Aalto University
School of Science
Master's Programme in Mathematics and Operations Research

Juhani Mutikainen

Optimization Model for Multicriteria Workforce Planning of Flexible Train Driver Resource

Master's Thesis
Espoo, July 4, 2019

Supervisor: Professor Antti Punkka
Advisor: Teemu Kinnunen M.Sc. (Tech.)

Aalto University
School of Science
Master's Programme in Mathematics and Operations Research

ABSTRACT OF MASTER'S THESIS

Author:	Juhani Mutikainen		
Title: Optimization Model for Multicriteria Workforce Planning of Flexible Train Driver Resource			
Date:	July 4, 2019	**Pages:**	v + 46
Major:	Systems and Operations Research	**Code:**	SCI3055
Supervisor:	Professor Antti Punkka		
Advisor:	Teemu Kinnunen M.Sc. (Tech.)		

Cost-efficient operations are increasingly important for VR, the Finnish state-owned railway company, as it faces more and more competition. Labour costs are a major expense for transportation companies, so there is a great incentive to find ways to minimize unnecessary labour costs, such as overtime work. One way to achieve this at VR is through improved rostering, the creation of work schedules for train drivers.

The current optimization model used for rostering at VR is suitable for minimizing planned labour costs of rosters. We modify the model to account for uncertainties related to freight trains. Using forecasts of probabilities of train cancellations and additional train orders, the model tries to decrease expected overtime costs in operations. The instrument for this is the planning of flexible driver resource, which refers to drivers who can receive their work information much later than regular drivers according to union agreements.

We compare the performance of two versions of the developed model with the original model with a simulation study. The results of the simulation indicate that the modified model can decrease expected overtime costs. However, more studies are needed to confirm the magnitude of cost savings in real-life situations as well as to improve the model further.

Keywords:	crew rostering, stochastic planning, multicriteria sorting, rail transport
Language:	English

Aalto-yliopisto
Perustieteiden korkeakoulu
Master's Programme in Mathematics and Operations Research

DIPLOMITYÖN TIIVISTELMÄ

Tekijä:	Juhani Mutikainen		
Työn nimi: Optimointimalli joustavan veturinkuljettajaresurssin monikriteeriseen työvoimasuunnitteluun			
Päiväys:	4. heinäkuuta 2019	**Sivumäärä:**	v + 46
Pääaine:	Systems and Operations Research	**Koodi:**	SCI3055
Valvoja:	Professori Antti Punkka		
Ohjaaja:	Diplomi-insinööri Teemu Kinnunen		

Kustannustehokkaat toiminnot ovat yhä tärkeämpiä VR:lle, Suomen valtion omistamalle rautatieyhtiölle, sen kohdatessa yhä enemmän kilpailua. Työvoimakulut ovat kuljetusalan yrityksille merkittävä kustannus, minkä takia niiden kannattaa panostaa turhien työvoimakulujen, kuten ylityökustannusten, välttämiseen. Yksi tapa VR:llä tähän on parantaa veturinkuljettajien työvuoroluetteloiden suunnittelua.

VR:llä tällä hetkellä käytössä oleva optimointimalli työvuoroluetteloiden suunnitteluun toimii hyvin suunniteltujen kustannusten minimoinnissa. Tässä työssä tätä mallia muokataan siten, että se ottaa huomioon tavarajuniin liittyviä epävarmuuksia. Malli vähentää toiminnoissa syntyviä odotusarvoisia ylityökustannuksia hyödyntämällä ennustettua tietoa junien peruutuksista ja junien lisätilauksista. Kustannussäästöt mahdollistaa joustavan veturinkuljettajaresurssin suunnittelu. Joustavalla veturinkuljettajaresurssilla tarkoitetaan niitä veturinkuljettajia, joille tarvitsee ilmoittaa työvuorot huomattavasti myöhemmin kuin muille kuljettajille työehtosopimuksen mukaisesti.

Työssä verrataan muokatun mallin kahden eri version suorituskykyä alkuperäiseen malliin verrattuna simulointitutkimuksella. Simulaation tulosten perusteella vaikuttaa siltä, että muokatulla mallilla on mahdollista vähentää odotusarvoisia ylityökustannuksia. Tarvitaan kuitenkin lisää tutkimuksia sekä selvittämään, kuinka suuria vaikutukset ovat tosielämässä, että parantamaan mallia entisestään.

Asiasanat:	työvuorosuunnittelu, stokastinen suunnittelu, monikriteerinen lajittelu, rautatieliikenne
Kieli:	Englanti

Acknowledgements

I wish to thank my family and friends for their support throughout my studies. The insights and support by the resource planning team at VR have been crucial for this thesis. I would also like to thank my advisor Teemu and supervisor Antti for their feedback and guidance.

Espoo, July 4, 2019

Juhani Mutikainen

Contents

Chapter 1

Introduction

1.1 Background

The main railway operator in Finland is the state-owned company VR Group (VR for short). VR has around 6300 employees, of which nearly 1000 are train drivers for freight and long-distance passenger trains (VR Group, 2019). The work schedules of the train drivers utilize synergy from these two train categories so that a work shift can contain trains from both categories. As VR faces competition from both rail and road, cost-efficient planning of the work schedules in general is crucial.

The planning process of work schedules consists of two main steps. The first step is shift planning. In this phase, train driving and other tasks are united manually into shifts. The second step is the assignment of shifts to the drivers of each depot for a planning period of 21 days. This phase is called *rostering* and is the focus of this thesis.

At maximum 17 % of train drivers at each depot can be reserved to be used as *flexible driver resource*. These drivers are called *extra drivers* in this thesis. Instead of a planned roster, extra drivers get to know only six guaranteed days off, while their shifts are announced 1-3 days beforehand. The purpose of the flexible driver resource is to have drivers available for demand not known at the time of rostering. Such unknown demand can result from possible *extra shifts* and shifts originally planned for drivers who happen to be on a sick leave. In addition, a portion of the shifts known before the rostering process cannot fit into the rosters of regular drivers, due to constraints such as maximum work time, so extra drivers should also take care of these *residue shifts*. If necessary, regular drivers can also provide some flexibility with overtime work. However, this is costly and requires permission of the driver, so optimal availability of extra drivers is important to minimize actualized costs.

Some freight trains have a high probability of being cancelled because their demand is inadequately known at the time of planning. Consequently, some shifts are more prone to cancellations than others. Because the rosters of regular drivers cannot be changed without extra costs and permission, these shifts should be left as residue shifts. If a residue shift is cancelled, there will be no cost, as the required work time has not been assigned to any driver. Therefore, residue shifts with low certainty increase the availability of extra drivers compared to more certain shifts. In addition, the availability of extra drivers could be increased by having less residue shifts on those days with more predicted extra shifts.

An optimization model and implementation for automatic rostering has been developed and is used at VR (Porokka, 2017). The implementation solves the rostering problem heuristically and excels in maximizing quickly the planned utilization rate of train drivers and indicators of drivers' work well-being subject to constraints that increase equality between drivers.

1.2 Objectives

This thesis further develops the existing roster optimization model to account for shift cancellations and extra shifts. The main objective of the modified model is to maximize the availability of flexible driver resource for possible extra shifts. The maximization of availability is divided into general and daily availability. General availability refers to the required amount of work hours from residue shifts after cancellations. Daily availability considers simultaneously the number of extra drivers able to work and how many extra shifts and not-cancelled residue shifts are expected on each day. The evenness of daily availability lowers the risk of cancellations due to lack of train driver and can enable more efficient use of extra drivers than general availability alone. The model should be intuitive and enable planners to easily modify it based on their local expertise. Therefore, we choose to utilize multicriteria sorting of shifts in modeling availability. The objectives related to well-being of drivers are untouched in the modified model. The implementation of the modified model should be approximately as quick and easy to use as the original one.

1.3 Structure of the thesis

The rest of the thesis is structured as follows. Chapter 2 reviews relevant literature of crew planning, uncertainty in crew planning context and multicriteria classification and sorting. Chapter 3 presents the modified optimization model. Chapter 4 compares two variants of the modified model to the original one with a simulation study and finally Chapter 5 concludes the thesis and discusses possibilities for future developments.

Chapter 2

Literature review

This chapter presents a review of literature relevant to the thesis. Section 2.1 presents a brief introduction to crew planning, while Section 2.2 examines crew planning under uncertainty. Section 2.3 presents the main approaches to multicriteria classification and sorting.

2.1 Crew planning

Crew planning is often split into two related sub-problems: scheduling and rostering. The problem in crew scheduling is finding a set of shifts that covers all required services or tasks (Valdes and Andres, 2010). Rostering, the focus of this thesis, is creating a work schedule (roster) for a planning period based on the shifts formed in crew scheduling (Lin and Tsai, 2019). Generally, the goal in crew planning is minimizing operational costs and maximizing the personal preferences of staff. However, the details and terminology of the sub-problems vary considerably. Some aspects depend on the field of application while others are specific to the company or organization in question (Valdes and Andres, 2010).

Labour legislation and union agreements set numerous compulsory constraints on both crew scheduling and rostering. In addition, some shifts may require a qualification, which places constraints especially on rostering. Non-compulsory constraints may also be placed to satisfy crew preferences, such as days off and shift equity (Ernst et al., 2004).

Rostering can be divided into two types: cyclic and non-cyclic rostering. In a cyclic roster, each worker follows the same long-term roster. The length of the long-term roster (in planning periods) is the same as the number of crew. After each period, the roster for each worker is the next row in the previous roster. In non-cyclic rostering, a unique roster for each worker is created for each planning period. Cyclic rostering requires identical qualifications and similar preferences between workers. While cyclic rosters can be fair and easy to maintain, they are quite inflexible regarding changes in demand (the number and timing of shifts) and supply (the number and availability of workers) (Xie and Suhl, 2015).

There are several types of methods and algorithms to solve crew planning problems. When the problem is highly constrained and the main issue is finding a feasible solution, constraint programming can be useful. For relatively simple problems, the best solutions can often be achieved with algorithms based on mathematical programming approaches, such as mixed-integer linear programming. Metaheuristics, including simulated annealing, tabu search and genetic algorithms, are generally suitable for problems that are too difficult for mathematical programming (Ernst et al., 2004).

Crew scheduling and planning are used in numerous fields. Examples include transportation, call centres, health care and protection and emergency services (Ernst et al., 2004). Lin and Tsai (2019) integrate train crew scheduling and rostering to a composite problem and solve it using a branch-and-price-and-cut and a depth-first search-based algorithm. Based on empirical studies with different group sizes (from 12 to 84 drivers), the authors conclude that the integrated approach can produce better solutions than scheduling and rostering separately. Burke et al. (2010) present a hybrid model of integer programming and variable neighbourhood search for nurse rostering. The authors compare the performance of the model to approaches using either a genetic algorithm or variable neighbourhood search. The hybrid model outperforms the compared algorithms on 12 rostering problems from a Dutch hospital.

Örmeci et al. (2014) develop a mixed-integer programming model for rostering in call centres with operational cost, agent satisfaction and customer service objectives as goals. They analyze the solutions of the model with through numerical experiments using data from a call centre in Istanbul with a total of 250 employees in four regions with 8 different types of calls corresponding to a required skill. Gendreau et al. (2006) discuss physician scheduling in emergency rooms using four different techniques, such as tabu search and constraint programming. They formulate the constraints of the problem based on practical experiences from five hospitals in Montreal area but do not numerically analyze the performance of different techniques.

2.2 Uncertainty in crew planning

While the real-life environments of crew planning usually involve considerable uncertainties, most crew planning models are deterministic, which means all forms of uncertainties are ignored. However, some literature exists on stochastic models as well. Van den Bergh et al. (2013) classify uncertainty in planning into three main categories: uncertainty of demand, arrival and capacity. Uncertainty of demand represents the predictability of the quantity of required workload. Uncertainty of arrival is concerned with the specific timing of required workload and is thus highly related to uncertainty of demand. For example, the number of patients in a hospital corresponds to uncertainty of demand, whereas the arrival of calls to a call center corresponds to uncertainty of arrival. Uncertainty of capacity indicates the predictability of available workforce (due to sickness, for example).

Approaches to uncertainty in crew rostering can be reactive or proactive (Ingels and Maenhout, 2015). In reactive approaches, the original roster is adapted after disruptions during operational planning to minimize resource shortages. Overtime work, schedule changes and using existing reserve shifts are typical methods to react to disruptions. Proactive approaches attempt to build robust rosters that will be less likely to be affected by these disruptions. A typical proactive method is introducing buffers. Time buffers, such as additional work time to a shift, can be used in crew scheduling. In crew rostering, capacity buffers, such as planning reserve shifts, can be used. While there is often a trade-off between planned costs and roster robustness, expected operational costs can be lowered by increasing robustness. The quality of robustness is usually measured by simulating disruptions to rosters and comparing to other approaches.

Ingels and Maenhout (2015) study the impact of reserve shifts on roster robustness. They implement five different strategies to proactively plan reserve shifts in rostering. The strategies are compared after a simulation of disruptions (demand and capacity) and reactive re-scheduling. While the authors conclude that capacity buffers in rosters are necessary, they note that the size and positioning of buffers is of great importance to robustness. The authors have also produced several other reports on roster robustness. The report from 2017 (Ingels and Maenhout, 2017) formulates the concept of employee substitutability to improve robustness. On a given day, each employee either has a shift requiring certain skill or a day off. However, three types of changes to the assignment can be possible: between-skill, within-skill and day-off-to-work substitutions. Individual employee substitutability is the weighted sum of the value of these substitution possibilities. Group employee substitutability measures the availability of these substitution possibilities after additional capacity buffers are created. After a simulation of demand and capacity disruptions and re-scheduling, rostering strategy based on individual employee substitutability is found to produce more robust rosters than group substitutability or minimum cost strategies.

A few studies apply uncertainty to a certain field in crew planning. Schaefer et al. (2005) study airline crew scheduling under uncertainty. They create a scheduling model that considers the expected costs of frictional disruptions, i.e. short-term delays due to various reasons. Based on simulation results, the authors conclude that the created schedules perform better in operations than those which use only planned cost.

Punnakitikashem (2007) develops an integrated model for nurse staffing and assignment under uncertainty. Nurse staffing and assignment are the operational phases of nurse planning after rostering. Staffing nurses to shifts is done 90 minutes before each shift, while assigning nurses to patients occurs 30 minutes before each shift. The goal of the two-stage stochastic integer programming model is to minimize excess workload on nurses under short-term demand uncertainty while keeping the budget in control.

Parisio and Jones (2015) develop a two-stage stochastic mixed integer programming model for rostering in retail outlets with demand uncertainty. The authors generate demand scenarios from historical data and run a simulation. The stochastic model performed better than the older deterministic model in the roster quality metric chosen by the authors, which measures the sum of differences between the number of assigned employees and actualized number of customers on each hour.

2.3 Multicriteria classification and sorting

The problem in multicriteria classification and sorting is assigning a set of alternatives into predefined homogeneous groups based on more than one criterion. Zopounidis and Doumpos (2002) present a literature review of the methods in multicriteria classification and sorting. The difference between classification and sorting is in the definitions of the groups. In classification, the groups are nominal, while groups that can be ordered based on the preferences of the decision maker belong to sorting. This thesis uses the formal notation by Zopounidis and Doumpos (2002). The finite set of n alternatives $A = a_1, a_2, \ldots, a_n$ should be assigned into q predefined groups $C_1, C_2, \ldots, C_q$. In a sorting problem, the group C_1 is the most preferred while the group C_q is the least preferred. A vector of m criteria $g = (g_1, g_2, \ldots, g_m)$ is used to classify/sort the alternatives. The performance of each alternative is evaluated on each criterion, so any alternative a_i can be described by vector $a_i = (g_{1i}, g_{2i}, \ldots, g_{mi})$, where g_{ji} is the performance on criterion g_j. In order to classify or sort an alternative, the performance vector needs to be aggregated by a suitable model.

There are three main types of approaches to aggregate the criteria: outranking, utility functions and models based on decision rules (Zopounidis and Doumpos, 2002). The central characteristic in outranking is the existence of thresholds determining when to prefer one over another and when to be indifferent (Yevseyva, 2007). In addition, alternatives can be incomparable.

According to Zopounidis and Doumpos (2002), the most widely used outranking method is ELECTRE TRI introduced by Yu (1992). In ELECTRE TRI, alternative a_i outranks alternative a_p if there are enough arguments to support that a_i is at least as good as a_p and if there is no argument to disprove this. The groups used in sorting are distinguished by reference profiles, which are defined by performance vectors equivalently to the alternatives. Reference profile $r_k = (r_{k1}, r_{k2}, \ldots, r_{km})$ distinguishes classes C_k and C_{k+1}. The comparison of alternatives and reference profiles is performed by pairwise comparison on all criteria. Alternative a_i is preferred to profile r_k if the weighted sum (based on importance) of criteria where a_i is at least as good as r_k is enough to reach the threshold used in sorting, while the equivalent weighted sum for r_k is not enough to reach the threshold (Cardinal et al. 2011). The mathematical formulation is

$$\sum_{j:g_{ji} \geq r_{kj}} w_j \geq \lambda$$
$$\sum_{j:r_{kj} \geq g_{ji}} w_j < \lambda,$$

where λ is the threshold and w_j is the weight of criterion j. However, if a_i is much worse than r_k on a single criterion, a veto evaluation can prevent a_i from outranking r_k. In those cases, the alternative and the profile are considered incomparable. If both the alternative and the profile reach the threshold, they are indifferent (Zopounidis and Doumpos, 2002).

Doumpos et al. (2009) present an approach based on differential evolutionary algorithm to help specifying the parameter values for ELECTRE TRI, while Damart et al. (2007) provide a methodology for that based on guided group discussions. Cardinal et al. (2011) apply ELECTRE TRI to student selection of a popular major in a French engineering school. Based on criteria such as GPA, motivation and personality, students are divided into four different classes to help the selection process.

Approaches based on utility functions (or value functions) create a marginal utility function for each criterion to transform the relevant scale of the criterion into terms of utility. Typically, the utility function is simply the sum of the marginal utility functions. In case of m criteria, the additive utility function is $U(g) = \sum_{j=1}^{m} u_j(g_j) \in [0, 1]$, where $u_j(g_j)$ is the marginal utility function for criterion g_j. The simplest approach to using utility theory in sorting is the UTADIS method (Jacquet-Lagreze and Siskos, 1982). In UTADIS, if the total utility of an alternative a_i is at least as high as the lower bound threshold of a class C_q, a_i belongs either C_q or a higher level class (Zopounidis and Doumpos, 2002). Köksalan and Özpeynirci (2009) introduce an interactive sorting method for additive utility functions. The interactive approach guarantees to classify alternatives correctly when the decision maker has preferences consistent with any additive utility function. Zopounidis and Doumpos (2000) present an interactive multicriteria decision support system for sorting problems. The decision maker has the possibility to use four different additive utility models: UTADIS and three of its variants (UTADIS I, UTADIS II and UTADIS III).

Rule-based models typically use "if... then..." type of decision rules, which are inferred from class assignment examples. A popular methodology for determining the rules is based on rough sets theory (Pawlak, 1982), which has the advantage of enabling creation of decision rules even if there are inconsistencies in the assignment examples. Each rule consists of a condition and decision part. The condition part states a specific profile on a subset of criteria, which is used to compare alternatives. The decision part determines an assignment to at least or at most a given class (Zopounidis and Doumpos, 2002). For decision makers, rule-based models are typically intuitive because of their connection to assignment examples, which can be of great benefit. Azibi and Vanderpooten (2002) propose method for creating decision rules that expresses the rules by linear constraints. The consistency of rules is tested by solving a series of linear programs. Chen et al. (2012) introduce an approach based on rough sets to determine linguistic decision rules.

Chapter 3

Optimization model

This chapter presents the modified optimization model for the rostering problem at VR. We present the background of the problem and a verbal description of relevant factors in Section 3.1. The mathematical formulation and solution methodology of the model is introduced in Section 3.2.

3.1 Description

3.1.1 Crew planning at VR

Union agreements specify that work for train drivers in Finland must be planned for three-week work periods and published at least one week before the beginning of each work period. The crew planning process consists of three different phases: scheduling, rostering and operational planning.

Crew scheduling creates shifts that cover all work demand known at the time of planning. The work can be driving freight and passenger trains or other train-related tasks, such as preparing a train for driving. The goal in scheduling is to combine the tasks to shifts efficiently, while also complying with labour law and union agreements. For example, there are limits to the lengths of shifts and tasks. Figures 3.1 and 3.2 present graphical examples of shifts. Typically, some shifts require that a driver should move from one location to another when there is no need for driving a train. In those cases, the shifts require additional traveling tasks. Scheduling is currently done manually using proprietary software and simultaneously for all driver depots in Finland. Each shift is assigned to a home depot, which is the start and end station of the shift. The capacity at each depot (skills and available drivers) must be considered while determining how the workload is divided between them. There are a few important shift concepts and types, which are explained below:

Night work
The amount of work time between 22:00 and 06:00

Artificial work time
Work time with added time compensation from evening (21:00 – 22:00) and night work

Night shift (type A)
A shift with three hours or more of night work in a single night

Night shift (type B)
A shift with work between 02:00 and 05:00

Shift with rest
A shift which contains a time period not counted as work time

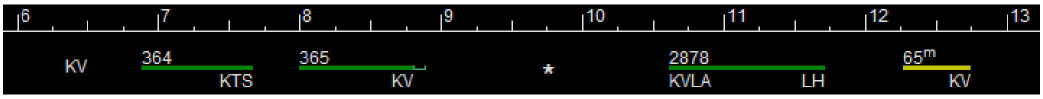

Figure 3.1: Example of a day shift. First driving a passenger trains from KV to KTS and back. After a break (the star sign), driving a freight train from KVLA to LH and back by traveling. KV and KVLA are at the same location in practice

Figure 3.2: Example of a night shift with rest. The shift starts with a short locomotive task to KUK. From KUK to RIT there is a freight train driving task, which is split into two parts with a break in KVLA. After a rest in RIT (two star signs), driving a freight train back to KV. The last train includes work between 02:00 and 05:00, which makes the shift a type B night shift. Hour 28 refers to 04:00, for example

Rostering takes place after scheduling and is the creation of rosters for each depot based on the shifts from scheduling. The depots are planned separately. Large depots have more than 100 drivers, while small depots have less than 20 drivers. Rostering is performed using optimization software implemented in C++ and R languages (Porokka, 2017), which is integrated into the same software crew scheduling uses. The drivers at each depot are split into regular and extra drivers. After rostering, extra drivers know only their six days off during the work period, unlike regular drivers, who need to know both their shifts and days off. Extra drivers act as flexible train driver resource, which is one form of capacity buffer (Ingels et al., 2015). According to union agreements, at maximum 17 % of drivers (rounded up) at each depot can be used as extra drivers. In practice, this usually means that each driver acts as an extra driver once in six work periods. Figure 3.3 presents a roster for regular drivers, while Figure 3.4 shows a roster for extra drivers. In addition to shift types and concepts, there are a couple of roster related concepts:

Double week rest
: Two consecutive calendar days without work

Work cluster
: A period of work between two double week rests

Sunday work
: Work time on Sundays and certain public holidays

Residue shift
: A shift not planned for regular drivers

Absence
: A driver has permission to be away from work due to vacation or a sick leave, for example

Because of either labour law and union agreements or the work well-being of drivers, rostering needs to account for the following constraints (Porokka, 2017):

C1. Each driver has a limit to the artificial work time during the work period. If the work period has no public holidays and the driver is a full-time worker without absences for the work period, the limit is 114 hours and 45 minutes

C2. At most 42 hours of night work during the work period

C3. At most five calendar days between two consecutive double week rests

C4. No more than 45 hours of artificial work between two consecutive double week rests

C5. At least ten hours of rest between two consecutive shifts

C6. No night shifts of type B on consecutive nights (without permission from union representative)

C7. The maximum number of consecutive nights with night shifts is two

C8. The drivers can only be assigned to shifts for which they have the required skills

C9. The drivers cannot be assigned to any shifts during absences

C10. Each shift is assigned to a driver

C11. Depot specific maximum on the amount of Sunday work

C12. Depot specific maximum on the number of night shifts

C13. Depot specific maximum on the number of shifts with rest

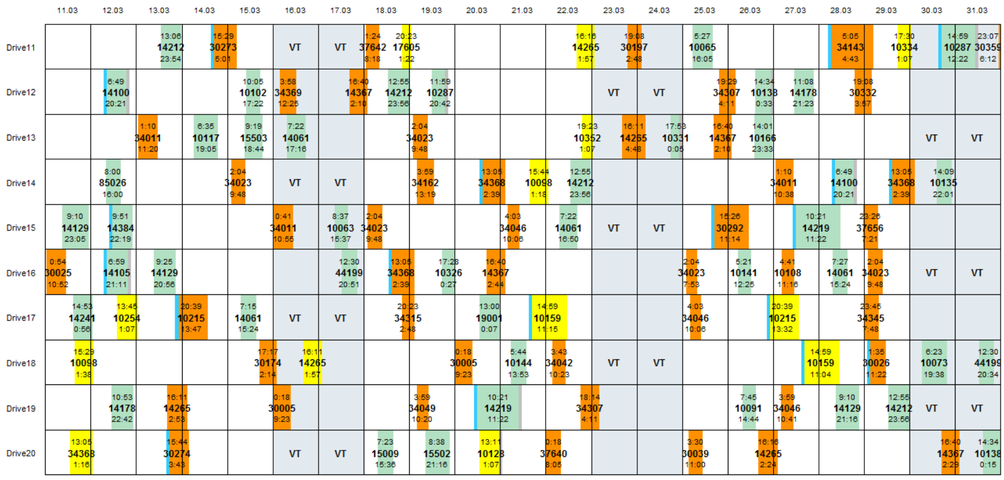

Figure 3.3: Example of roster for regular drivers. Orange shifts are type B night shifts and yellow ones are type A night shifts. Blue line in left side of shift indicates a shift with rest. "VT" indicates a compulsory weekend rest during the work period. Driver 11 has four work clusters, which are separated with double week rests.

	11.03	12.03	13.03	14.03	15.03	16.03	17.03	18.03	19.03	20.03	21.03	22.03	23.03	24.03	25.03	26.03	27.03	28.03	29.03	30.03	31.03
Drive11						v	v	v	v								v	v			
Drive12						v	v				v	v				v	v				
Drive13							v	v					v	v				v	v		
Drive14	v	v	v										v	v	v						
Drive15	v	v	v										v	v	v						
Drive16	v	v											v	v	v	v					
Drive17				v	v							v	v	v	v						
Drive18	v	v							v	v										v	v
Drive19				v	v				v	v										v	v
Drive20	v	v								v	v									v	v
Drive21	v	v									v	v								v	v

Figure 3.4: Example of a days off roster for extra drivers. Each extra driver is planned six days off. One weekend rest at minimum and at least one day off each week. Single days off are unpreferred

After the rostershave been published to drivers, operational planning oversees the work period. Operational planners need to adjust the rosters to demand, arrival and capacity disruptions, such as additional work demand (extra shifts) and sick leaves. Each residue shift which is not cancelled must also be assigned to a driver. The primary instrument for operational planning is the use of flexible train driver resource (extra drivers). The secondary instrument is to assign overtime work to regular drivers. However, this is costly and always requires permission from the driver. In addition, it is possible to use extra drivers after the deadline for shift announcement, but this demands permission and additional compensations.

For example, the extra driver roster in Figure 3.4 has five extra drivers available on the second day of the period. If there are three residue shifts for the second day, extra drivers are able to do at most two extra shifts. If the number of extra shifts is higher, the rest will be overtime work for regular drivers, even if extra drivers were underused on all other days. However, if one of the three residue shifts is cancelled, extra drivers are able to do one more extra shift, which can decrease overtime work. The regular driver number 15 in Figure 3.3 has shift 14129 assigned on the first day and 14384 on the second day of the period. If the customers of the freight trains of shift 14384 want to change the departure day to the first day, driver 15 cannot do the shift regardless of permission due to shift 14129. The shift on the second day will be cancelled for driver 15, while an equivalent shift will be an extra shift for the first day.

3.1.2 Shift structure

A shift typically consists of two or more tasks. The first task of a shift always starts at the home depot of the driver and the last task ends there. The most common type of task is driving a train from one location to another. The driver can also be required to drive only a locomotive or travel as a passenger in a train. Other possible tasks, such as preparing the next train for driving, are cancelled if the trains of the shift are cancelled, so we can ignore them in this thesis. Figure 3.5 presents a simple shift structure, while Figure 3.6 presents a more complicated shift.

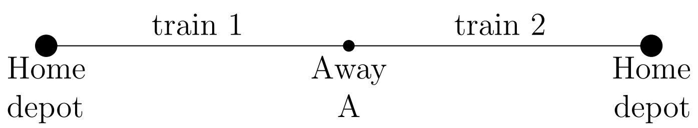

Figure 3.5: A shift consisting of two train driving tasks

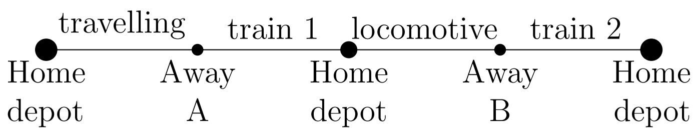

Figure 3.6: A shift with traveling as passenger, driving a locomotive and two train driving tasks with a visit in the home depot in the middle

3.1.3 Shift certainty and adaptability

The certainty of a shift refers to the probability that its required working time will not change after rostering. The working times of regular drivers cannot be altered without permission and extra cost, even if their shifts are cancelled, while extra drivers are more flexible. Therefore, residue shifts should be those with low shift certainty.

After the rosters have been planned and published, some trains in the shifts can be cancelled for various reasons. Depending on shift structure, train cancellations can lead to a) cancellation of the whole shift, b) cancellation of a part of the shift or c) none or only minor changes to the working time of the shift. As every shift must start and end at home depot, a whole shift is cancelled only if all tasks in it are cancelled. Otherwise a cancelled train is replaced by traveling as a passenger or driving only a locomotive. If the original shift includes traveling tasks, they are also cancelled if all train driving tasks are cancelled, because there is no need for them.

Some shifts can be divided into multiple parts if home depot is visited in the middle of the shift at least once. Such shifts can be partially cancelled if all tasks in one of the parts are cancelled. The different types of shift structures and effects of cancellations on them are presented in Figures 3.7, 3.8, 3.9 and 3.10.

Occasionally, the first or the last task of a shift is simply traveling as a passenger from a depot to another. In these cases, the home depot of the shift could be changed if necessary. For example, if a depot A shift begins with traveling from A to depot B, the shift could easily be changed to a depot B shift that ends with traveling from A to B. Thus, such a shift can be helpful as a residue shift if depot B has more flexible driver resource available during that time. In this thesis, we call the possibility to change depot of a shift as shift adaptability. Example of a depot change is presented graphically in Figure 3.11.

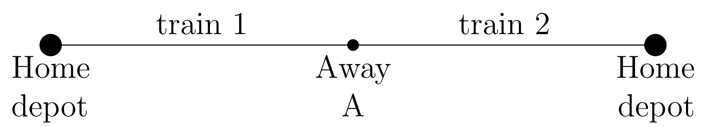

Figure 3.7: A shift cancelled only if trains 1 and 2 are both cancelled. Otherwise cancellations are replaced by traveling or driving a locomotive

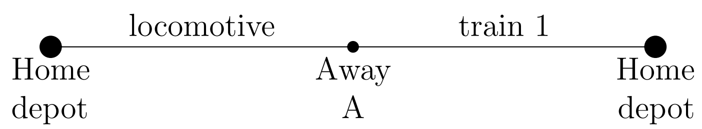

Figure 3.8: A shift cancelled if train 1 is cancelled

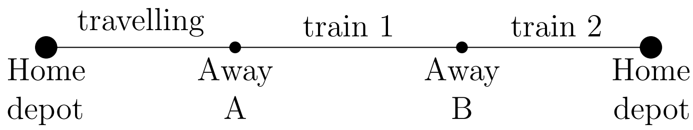

Figure 3.9: A shift cancelled only if trains 1 and 2 are both cancelled. If train 1 is cancelled, the shift starts with either traveling or locomotive from home depot to B. In some cases, this leads to cancellation of some working time

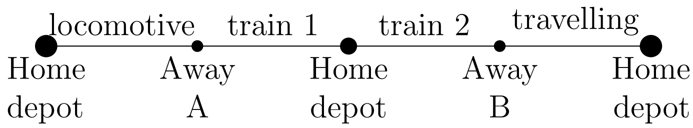

Figure 3.10: A shift cancelled only if trains 1 and 2 are both cancelled, while cancellation of either train is enough for a partial cancellation of shift, because of a visit to the home depot. The visit to the home depot divides the shift into two parts: trip to A (first part) and trip to B (second part). For example, if train 1 is cancelled, the first part of the shift is cancelled and only the trip to B is required

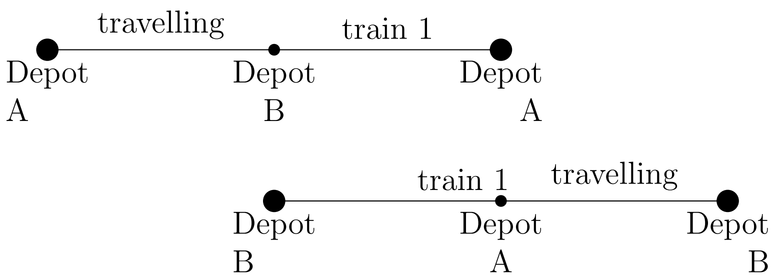

Figure 3.11: A shift before and after a depot change from A to B

3.1.4 Objectives

The first and main objective of the model is to maximize the availability of flexible train driver resource for expected extra shifts. The purpose of availability is to decrease expected costs. Availability consists of general and daily availability.

General availability depends on both the shift certainty and adaptability of residue shifts. We do not try to calculate expected artificial work hours of residue shifts, because cancellations of whole shifts usually increase availability for extra shifts more than partial cancellations, even if the cancelled work time is equal. Therefore, we minimize the planned artificial work time of residue shifts while giving shifts not prone to cancellations priority for regular drivers. In addition, the possibility to change the depot of a residue shift can increase availability at the depot where the shift would be moved from. Thus, we also give some priority to shifts with no adaptability for regular drivers.

Daily availability considers both the number of extra drivers available and the distribution of shifts predicted for extra drivers. The predicted shifts for a given day are the sum of the forecast number of extra shifts and the forecast number of residue shifts not cancelled. The evenness of daily availability during a planning period should be maximized in order to prevent high peaks in the number of shifts which require drivers to accept overtime work.

The second objective is the compactness of rosters, which is a shared objective with the old model (Porokka, 2017). Rest time between two consecutive shifts should be as close to 10 hours as possible, unless there is a double week rest between them. Also, the number of work clusters should be minimized. Compactness is aimed at improving the work well-being of train drivers.

3.2 Formulation and solution

3.2.1 Solution methodology

This thesis uses the solution methodology developed by Porokka (2017), as the constraints of the problem have not changed. The methodology is based on a heuristic *Adaptive Large Neighborhood Search (ALNS)* algorithm by Ropke and Pisinger (2006). The basic idea of the algorithm is to improve solutions by removing some shifts from a roster and then trying to put the shifts back to different drivers.

There are three stages to the search of a solution: initial solution, feasible solution and solution improvement. After the initial solution, some shifts are left unassigned but all constraints are met by the assigned shifts. After the feasible solution, all shifts are assigned to drivers. In the first two stages, the actual objective function is irrelevant as feasibility is the only goal in improvement.

After a feasible solution is found, the solution can be improved with regard to the objective function with Algorithm 1 by Porokka (2017). While generally only improvements to the solution should be permitted, the algorithm sometimes allows feasible but inferior solutions with probability p_{acc} to enable escaping local optima. Possible insertion strategies used in the algorithm include greedy insertion, greedy random insertion and insertion based on a regret heuristic. Greedy random insertion is the fastest strategy and adequate for solving the problem. The shifts to be inserted are put in a random order and then sequentially inserted back to the roster. If there are multiple drivers to whom a shift can be inserted, the insertion with the lowest increase in objective function is chosen.

Algorithm 1 Solution improvement

feasible roster R
$R_{best} \leftarrow R$
repeat
    $R' \leftarrow R$
    choose a random number r from uniform distribution $U(0,1)$
    remove $q \in \{q_{min}, ..., q_{max}\}$ shifts from roster R
    choose insertion strategy s
    insert the q removed shifts back to R using strategy s
    if insertion did not result in a feasible solution **then**
        $R \leftarrow R'$
    else if $(f(R) > f(R'))$ and $(p_{acc} < r)$ **then**
        $R \leftarrow R'$
    else if $f(R) < f(R_{best})$ **then**
        $R_{best} \leftarrow R$
    end if
until the stopping criterion is met
R_{best} contains the best feasible solution found

3.2.2 General availability

For simplification purposes, we divide shifts into four priority classes based on shift certainty and adaptability. We split shift certainty into two criteria (whole and partial cancellations), so we perform multicriteria sorting with three criteria. Each priority class corresponds to a certain level of adjustment to the artificial work time of shifts. General availability is then achieved by minimizing the adjusted work time of residue shifts.

The first criterion is whole cancellations, q_1, which is the probability that the whole shift is cancelled. We assume that each train task has a cancellation probability p_i that is independent from other train tasks. Locomotive and traveling tasks are assumed to be always cancelled if the related train tasks are cancelled. Figure 3.12 presents an example of a shift with four tasks. The probability of all tasks in the shift being cancelled is

$$\begin{aligned} q_1 &= p_1 \cdot p_2 \cdot p_3 \cdot p_4 \\ &= p_1 \cdot p_3, \end{aligned}$$

where $p_2 = 1$ and $p_4 = 1$ are the cancellation probabilities of locomotive and traveling tasks, respectively. For a shift with n tasks,

$$q_1 = \prod_{i=1}^{n} p_i. \tag{3.1}$$

The second criterion is partial cancellations, q_2, which is the probability that only a part of the shift is cancelled. This excludes the probability of the whole shift being cancelled (q_1). Partial cancellations are not as valuable for availability as whole cancellations, because they require a driver. The shift in Figure 3.12 has two parts, A and B. According to addition law of probability, if $P(A)$ is the individual cancellation probability of part A, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ is the probability that either part A or part B is cancelled (or both). Thus, we can calculate the probability that only part A or B is cancelled in shift of Figure 3.12 as

$$\begin{aligned} q_2 &= P(A \cup B) - q_1 \\ &= P(A) + P(B) - P(A \cap B) - q_1 \\ &= p_1 \cdot p_2 + p_3 \cdot p_4 - p_1 \cdot p_2 \cdot p_3 \cdot p_4 - q_1 \\ &= p_1 + p_3 - 2q_1. \end{aligned}$$

For a shift with n tasks and a visit to home depot after task j,

$$q_2 = \prod_{i=1}^{j} p_i + \prod_{i=j+1}^{n} p_i - 2q_1.$$

Occasionally, a shift can have more than one visit to the home depot, which means the shift can be split into three or more parts. The general form of q_2 for a shift with that can be split into parts $A, B, ..., N$ can be stated as

$$q_2 = P(A \cup B \cup ... \cup N) - q_1.$$

The third criterion is shift adaptability, q_3, which is whether the home depot of the shift can be changed. In practice, this means whether the first or last task of the shift is travelling to or from another depot. The value of q_3 is either true (1) or false (0).

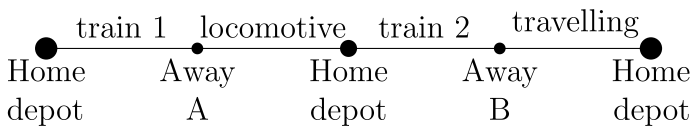

Figure 3.12: A shift with four tasks that has a visit to the home depot in the middle and can thus be split into two parts: the first part is the trip to A and back and the second part is the trip to B and back. For example, if train 1 is cancelled, the first part is cancelled

Based on the three criteria, we assign all shifts to one of four priority classes:

$\mathbf{C_1}$ Extra preference for regular drivers

$\mathbf{C_2}$ No preference between extra and regular drivers

$\mathbf{C_3}$ Extra preference for extra drivers

$\mathbf{C_4}$ Always to extra drivers

Four was considered a suitable number of priority classes by planners, but the number of classes can easily be increased for future uses of the model. The fixed amount of work hours from shifts and available work hours from drivers at a depot mean that not all shifts can fit into the rosters of regular drivers. While some of the shifts must be left as residue shifts, we do not intend to increase residue shift work, so C_4 should be defined in such a way that work hours of C_4 shifts are significantly less than the required residue shift work time. In practice, the number of shifts with properties to belong to this class is always very small, so this does not become a problem. To help planners make assignment examples for this thesis, we divide the probabilities of q_1 and q_2 into four classes: very low, low, medium and high. Table 3.1 presents the numerical thresholds for the probability classes used in this thesis. We chose the thresholds based on interviews with planners and analysis of cancellation probabilities of shifts. While this thesis uses identical thresholds for q_1 and q_2, the thresholds can be different.

Table 3.1: Thresholds of probability classes of whole and partial cancellations

Criterion	Very low	Low	Medium	High
q_1	0	0.1	0.15	0.3
q_2	0	0.1	0.15	0.3

We prioritized clarity, simplicity and adjustability when choosing the method of sorting, because the method needs to be understandable for planners and easy to modify if the optimization model needs to be developed further. We found methods based on utility functions too complex and unintuitive for this project and some of the characteristics of ELECTRE TRI, such as veto evaluations, were unnecessary. Ultimately, we chose a model based on decision rules as the best alternative because of its intuitiveness. Based on discussions with planners and practical experiences from rostering, the following statements were obtained guide the sorting:

1. Only shifts with high probability of whole cancellations should belong to C_4

2. Only shifts with very low probability of whole and partial cancellations and no adaptability should belong to C_1

3. Adaptability makes little to no difference between C_2, C_3 and C_4

Statements 1 and 2 specify C_4 and C_1, respectively. In order to differentiate between C_2 and C_3, we obtained the following assignment examples from planners:

Table 3.2: Shift assignment examples

q_1	q_2	Class
medium	high	$C3$
low	high	$C3$
medium	medium	$C3$
low	medium	$C2$
low	low	$C2$

Because the number of possible combinations is very small, the assignment examples are enough to determine decision rules without inconsistencies. The algorithm to obtain the class of a shift based on the decision rules is presented as a flowchart in Figure 3.13.

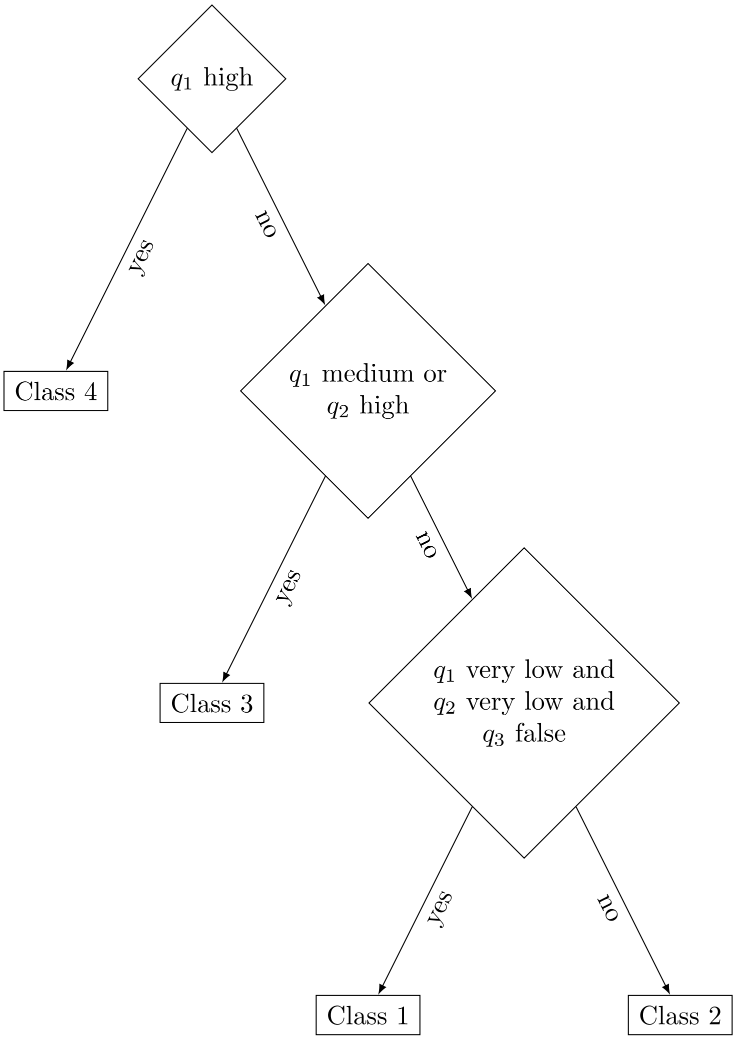

Figure 3.13: The algorithm for determining shift certainty class

In the original model by Porokka (2017), the main goal is to minimize the cost of artificial working time of residue shifts in roster R:

$$C_e(R) = T_\alpha,$$

where T_α is the total artificial working hours of residue shifts. This is equivalent to the sum of artificial working hours of all residue shifts:

$$C_e(R) = \sum_{i=1}^{n} t_\alpha(i), \tag{3.2}$$

where n is the number of residue shifts and $t_\alpha(i)$ is the artificial working time of shift i. For general availability, the artificial working time of a shift should be adjusted based on the priority class of the shift. We make the adjustment by multiplication with weight term $w_\alpha(i)$, whereafter the cost function becomes

$$C_e(R) = \sum_{i=1}^{n} w_\alpha(i) t_\alpha(i) \tag{3.3}$$

If $w_\alpha(i) = 1 \ \forall i$, the cost function is identical to 3.2.

The adjustments to the artificial work times of shifts should not significantly impact the total planned working time of residue shifts after rostering. Instead, the goal is to change what kind of shifts are left as residue shifts. However, small increases to the total planned working time of residue shifts are acceptable when the availability of extra drivers is increased. The class C_2 is neutral in relation to artificial working time adjustments. The class C_1 should cost more than its artificial working time, while class C_3 should cost less than its artificial working time. The weight for class C_4 should be negative and have a large absolute value to ensure they go to extra drivers. Otherwise, the model may prioritize compactness of rosters over availability. As the number of C_4 shifts is small compared to required extra shifts, the negative weight does not increase the work hours of residue shifts.

$$w_\alpha(i) = \begin{cases} w_\alpha(C_1) > 1 & \text{if i} \in C_1 \\ w_\alpha(C_2) = 1 & \text{if i} \in C_2 \\ w_\alpha(C_3) < 1 & \text{if i} \in C_3 \\ w_\alpha(C_4) < 0 < w_\alpha(C_3) & \text{if i} \in C_4 \end{cases}$$

The weight terms can be seen as trade-offs between shifts of different classes. For example, C_3 shifts need only have $w_\alpha(C_3)$ times the artificial working time as C_2 shifts in order to have the same cost. Thus, if a C_1 shift needs to have half the working time of a C_2 shift to be equivalent in value as a residue shift, $w_\alpha(C_1)$ should be 1.5. C_4 shifts have trade-offs only with the other objective of the model, compactness of rosters. Experimenting with different weight values and analyzing the results is needed, because of the complexity of the model.

3.2.3 Daily availability

Optimal general availability is not enough to minimize overtime work, because it does not take into account how many extra drivers are required for extra shifts on each day. Extra drivers should be evenly available for forecast extra shifts. The daily distribution of extra shifts can be hard to predict and can vary from planning period to another. Nevertheless, some weekdays can be more prone to extra shifts than others. If historical data is available, then it should be taken into account during rostering.

Like shifts in general, extra shifts can start and end at any time of day. In addition, overnight extra shifts are quite common. As the precise information of extra shifts, such as start time and duration, is by definition not known in advance, the numerous constraints on shift allocation (3.1.1) make modeling of daily availability highly complex. We make the following assumptions and simplifications to enable modeling:

- Each extra driver is available for one shift each day if not absent
- Overnight shifts are allocated to both start and end days based on the proportion of duration before and after midnight
- Shifts that are not overnight shifts are allocated fully to the start day

Some residue shifts are cancelled and do not need extra drivers. The expected number of residue shifts after cancellations on day i is

$$E[n_r(i)] = n_t(i) - \sum_{j=1}^{n_t(i)} q_1(j),$$

where n_t is the total number of residue shifts and $q_1(j)$ is the probability of total cancellation of residue shift j (3.1). The total number of required drivers on each day is the sum of absences, forecast residue shifts and extra shifts. It is possible that the highest number of required drivers cannot be lowered in cases while a high number of required drivers on some other day could be lowered. Therefore, simply minimizing the peak could have significant shortcomings. We take into account the whole distribution by minimizing the sum of the squared number of required drivers each day. The cost function of daily availability is then defined as

$$C_d(R) = \sum_{i=1}^{21} (n_a(i) + E[n_r(i)] + E[n_e(i)])^2 \tag{3.4}$$

$$= \sum_{i=1}^{21} (n_a(i) + E[n_s(i)])^2 \tag{3.5}$$

where R is the roster in question, n_a is the number of absent extra drivers, n_r is the forecast number of residue shifts after cancellations, n_e is the forecast number of extra shifts and n_s is the total number of expected shifts for extra drivers. The total daily cost should be neutral to different sorts of overnight shifts. For example, a shift with half of duration on day 1 and half on day 2 is allocated $\sqrt{\frac{1}{2}}$ times to days 1 and 2. Figures 3.14 and 3.15 present two examples of daily availability. The sum of total required number of extra driver days- during the period is the same in the examples (50), but the cost of daily availability is 148 in Figure 3.14 and only 126 in Figure 3.15.

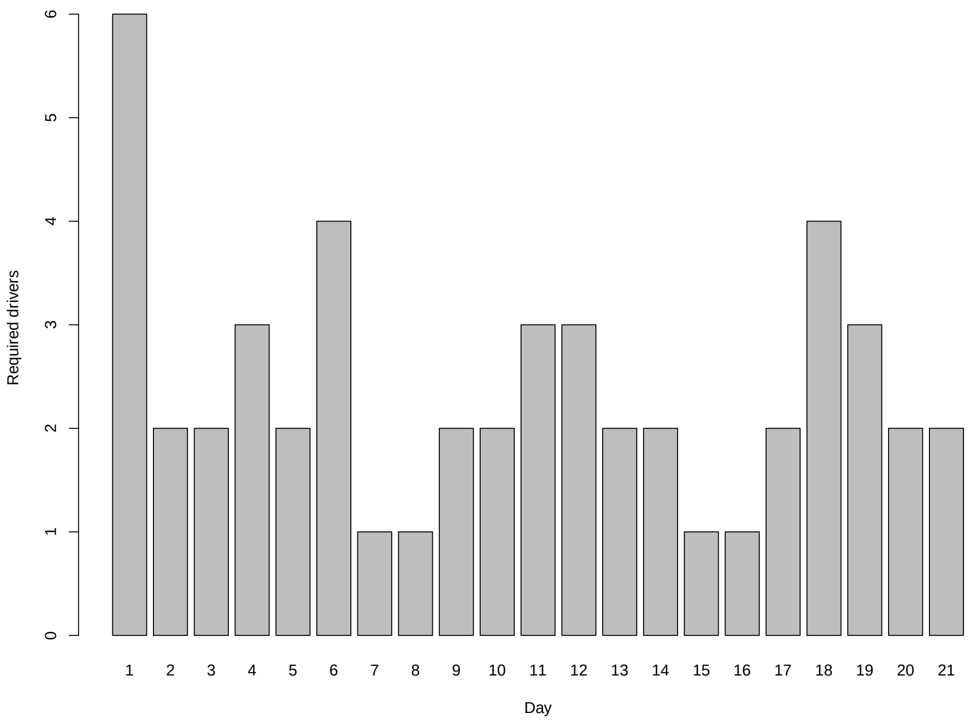

Figure 3.14: Example of a relatively uneven distribution of daily availability

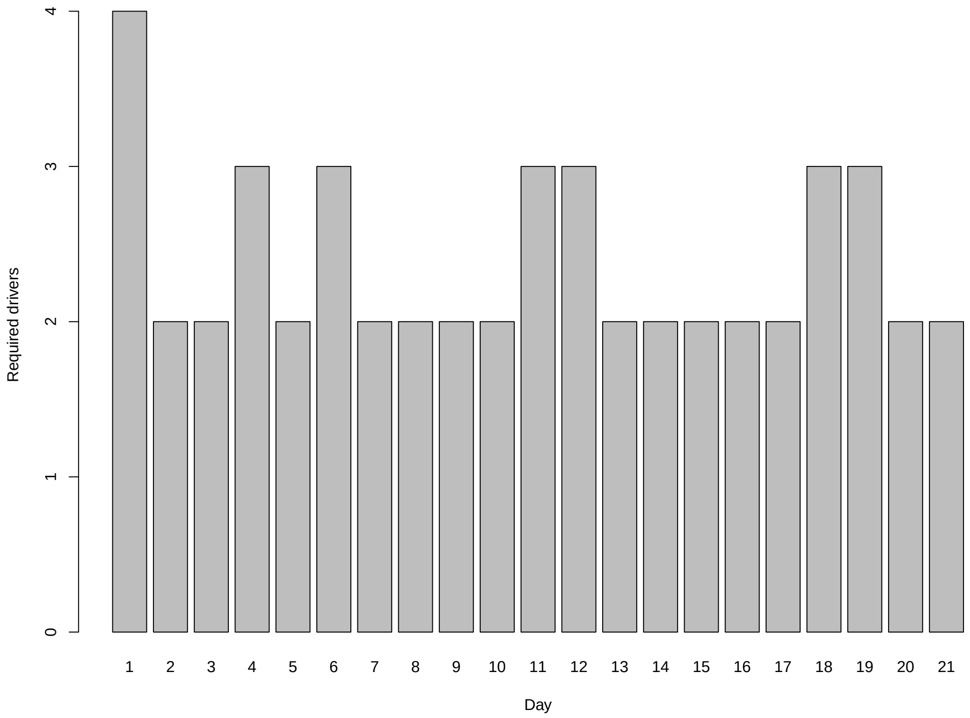

Figure 3.15: Example of a relatively even distribution of daily availability

In the thesis by Porokka (2017), the total cost of a roster is simply the sum of the costs of roster of an individual driver. This allows for straightforward calculations of total cost while removing and inserting shifts in Algorithm 1. In order to avoid having to recalculate daily availability of the roster each time a shift is to be inserted to a driver, we reformulate the cost function of daily availability (3.4) to be similar in form to general availability (3.3).

Each shift assigned to an extra driver has a cost based on the artificial work time and the certainty class of the shift. If a cost based on the date of the shift is added, the solution algorithm should even out the distribution of residue shifts in regard to absences and expected extra shifts. The method of determining the daily cost of shift is presented in figure 3.16. The modified cost function of daily availability is

$$C_d(R) = \sum_{i=1}^{n} w_\beta(i), \tag{3.6}$$

where n is the total number of residue shifts and $w_\beta(i)$ is the weight corresponding to how much availability is required on the day of shift i.

The importance of daily availability depends on the probability of overtime work. If the expected utilization rate of extra drivers is very low, no overtime work is needed regardless of daily availability. In such cases, daily availability should be given lower priority. In addition, if the utilization rate of extra drivers is very high from residue shifts alone, all extra shifts will require overtime work.

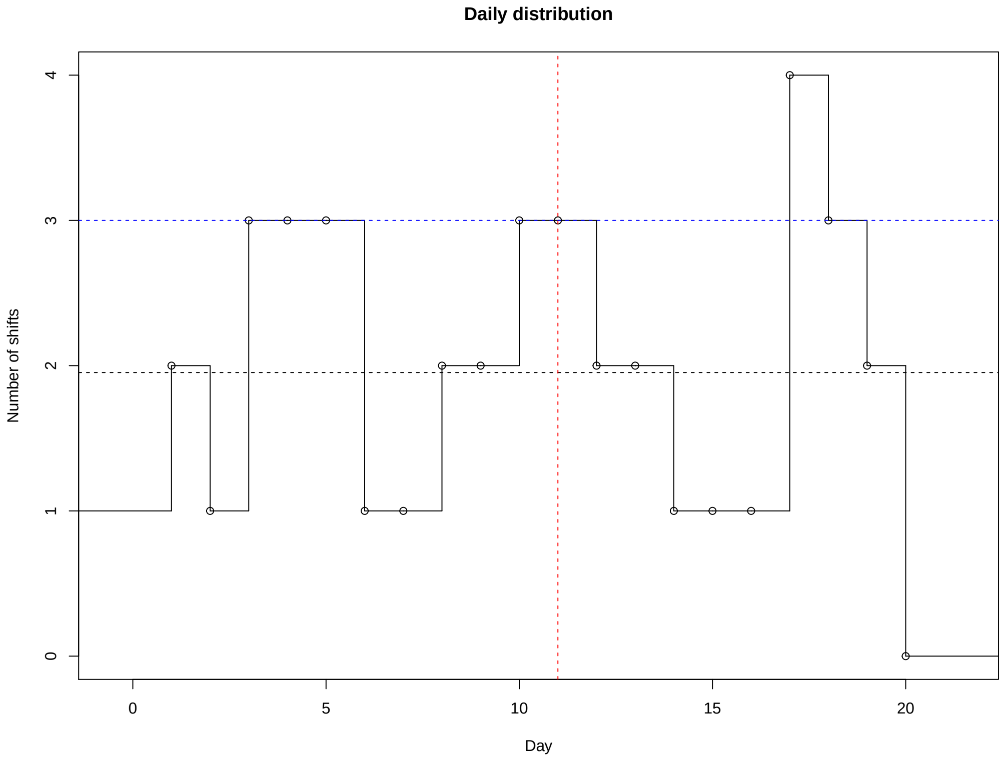

Figure 3.16: Example of a sum distribution of expected extra shifts and residue shifts after cancellations. The forecast number of shifts (blue line) for the day of the shift in question (red line) in relation to the average level of forecast shifts (dashed black line) determines the daily weight of a shift. In this example, the weight of the shift would be around $3/2 = 1.5$. The highest possible weight would be $4/2 = 2$ on day 17

3.2.4 Objective function

The cost functions of both general availability (3.3) and daily availability (3.6) should be minimized. However, the functions may not be of equal importance and their importance might depend on the depot in question. Therefore, we define total availability of Roster R as a weighted sum of the cost functions:

$$f_a(R) = W_e C_e(R) + W_d C_d(R),$$

where W_e and W_d are weights capturing the importance of general and daily availability, respectively. The costs of compactness (f_c) are the same as those described by Porokka (2017). Thus, the total objective is

$$f(R) = f_c(R) + f_a(R).$$

Chapter 4

Simulation

4.1 Implementation

To find out how well the model will perform, we implement a simulation of operational planning. The simulation environment is implemented in C++ and R languages and is fully integrated to the rostering implementation (Porokka, 2017). The goal is to measure how well rosters could theoretically perform after the planning process depending on the optimization model in use. Simulations of disruptions such as extra shifts during operational planning are often done on a day-to-day basis (Ingels et al., 2015 and 2017). This would be suitable for VR as well, as the cancellation of trains and other disruptions often happen quite late. However, we simplify the simulation to consider all days of the planning period simultaneously in order to decrease required computation significantly and to have less need to create completely new software. Consequently, the simulation assumes all disruptions to be known at the start of the period. This means that the simulated use of extra drivers is somewhat more efficient than is possible in practice, regardless of the optimization model in use. On the other hand, as we perform the simulation on a single depot, the possibility to change the depot of shift is ignored, which can decrease efficiency to some extent. We leave out disruption types that are not considered in the optimization model, such as sick leaves. This decreases unnecessary variability between different simulation runs. As the forecast distribution of sick leaves would be more or less uniform, simulating sick leaves would not differentiate the different versions of the model. In short, we make the following simplifications or assumptions for computational purposes:

- All disruptions known at the start of operational planning
- No sick leaves

- Same extra shifts in all simulation runs

The demand and arrival uncertainties in the simulation are the cancellation of trains/shifts and the precise information of extra shifts (arrival and duration).

We use a large depot with relatively many freight trains prone to cancellations as our test depot. Table 4.1 presents statistics on the shifts of the depot. The forecast utilization rate of drivers after extra shifts is around 100 %. The shift data for both planned and extra shifts we use is from a single representative past period of the depot. For realistic train cancellations, the cancellation probabilities of freight trains are estimated using 6 months of historical cancellation data.

Table 4.1: Shift attributes in the simulation depot

Attribute	Planned	Extra
Shifts	895	63
Night shifts	537	26
Night hours	3145	182
Artificial work hours	9834	660
Class 1 shifts	96	-
Class 2 shifts	774	-
Class 3 shifts	17	-
Class 4 shifts	8	-

Each run of the simulation consists of four steps: rostering, days off planning, the calculation of train and shift cancellations and minimization of overtime work. The rostering step uses Algorithm 1 to improve the solution after a feasible roster has been found. Rostering has stochastic elements, such as choosing the shifts to be removed while improving a solution, which is why each simulation run produces a unique roster. After rostering, we plan the days off for extra drivers automatically. Section 4.2 presents the method of automatic days off planning.

The actual simulation of disruptions can start after extra drivers have their days off planned. The simulated cancellations of individual trains follow a binomial distribution based on their cancellation probability. We calculate shift cancellations for each simulation run from individual train cancellations using the same logic as with shift priorities in section 3.1.3. We then remove the cancelled shifts from residue shifts (and cancelled parts from partly cancelled shifts). In the final step, we introduce the extra shifts and seek to maximize the working hours of extra drivers using both residue shifts from rostering and extra shifts. The shifts that cannot be covered by extra drivers are assumed to be left to overtime workers. To have a clearer view of the properties of the different versions of the model, we do not try to minimize the unevenness of overtime shifts at this stage. The problem of minimizing overtime work is very similar to the original solution improvement of rostering in section 3.2.1. The difference is that here extra drivers are equivalent to regular drivers in the original problem, while overtime workers are equivalent to extra drivers. Thus, we can use Algorithm 1 and choose the artificial work hours of overtime workers as objective function.

We compare three different versions of the optimization model. The first version is the original model, which ignores uncertainties. The second version of the model uses priority classes to achieve general availability (Equation 3.3), while the third version uses both priority classes and daily weights for shifts based on forecast extra shift distribution (Equation 3.6) to achieve general and daily availability. The models are identical on other aspects. We measure the performance of the models by two indicators: the mean amount of overtime work (general availability) and the mean peak number of simultaneous overtime shifts (daily availability) after the simulations.

We simulate rostering and minimization of overtime with computational times of 20 and 10 seconds, respectively. Based on practical experiences, these seem to be enough for good-enough solutions. The total computational time of each run is around 35 seconds using a computer with 2.3 GHz Intel Core i5-5300U processor and 8 GB of RAM.

4.2 Automatic days off planning

For the purposes of this simulation, we develop a multicriteria optimization model and heuristic for choosing the days off for extra drivers after rostering automatically but with similar results to manual choosing. Each extra driver has one weekend off, which is preselected before the simulation. The goal is to choose the rest of the days off so that there is evenly room for extra shifts throughout the planning period, while also considering the preferences of the extra drivers. Extra drivers generally wish to avoid single days off and work periods (between days off) of less than three days, which are also often inefficient. In the future, the model should also look at the forecast distribution of extra shifts in order to properly optimize the days off.

We use a typical additive model for the objective function. To even out the number of extra drivers that are “taken” (either a day off or a residue shift) each day, we use the least squares of taken drivers. For example, minimizing the maximum number of taken drivers would not work in cases where there is simply too many residue shifts for a certain day. Let a be the cost of a single day off for an extra driver, while b and c are the costs of single- and double-day work periods, respectively. We define the objective function for a days off plan P automatically as

$$f(P) = a \cdot n_o + b \cdot n_s + c \cdot n_d + \sum_{i=1}^{21} n_{t,i}^2,$$

where n_o is the total number of single days off, n_s is the total number of single-day work periods, n_d is the total number of double-day work periods and $n_{t,i}$ is the number of taken extra drivers on day i. Optimizing the planning of days off is much simpler than rostering, but there are many similarities. Therefore, we propose a simple heuristic inspired by the solution improvement Algorithm 1 from section 3.2.1. that starts with a feasible plan.

Figure 4.1 presents an example of automatically planned days off for extra drivers along with residue shifts from rostering. The main purpose of optimizing the days off is enabling simulation without manual input between runs. Therefore, we determine suitable costs for the objective function by simply looking at the resulting graphical plans iteratively. The problem is simple enough that Algorithm 2 can find a good enough solution very quickly, despite it consisting mostly of trying at random. We find removing those days off that have currently the most taken drivers to be an intuitive and satisfactory removal strategy for Algorithm 2.

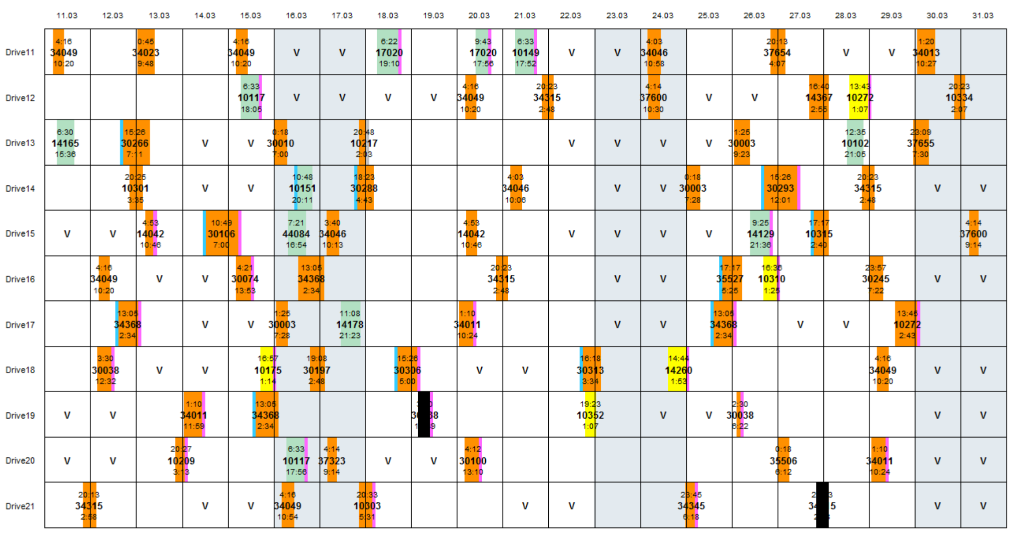

Figure 4.1: Days marked as "V" are the automatically planned days off for the extra drivers based on residue shifts from rostering

Algorithm 2 Automatic planning of days off

feasible plan P
$P_{best} \leftarrow P$
repeat
  $P' \leftarrow P$
  choose a random extra driver d
  choose removal strategy s
  remove $q \in \{q_{min}, ..., q_{max}\}$ days off from driver d using strategy s
  insert q number of removed days off back to driver d to randomly chosen days
  if insertion did not result in a feasible solution **then**
    $P \leftarrow P'$
  else if $(f(P) > f(P'))$ **then**
    $P \leftarrow P'$
  else if $f(P) < f(P_{best})$ **then**
    $P_{best} \leftarrow P$
  end if
until the stopping criterion is met
P_{best} contains the best feasible solution found

4.3 Results

We have chosen the number of simulation runs to be 100. The graphical representations of the cumulative means of indicators overtime work from simulations in Figures 4.2, 4.3, 4.4, 4.5, 4.6 and 4.7 indicate that 100 runs is enough for convergence of mean for the purposes of this simulation. Table 4.2 presents the numerical results of the simulations.

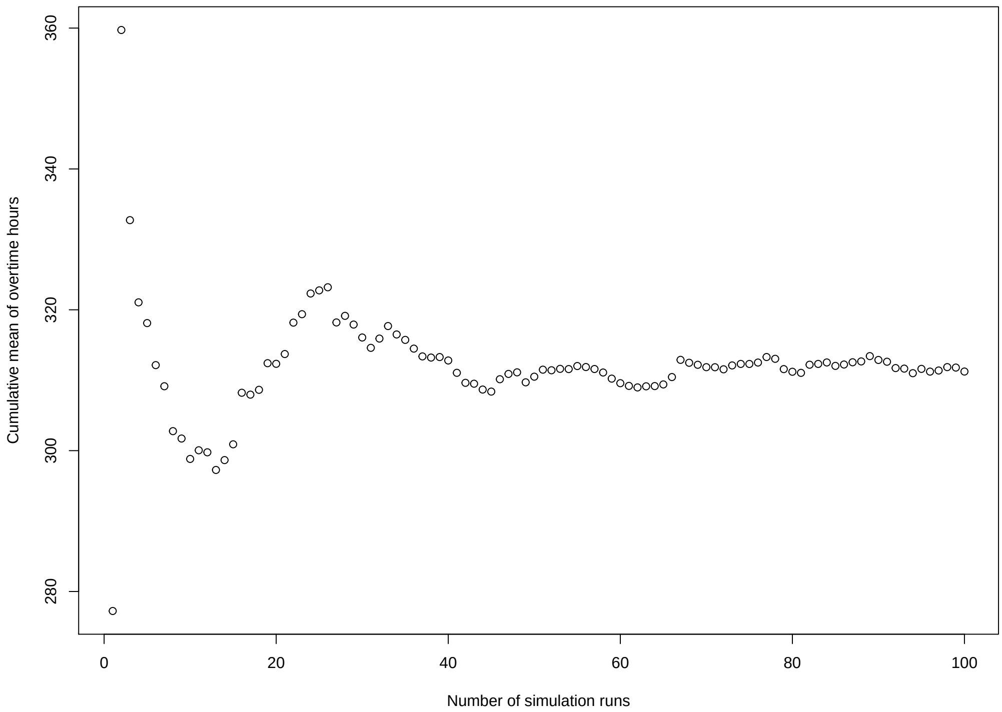

Figure 4.2: The convergence of cumulative mean of overtime work using model 1

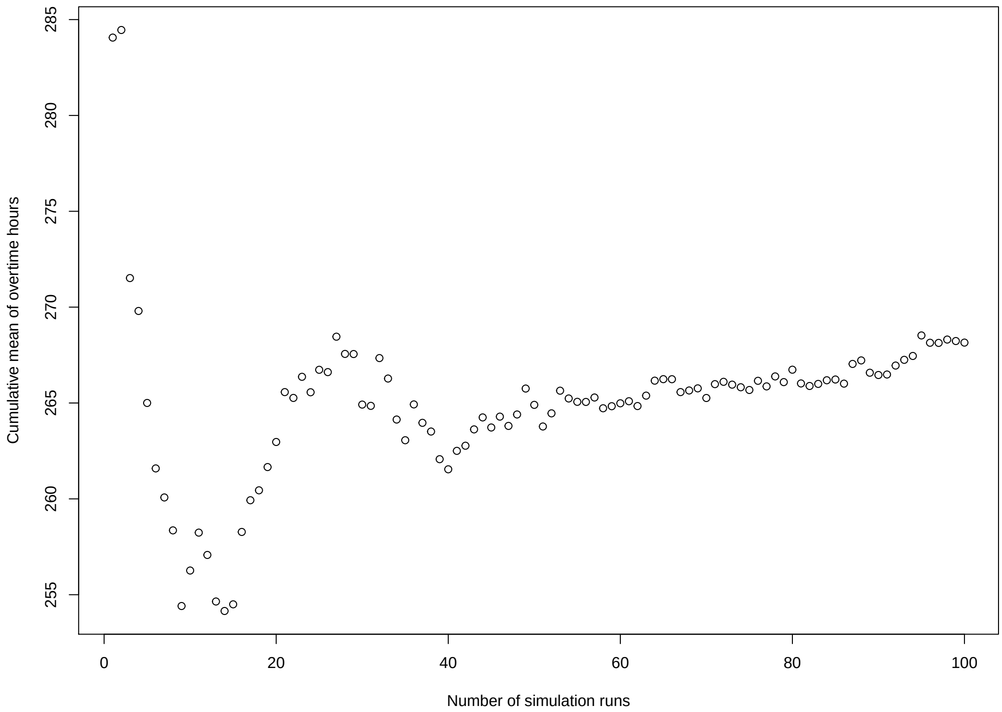

Figure 4.3: The convergence of cumulative mean of overtime work using model 2

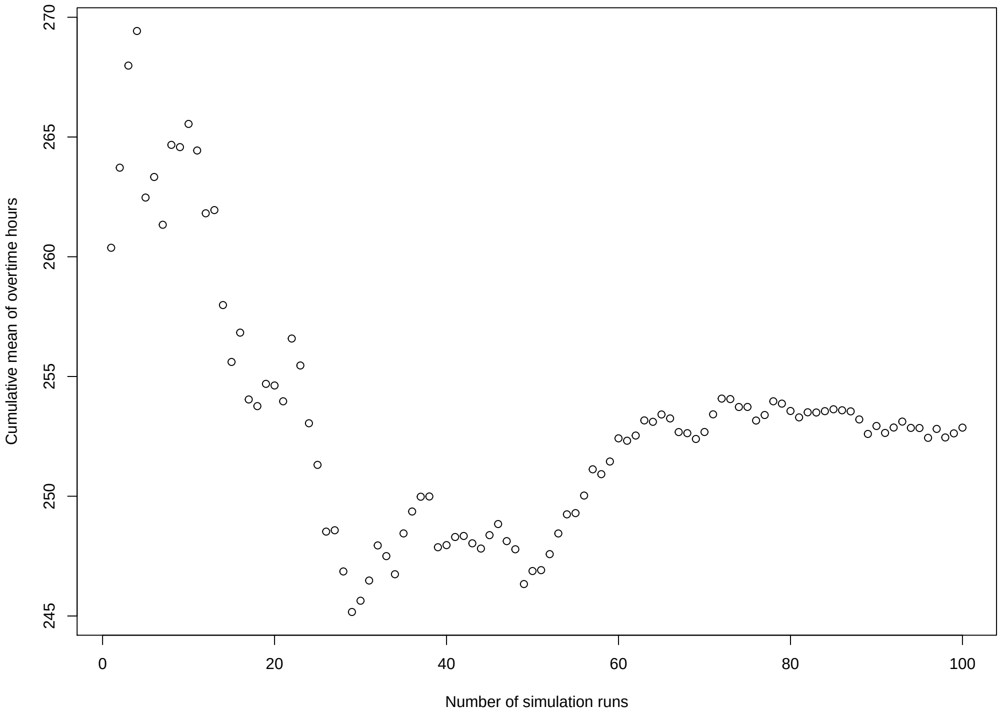

Figure 4.4: The convergence of cumulative mean of overtime work using model 3

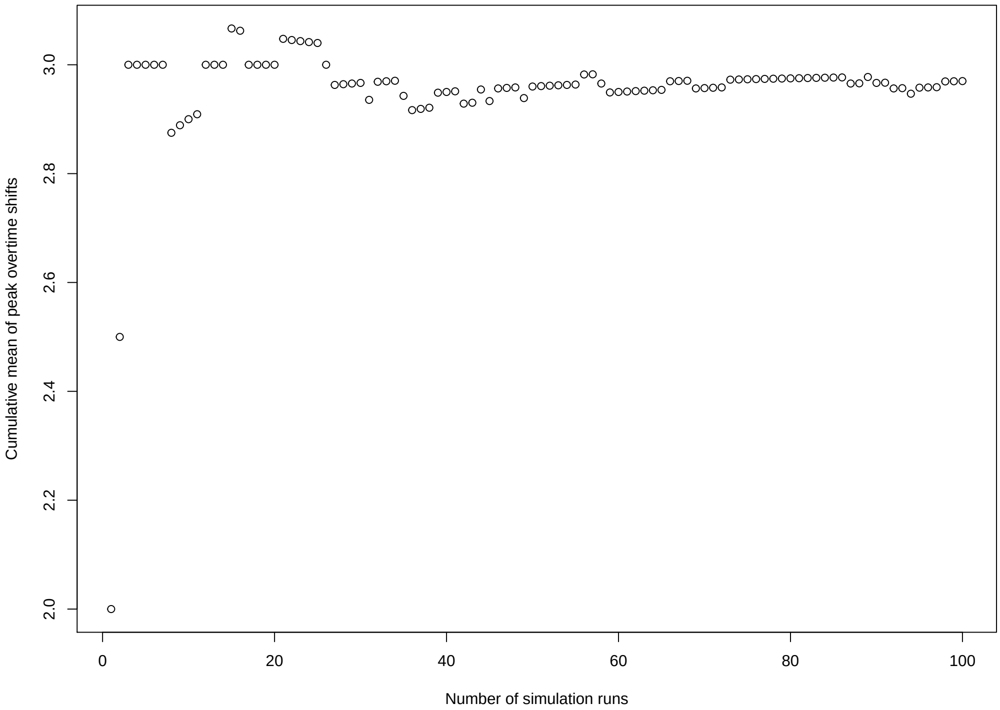

Figure 4.5: The convergence of cumulative mean of overtime peak using model 1

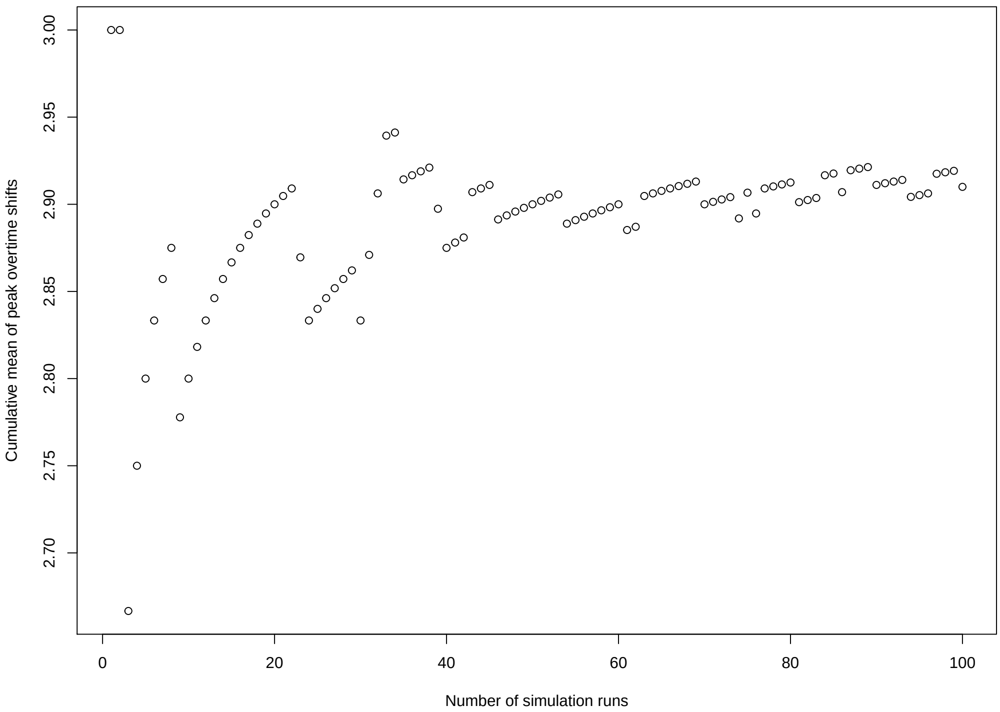

Figure 4.6: The convergence of cumulative mean of overtime peak using model 2

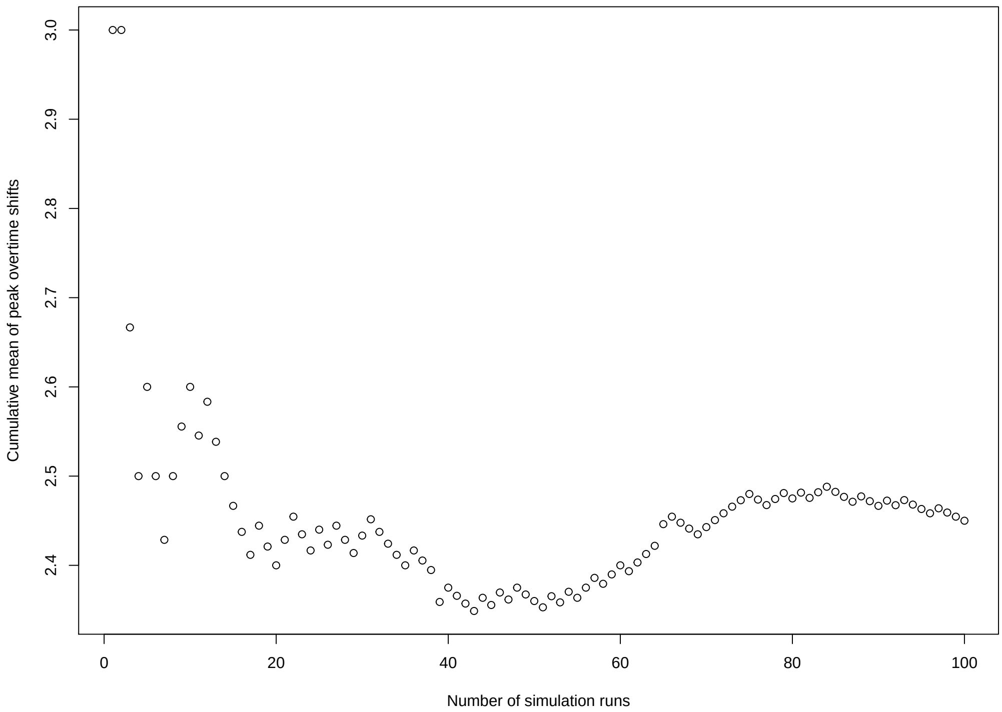

Figure 4.7: The convergence of cumulative mean of overtime peak using model 3

Table 4.2: Simulation results for a large depot (means)

Model	Overtime work (h)	Overtime peak
model 1	311	3.0
model 2	268	2.9
model 3	252	2.5

Compared to model 1, model 2 makes a clear decrease to overtime work (over 40 hours, 14 %), while there is little to no difference in overtime peak. As model 2 does not take extra shift distribution into account, these results are quite expected and suggest that the model works as intended. While 14 % decrease may seem modest, the proportion of class 3 and 4 shifts is quite small, so a radical decrease is not possible in this context.

Model 3 performs slightly better than model 2 on both indicators, which confirms that the most developed model is the best performing alternative for the two indicators we have chosen.

Chapter 5

Conclusion

The goal of this thesis was to improve an existing roster optimization model to account for uncertain freight trains. Because of these uncertainties, some trains will be cancelled after planning while some are not yet known at the time of planning. Using forecast information on these types of trains during roster planning could improve the use of flexible train driver resource, which decreases overtime costs.

In contrast to the original model, the modified model does not simply minimize the planned artificial work time of residue shifts but also prioritized shifts that are not prone to cancellations for regular drivers. We used multicriteria sorting to determine the priority of shifts. In addition, we increased the priority of shifts on days when there are more forecast extra shifts. Both of these modifications were intended to increase the availability of extra drivers.

Based on the results of a simulation study, we managed to create a modification to the optimization model, which succeeds in the goal of the thesis. However, more studies need to be performed to verify the strength of improvements and to see if models with different approaches could perform even better. In addition, the accuracy of forecasting needs attention for good performance in real-life planning.

Bibliography

[1] Azibi, R., and Vanderpooten, D. Construction of rule-based assignment models. *European Journal of Operations Research 138*, 2 (2002), 274–293.

[2] Burke, E. K., Li, J., and Qu, R. A hybrid model of integer programming and variable neighbourhood search for highly-constrained nurse rostering problems. *European Journal of Operational Research 203*, 2 (2010), 484–493.

[3] Cardinal, J. S.-L., Mousseau, V., and Zheng, J. An Application of Constrained Multicriteria Sorting to Student Selection. In *Portfolio Decision Analysis*, A. Salo, J. Keisler, and A. Morton, Eds., vol. 162. International Series in Operations Research & Management Science, 2011, pp. 213–240.

[4] Chen, Y., Kilgour, D. M., and Hipel, K. W. A Decision Rule Aggregation Approach to Multiple Criteria-Multiple Participant Sorting. *Group Decision and Negotiation 21*, 5 (2012), 727–745.

[5] Damart, S., Dias, L. C., and Mousseau, V. Supporting groups in sorting decisions: Methodology and use of a multi-criteria aggregation/disaggregation DSS. *Decision Support Systems 43*, 4 (2007), 1464–1475.

[6] Doumpos, M., Marinakis, Y., Marinaki, M., and Zopounidis, C. An evolutionary approach to construction of outranking models for multicriteria classification: The case of the ELECTRE TRI method. *European Journal of Operational Research 199*, 2 (2009), 496–505.

[7] Ernst, A., Jiang, H., Krishnamoor, M., and Sier, D. Staff scheduling and rostering: A review of applications, methods and models. *European Journal of Operational Research 150*, 1 (2004), 3–27.

[8] Gendreau, M., Ferland, J., Gendron, B., Hail, N., Jaumard, B., Lapierre, S., Pesant, G., and Soriano, P. Physician Scheduling in Emergency Rooms. *International Conference on the Practice and Theory of Automated Timetabling 6* (2006), 53–66.

[9] Ingels, J., and Maenhout, B. The impact of reserve duties on the robustness of a personnel shift roster: An empirical investigation. *Computers & Operations Research 61* (2015), 153–169.

[10] Ingels, J., and Maenhout, B. Employee substitutability as a tool to improve the robustness in personnel scheduling. *OR Spetcrum 39*, 3 (2017), 623–658.

[11] Jacquet-Lagreze, E., and Siskos, J. Assessing a set of additive utility functions for multicriteria decision-making, the UTA method. *European Journal of Operational Research 10*, 2 (1982), 151–164.

[12] Köksalan, M., and Özpeynirci, S. B. An interactive sorting method for additive utility functions. *Computers & Operations Research 36*, 9 (2009), 2565–2572.

[13] Lin, D.-Y., and Tsai, M.-R. Integrated Crew Scheduling and Roster Problem for Trainmasters of Passenger Railway Transportation. *IEEE Access 7* (2019), 27362 – 27375.

[14] Parisio, A., and Jones, C. N. A two-stage stochastic programming approach to employee scheduling in retail outlets with uncertain demand. *Omega 53* (2015), 97–103.

[15] Pawlak, Z. Rough Sets. *International Journal of Computer & Information Sciences 11*, 5 (1982), 341–356.

[16] Porokka, A. Train driver rostering in Finland considering driver satisfaction. Master's thesis, School of Science, Aalto University, 2017.

[17] Punnakitikashem, P. *Integrated Nurse Staffing And Assignment Under Uncertainty.* PhD thesis, University of Texas at Arlington, 2007.

[18] Ropke, S., and Pisinger, D. An Adaptive Large Neighborhood Search Heuristic for the Pickup and Delivery Problem with Time Windows. *Transportation Science 40*, 4 (2006), 455–472.

[19] Schaefer, A. J., Johnson, E. L., Kleywegt, A. J., and Nemhauser, G. L. Airline Crew Scheduling Under Uncertainty. *Transportation Science 39*, 3 (2005), 289–441.

[20] Valdes, V., and Andres, V. *Integrating Crew Scheduling and Rostering Problems.* PhD thesis, University of Bologna, 2010.

[21] Van den Bergh, J., Beliën, J., Bruecker, P. D., Demeulemeester, E., and Boeck, L. D. Personnel scheduling: A literature review. *European Journal of Operational Research 226*, 3 (2013), 367–385.

[22] VR Group. VR Group. `https://www.vrgroup.fi/en/`, 2019. Accessed: 2019-06-11.

[23] Xie, L., and Suhl, L. Cyclic and non-cyclic crew rostering problems in public bus transit. *OR Spectrum 37*, 1 (2015), 99–136.

[24] Yevseyeva, I. *Solving Classification Problems with Multicriteria Decision Aiding Approaches.* PhD thesis, University of Jyväskylä, 2007.

[25] Yu, W. E. S. *Aide multicritère à la décision dans le cadre de la problématique du tri : concepts, méthodes et applications.* PhD thesis, Universite de Paris-Dauphine, 1992.

[26] Zopounidis, C., and Doumpos, M. PREFDIS: a multicriteria decision support system for sorting decision problems. *Computers & Operations Research 27*, 7-8 (2000), 779–797.

[27] Zopounidis, C., and Doumpos, M. Multicriteria classification and sorting methods: A literature review. *European Journal of Operational Research 138*, 2 (2002), 229–246.

[28] Örmeci, E. L., Salman, F. S., and Yücel, E. Staff rostering in call centers providing employee transportation. *Omega 403* (2014), 41–53.