becomes the new start time, as shown in Figure A.4B. Otherwise, if the shift starts after the start time, the missing QL3 hours for the beginning of the day are calculated. Which is the difference between the start of the shift and the start time of the day, as can be seen in Figure A.4A. The missing hours for the end of the day are set equal to the difference between the end time of the day and the end of the shift. If the second shift overlaps with the first shift, as in Figure A.4B, the end time of the second shift becomes the new start time. The penalty for the end of the day is recalculated as the gap decreases. However, if the start time of the second shift lies beyond the end time of the first shift, we increase the missing QL3 hours of the beginning of the day with the difference. As the shifts are sorted by their starting time, no shift can cover this gap of the beginning of the day. This is repeated for all shifts on each day. In Appendix **??** the pseudocode can be found.

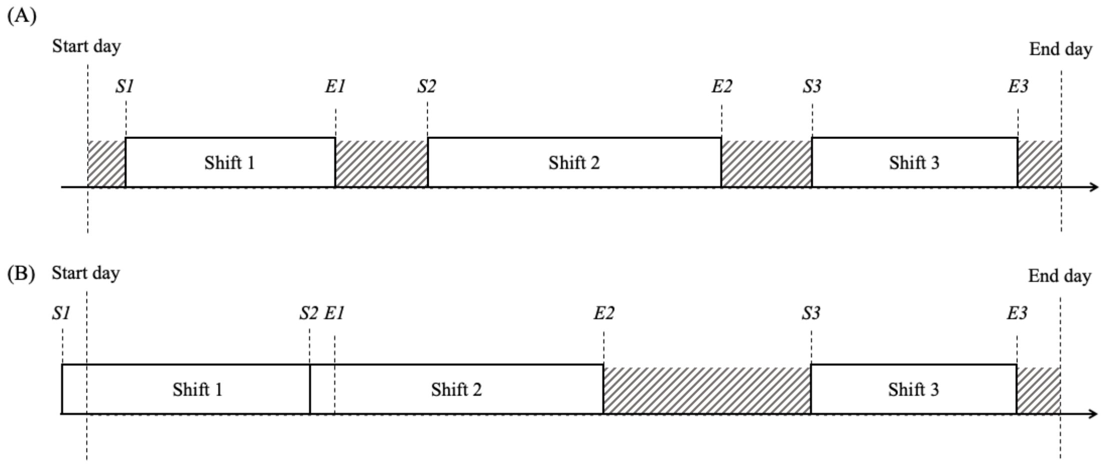


Figure A.4: Example calculation penalty 24-hour coverage qualification level 3. The stripes represent the missing hours of qualification level 3.

SC12: Underqualification.

To deliver the right care at the correct time, qualification levels should be taken into consideration. As stated before, it is not evident that the staffing capacity is fitting. Also, in practice, if the capacity is not sufficient to meet demand, underqualified nurses nurses assigned to shifts to maximize the number of covered shifts. However, we try to prevent this by minimizing the hours worked by an underqualified nurse per day. The pseudocode can be found in **??**.

Appendix B

Algorithms

In this section, a detailed explanation is provided of the calculation process for the penalties associated with TRCs and ORCs.

Algorithm 3: Check if a shift can be assigned to a nurse

```
Input  : Shift, remaining minutes for each nurse
Output: Can be assigned or not
Sort the nurses based on the remaining minutes to assign;
Assigned ← False;
while Assigned is False and not all nurses have been checked do
    foreach Nurse in sorted nurses do
        Check if the nurse has the correct age for the shift;
        Check if the nurse has the correct qualification level;
        Check if the nurse has minutes left to assign;
        Check if the nurse is available;
        Check if the nurse has already a shift on this day;
        if All checks are true then
            Assigned ← True;
        else
            Go to the next nurse in the list
        end
    end
end
return Assigned
```

Appendix C

Parameter tuning SA

C.1 Results Parameter Tuning

As mentioned, the starting temperature, T_{start}, is based on the objective value of the initial solution to provide an instance-based value. In this research, T_{start} is chosen such that a solution that is 1.8 worse than the initial solution is accepted with a probability of 0.5.

Table C.1: Average outcomes of the objective value over five runs for the three case studies using different parameter values.

Case study	T_{start}	α	T_0	Objective value		
				MCL = 1	MCL = 100	MCL = 1000
Small	0.1169	0.8	0.01	0.0981	0.0721	0.0629
			0.001	0.0956	0.0237	0.0663
			0.0001	0.0954	**0.0064**	0.0632
		0.9	0.01	0.0991	0.0555	0.0616
			0.001	0.0903	0.0381	0.0708
			0.0001	0.0885	0.0397	0.0727
		0.99	0.01	0.0883	0.0734	0.0693
			0.001	0.0721	0.0718	0.0679
			0.0001	0.0410	0.0684	0.0688
Medium	0.3199	0.8	0.01	0.2044	0.1759	0.1659
			0.001	0.2477	0.1302	0.1935
			0.0001	0.2348	0.1361	0.1812
		0.9	0.01	0.2315	0.1931	0.1931
			0.001	0.2271	0.1897	0.1920
			0.0001	0.1924	0.1994	0.2043
		0.99	0.01	0.1793	0.2036	0.1692
			0.001	0.1508	0.1979	0.1964
			0.0001	**0.0988**	0.1934	0.1943
Large	0.6765	0.8	0.01	0.3981	0.3424	0.3506
			0.001	0.3924	0.3345	0.3640
			0.0001	0.3932	0.3492	0.3515
		0.9	0.01	0.3934	0.3477	0.3563
			0.001	0.3909	0.3538	0.3610
			0.0001	0.3748	0.3590	0.3628
		0.99	0.01	0.3595	0.3512	0.3775
			0.001	0.3564	0.3661	0.3673
			0.0001	**0.2674**	0.3225	0.3402

C.1.1 Outcomes parameter tuning Small case study

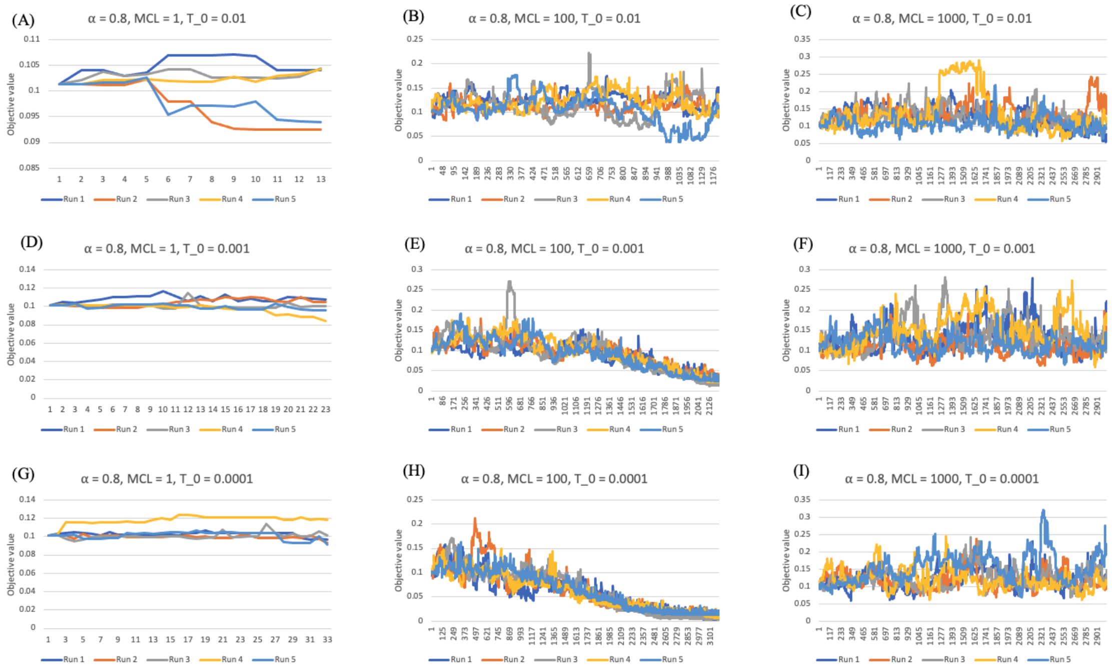


Figure C.1: The outcomes for parameter tuning of the SA algorithm for the small case study with α of 0.8

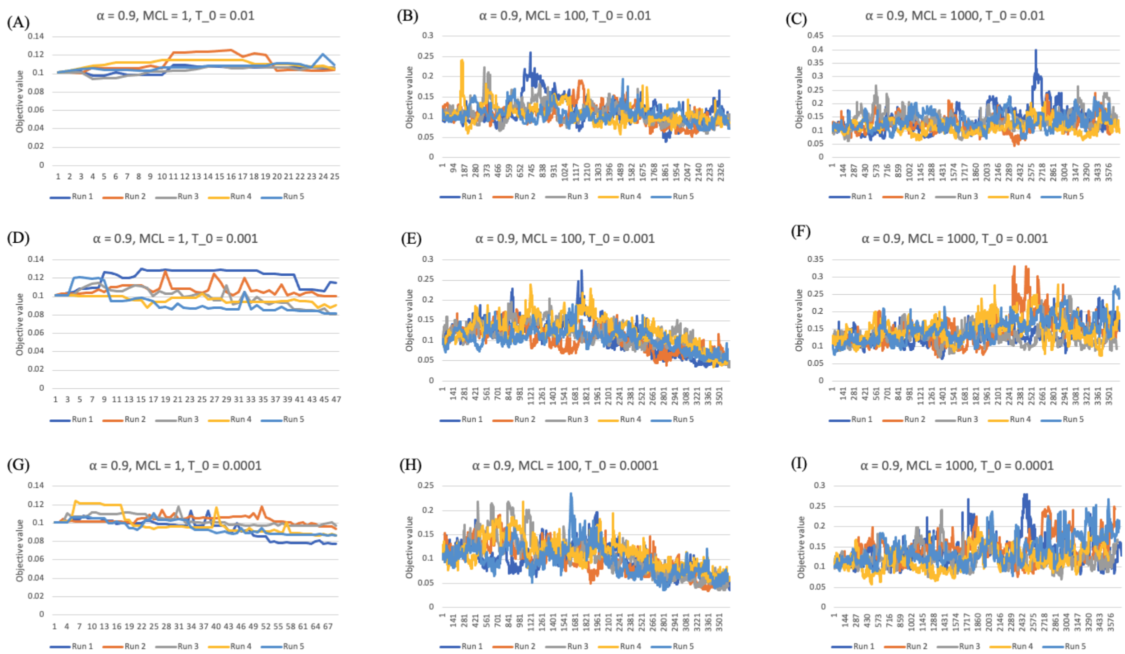


Figure C.2: The outcomes for parameter tuning of the SA algorithm for the small case study with α of 0.9

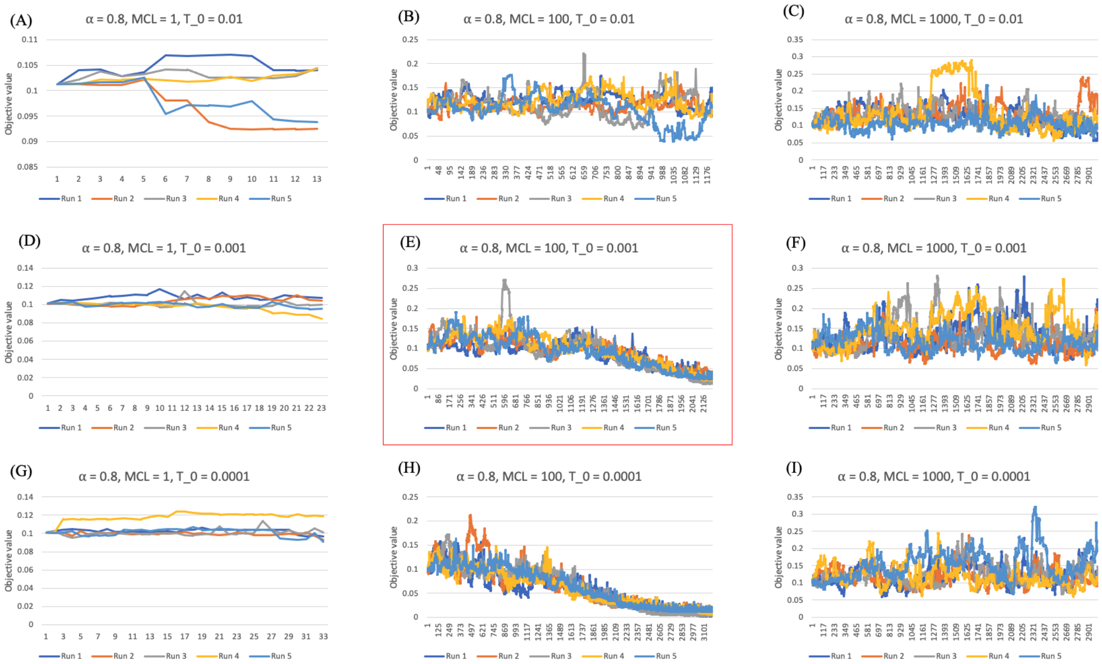


Figure C.3: The outcomes for parameter tuning of the SA algorithm for the small case study with α of 0.99

C.1.2 Outcomes parameter tuning Medium case study

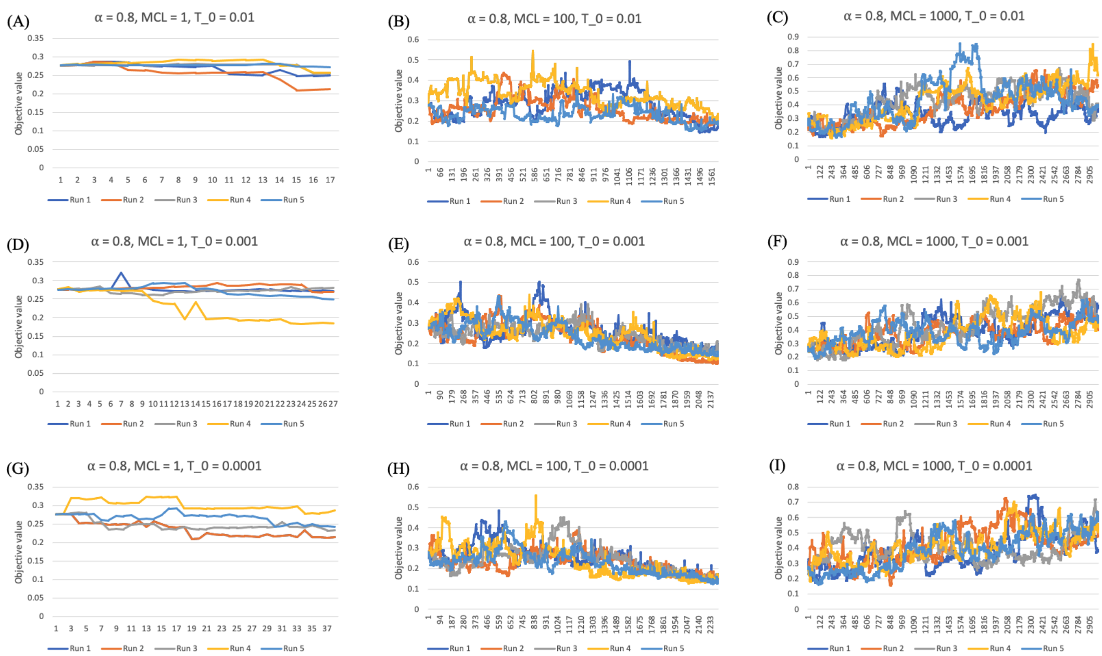


Figure C.4: The outcomes for parameter tuning of the SA algorithm for the medium case study with α of 0.8

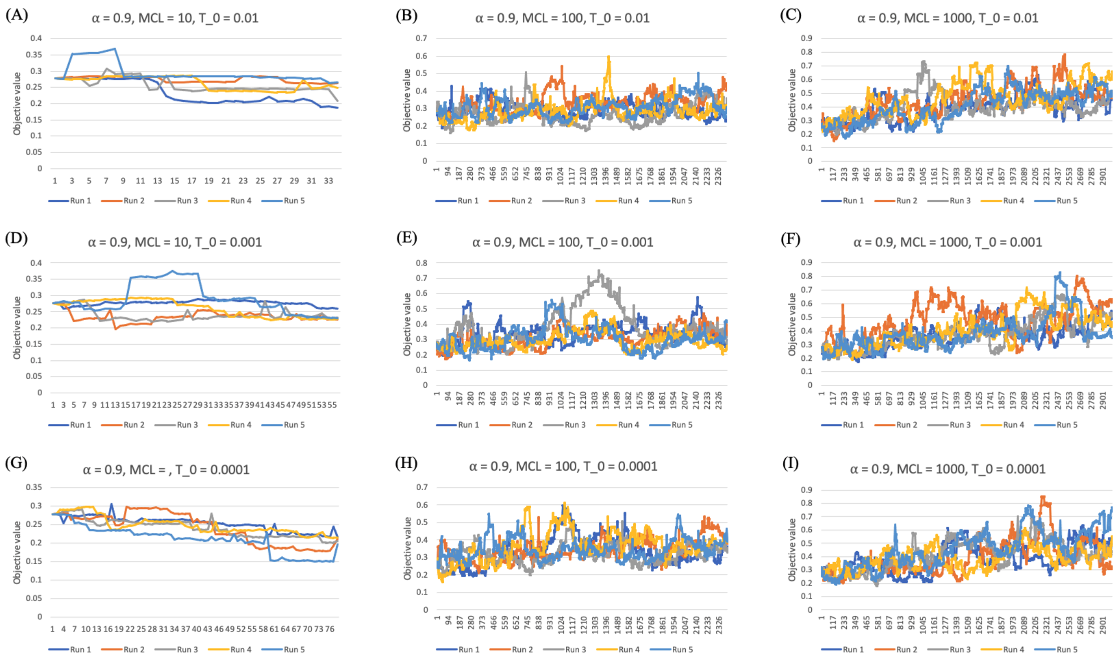


Figure C.5: The outcomes for parameter tuning of the SA algorithm for the medium case study with α of 0.9

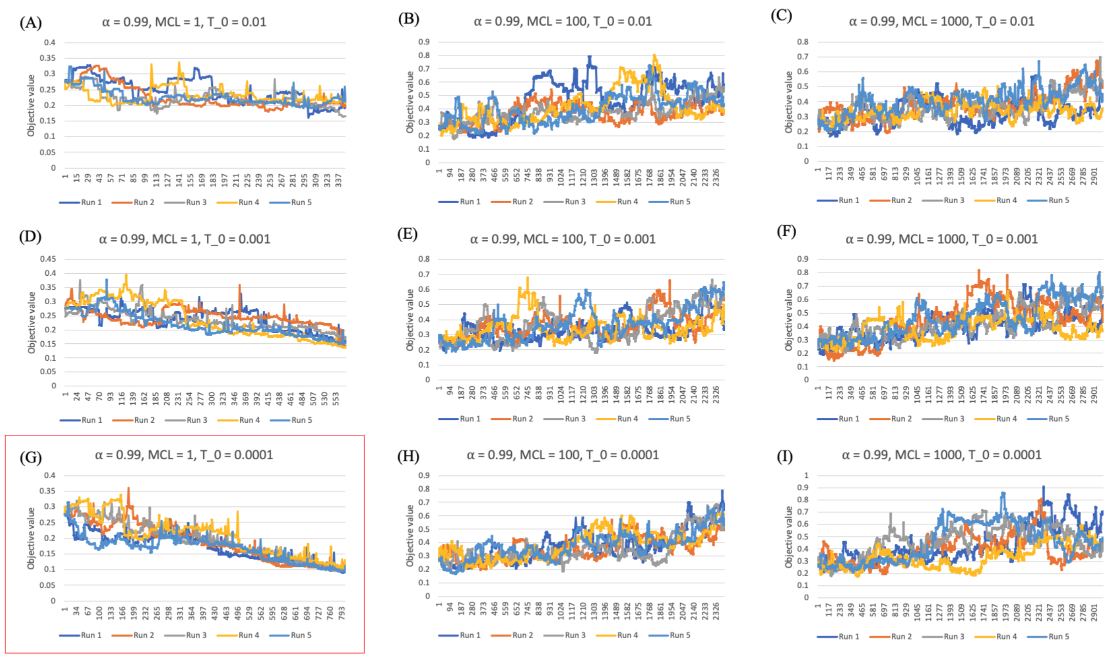


Figure C.6: The outcomes for parameter tuning of the SA algorithm for the medium case study with α of 0.99

C.1.3 Outcomes parameter tuning Large case study

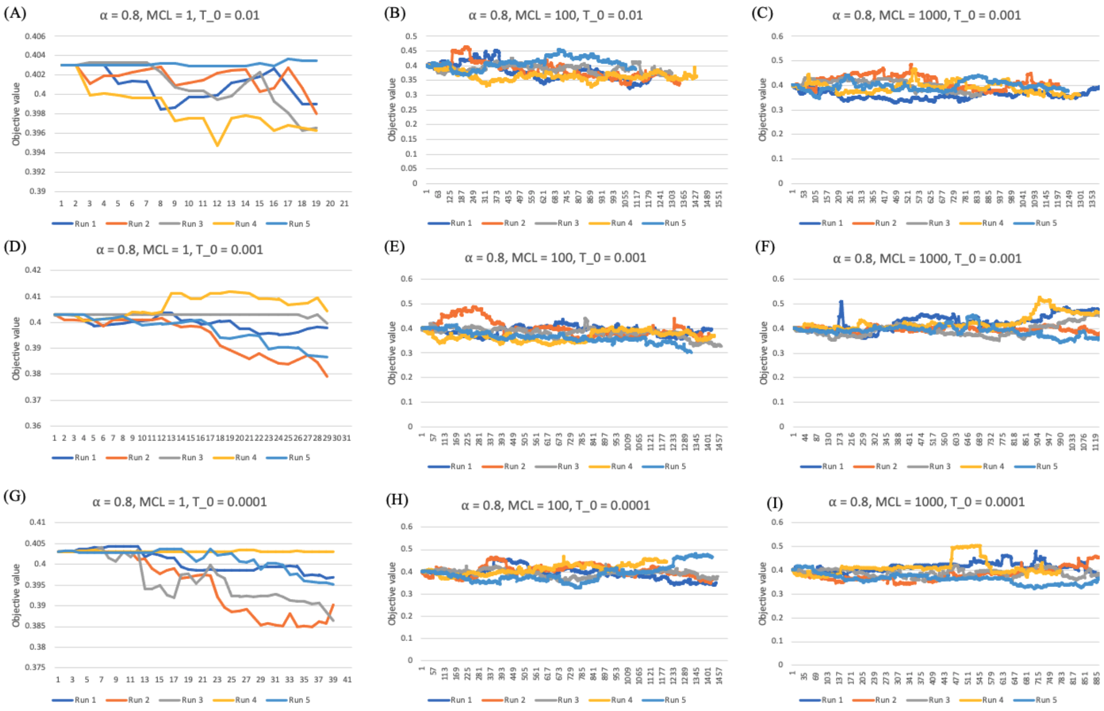


Figure C.7: The outcomes for parameter tuning of the SA algorithm for the large case study with α of 0.8

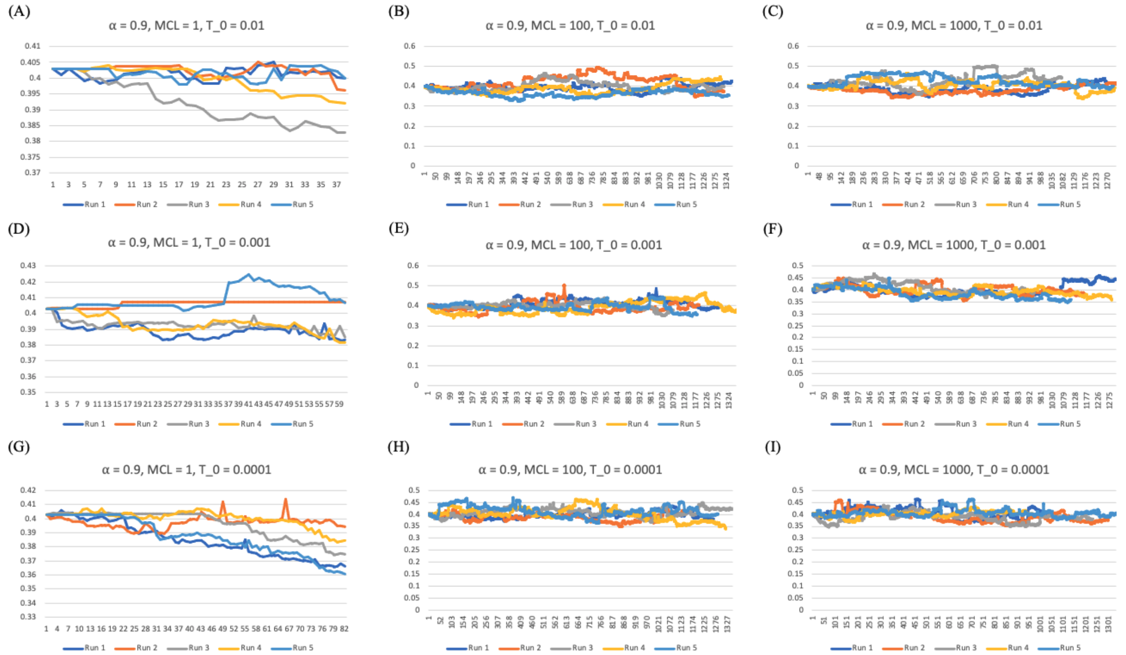


Figure C.8: The outcomes for parameter tuning of the SA algorithm for the large case study with α of 0.9

α = 0.99, MCL = 1000, T_0 = 0.0001

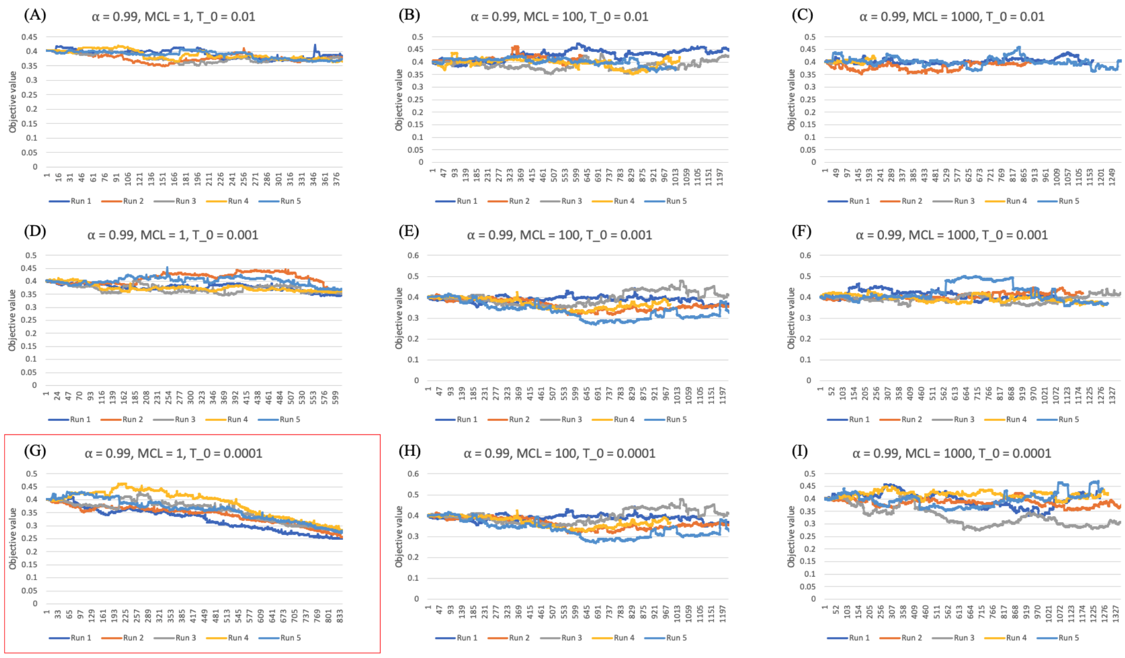
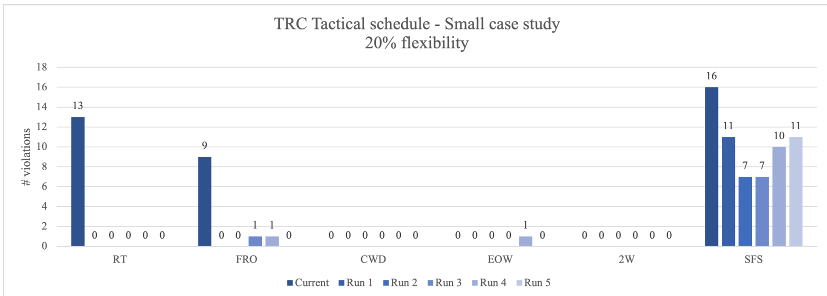


Figure C.9: The outcomes for parameter tuning of the SA algorithm for the large case study with α of 0.99

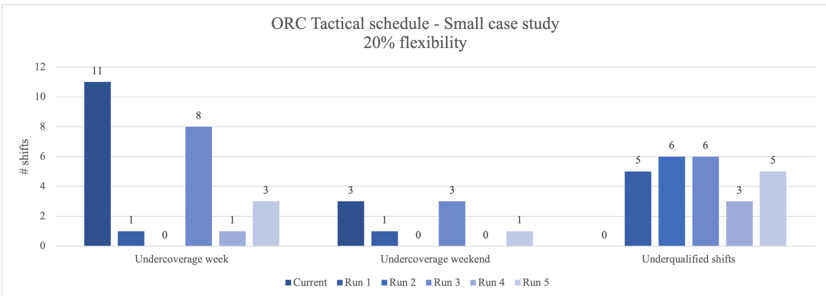
Appendix D

Experimental results case studies

D.1 Experimental results small case study

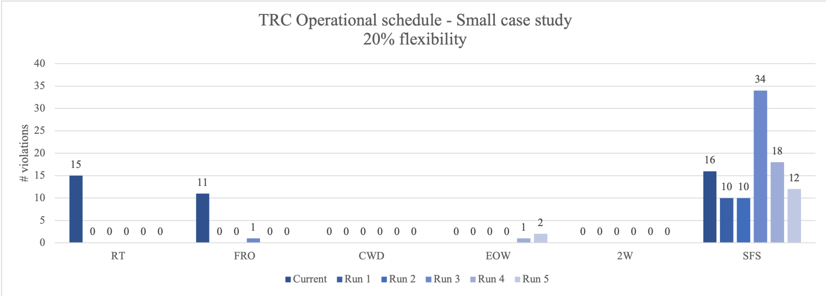


(a) TRC violations Tactical schedule

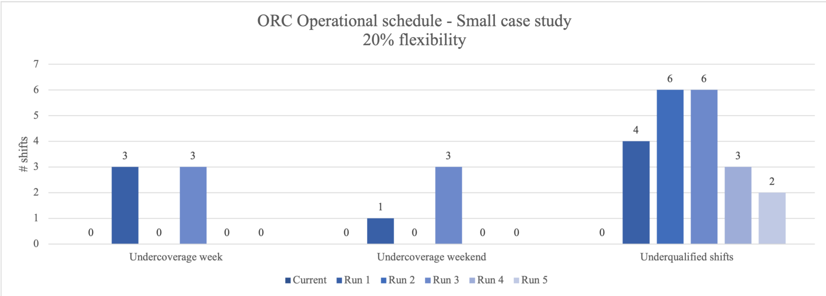


(b) ORC violations Tactical schedule

Figure D.1: Small case study: Results for time-related and organizational violations for tactical schedule for the manual schedule and schedule generated by the proposed method.



(a) TRC violations Operational schedule



(b) ORC violations Operational schedule

Figure D.2: Small case study: Results for time-related and organizational violations for operational schedule for the manual schedule and schedule generated by the proposed method.

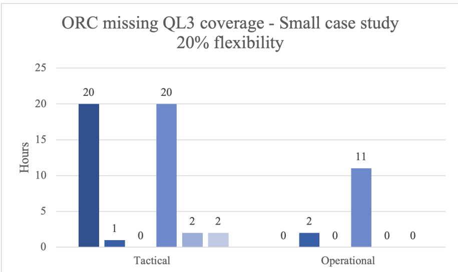
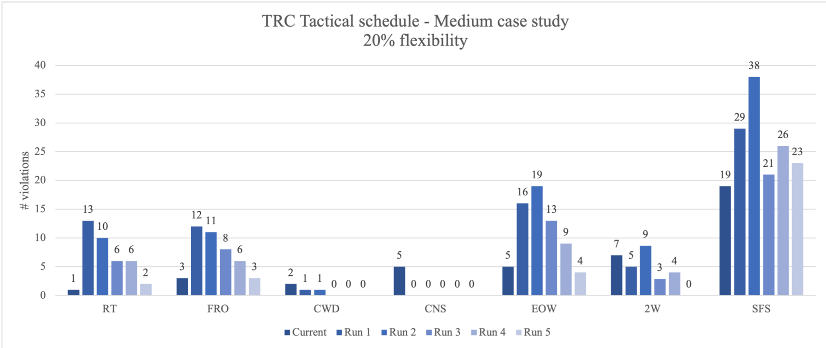
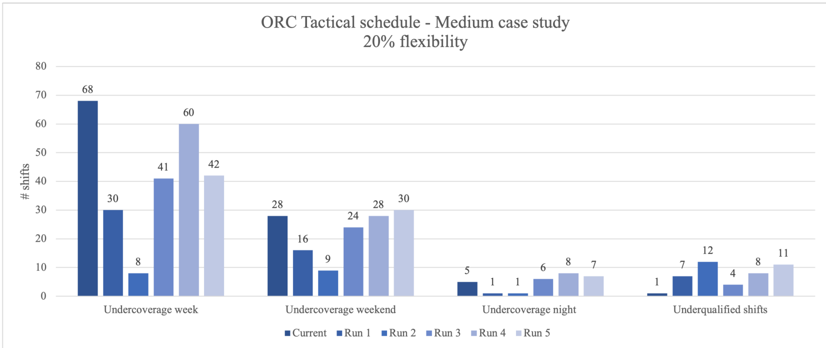


Figure D.3: Missing coverage QL3 in hours for the tactical schedule small case study.

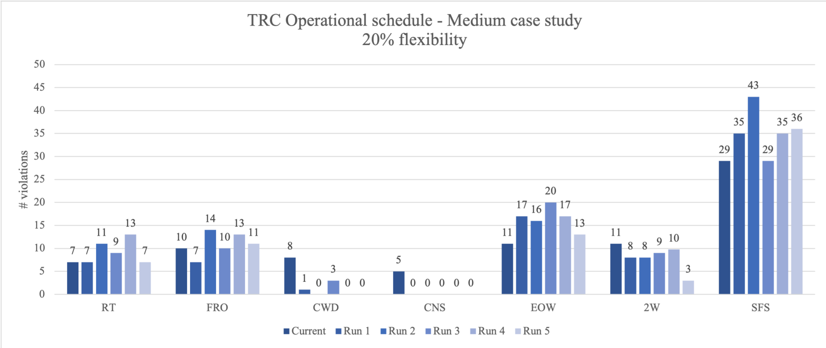
D.2 Experimental results medium case study



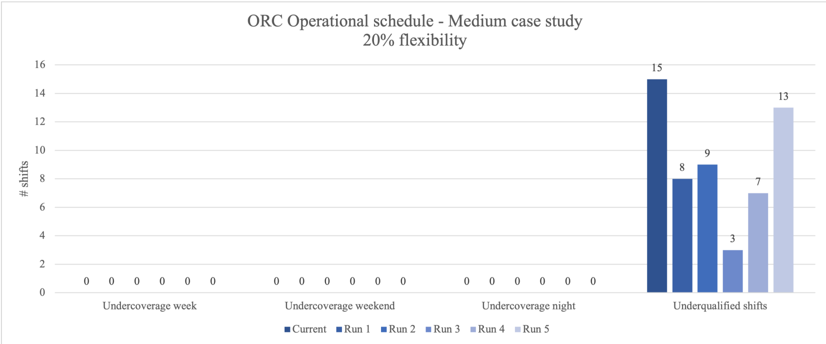
(a) TRC violations Tactical schedule



(b) ORC violations Tactical schedule



(a) TRC violations Operational schedule



(b) ORC violations Operational schedule

Figure D.5: Medium case study: Results for time-related and organizational violations tactical and operational schedules for the manual schedule and schedule generated by the proposed method.

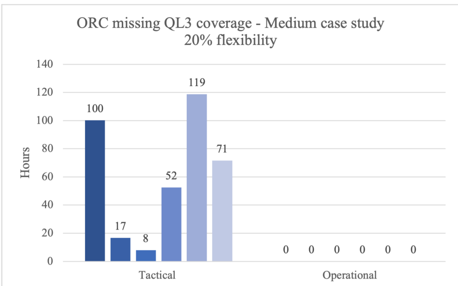
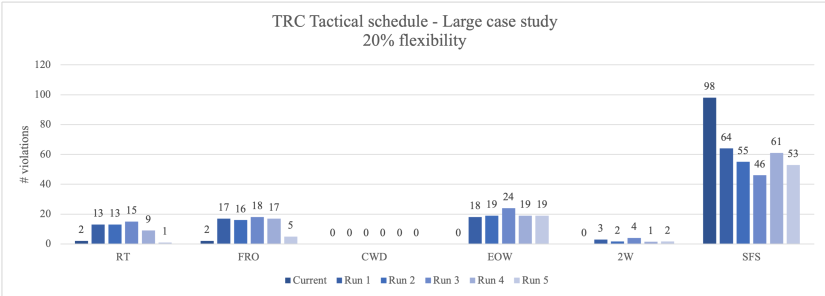
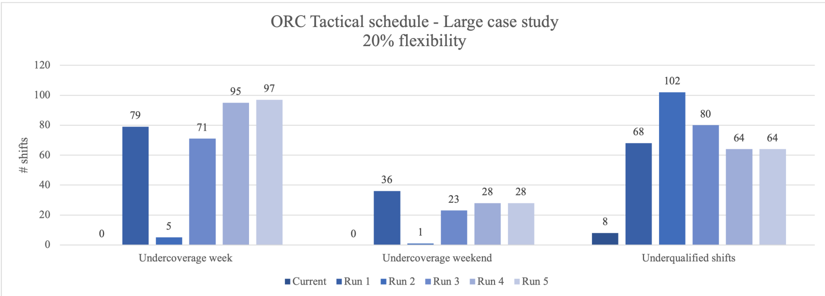


Figure D.6: Missing coverage QL3 in hours for the tactical schedule medium case study.

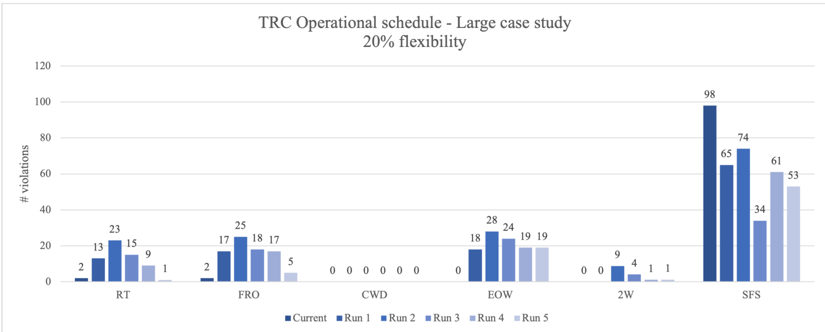
D.3 Experimental results large case study



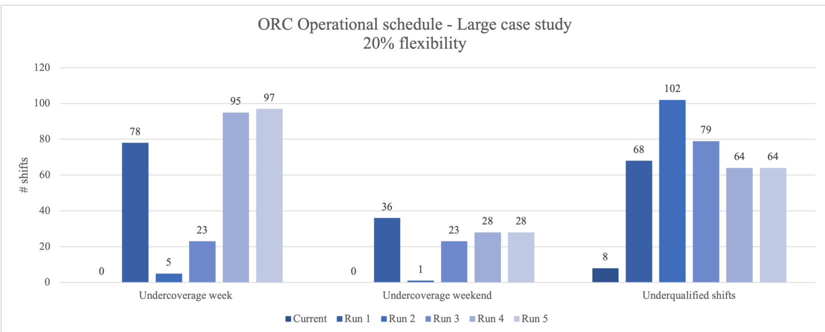
(a) TRC violations Tactical schedule



(b) ORC violations Tactical schedule



(c) TRC violations Operational schedule



(d) ORC violations Operational schedule

Figure D.7: Large case study: Results for time-related and organizational violations for tactical and operational schedules for the manual schedule and schedule generated by the proposed method.

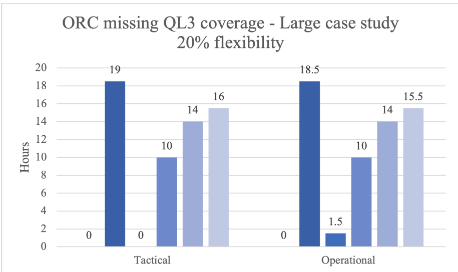


Figure D.8: Missing coverage QL3 in hours for the tactical schedule large case study.