UNIVERSITY OF TWENTE.

merem

MEDISCHE REVALIDATIE

Master Thesis Applied Mathematics

Designing a DLA-CFA policy for patient admission within a rehabilitation clinic's multi-disciplinary, multi-appointment environment

Renske Janssen

Stochastic Operations Research
University of Twente

December 2024

Supervisors:
prof. dr. R. Boucherie (UT)
dr. ir. A. Braaksma (UT)
MSc. B. Boezeman (Merem)

Graduation committee:
prof. dr. R. Boucherie (UT)
dr. ir. A. Braaksma (UT)
dr. A. Antoniadis (UT)
MSc. B. Boezeman (Merem)

Preface

This thesis, titled 'Designing a DLA-CFA policy for patient admission within a rehabilitation clinic's multi-disciplinary, multi-appointment environment' was written to complete my master 'Applied Mathematics' at the University of Twente. For this research, I had the opportunity to apply my mathematical skills to a real-world problem, allowing me to engage creatively with the concepts I had mastered. This thesis has transformed into what it is today due to the invaluable support of several remarkable individuals, whom I would like to acknowledge below.

Firstly, I want to thank the rehabilitation clinic *Merem* for the opportunity to perform my thesis there. For this, I want to thank Bauke for welcoming me and ensuring I had all at my disposal during my research. I also have to thank the planners in the planning department. A special thanks to Thijs, who provided me with many life lessons and made me a little bit more of a 'Randstedeling'. Despite the long trip, I always enjoyed traveling to Hilversum weekly.

I also want to thank my supervisors from the University of Twente, Richard and Aleida. Thank you for your insight and ideas, which helped me during my research, and for offering a workplace at the university, which helped me to get in contact with the PhD'ers of the research group SOR. In addition to the theoretical insight they gave me, the daily lunch walks also helped me to enlighten myself from the hard work. Thank you all, especially the people of Zilverling 4006 and Zilverling 4054.

This thesis also marks the end of my studies at the University of Twente. Therefore, I would like to thank all the friends I have made over the years, especially the members of D.W.V. Klein Verzet, with whom I have shared numerous small and big adventures over the years. Furthermore, I want to thank Larissa, with whom I shared my five and a half years of study in Enschede. For the last couple of months, I always looked out for our daily coffee breaks. Also, thanks to my parents and little brother for their unwavering support and encouragement during my study time. Knowing they are always by my side, no matter what path I choose. And thanks to Merel and Dennis for giving me valuable feedback.

Finally, a very special thanks go to Martijn for always being there for me, for listening if I had to blow off some steam, and for distracting me when I should not be thinking about my thesis.

Renske Janssen,
Enschede, December 2024

Abstract

In rehabilitation clinics, inpatient and outpatient patients follow treatment with different health professionals, e.g., physiotherapists, occupational therapists, and speech therapists. The type and number of therapy sessions and the length of the total admission or amount of outpatient visits differ per patient. A clinic often has limited capacity, but a fast start of treatment is often crucial to gain the best improvement. We developed a method for an admission policy for outpatients that balances the waiting times and the overuse of capacity. We design a Markov Decision Process (MDP) that models the arrivals and length of stay of both inpatients, who are directly admitted, and outpatients, who first enter the waiting list. We define set decision moments where the patient's length of stay can be increased, resulting in extra unknown demand. We solve the obtained MDP by using a combined Direct Lookahead and Cost Function Approximation policy (DLA-CFA) policy. This policy forecasts the demand resulting from exogenous information for a couple of weeks in advance. We use a stochastic approximation algorithm and a newsvendor model to forecast the demand. We apply the model to a simple multi-appointment but single-discipline setting and a system resembling a multi-disciplinary rehabilitation clinic, and compare the results with a myopic policy without such a forecast. We conclude that our developed method outperforms the myopic policy. We note that the stochastic approximation and newsvendor models achieve similar objective function costs. However, their performance may differ in other assessment factors.

Management samenvatting

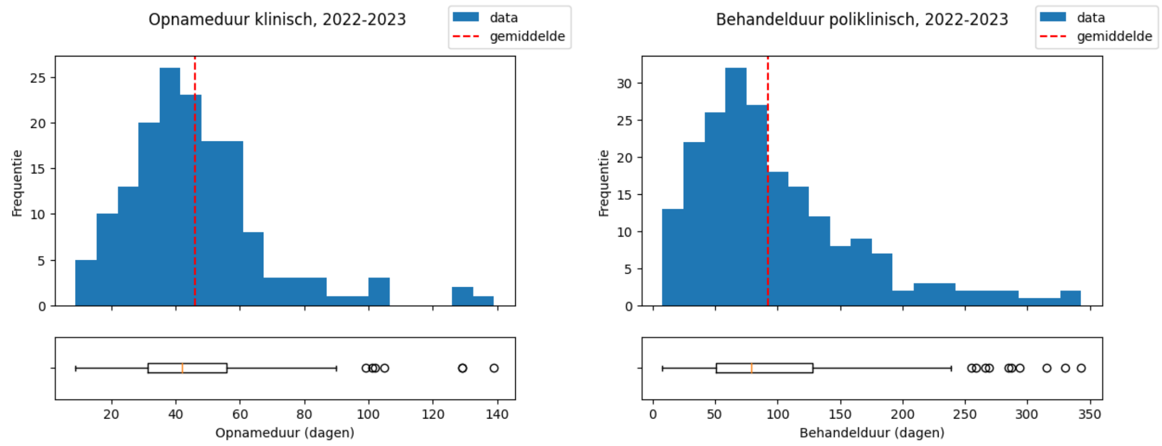
‘Kan zo’n AI niet voor ons plannen?’

Een data-analyse en verbetermodel voor de planning voor de neurologie-chirurgie afdeling

Door Renske Janssen, stagiaire planning
‘Verbeter de planning op de neurologie/chirurgie afdeling van *Merem*’ was de opdracht voor mijn masterscriptie Applied Mathematics. Dit is echter niet eenvoudig. Het gebruik van toepasselijke wiskundige modellen, in de volksmond vaak ‘Artificial Intelligence’ genoemd, werkt het best als de variabiliteit van het model, tot op zekere hoogte, niet al te groot is. Bij de planning van een multidisciplinair systeem met meerdere afspraken per discipline voor één patiënt, zoals dat van de neurologie/chirurgie afdeling, is dit niet het geval. Het grote aandeel telefoontjes, mailtjes en papierwerk dat de planner van de afdeling momenteel verwerkt met mutaties, ziektegevallen en plotselinge nieuwe opnames, maakt dat deze directe handelingen moeilijk in een wiskundig programma kunnen worden gevat. Dit is dan nog afgezien van het sociale aspect waarbij een menselijke planner de behoeftes van zowel patiënten als behandelaars kan inschatten. Aannemende dat mutaties buiten beschouwing gelaten kunnen worden en dat de capaciteit vaststaat, kan wiskunde dan inzicht geven hoe er niet te veel en niet te weinig afspraken worden ingepland?

Data-analyse

Een data-analyse laat zien wat de variabelen in het planningsmodel zijn, en geeft ook een idee van de grootte van het systeem. De data-analyse is gedaan over de jaren 2022 en 2023 en omvat alle patiënten die klinisch en/of poliklinisch in deze jaren onder behandeling waren. De analyse toont dat er per week gemiddeld drie nieuwe patiënten klinisch en vier patiënten poliklinisch (inclusief patiënten die eerst klinisch behandeld werden) beginnen met behandelingen. Maar het gemiddelde zegt niet alles, het aantal opnames kan wekelijks variëren tussen nul en zeven nieuwe klinische patiënten en nul en negen nieuwe poliklinische patiënten.



Figuur 1: Opnameduur voor klinische patiënten (links), en behandelduur voor poliklinische patiënten (rechts).

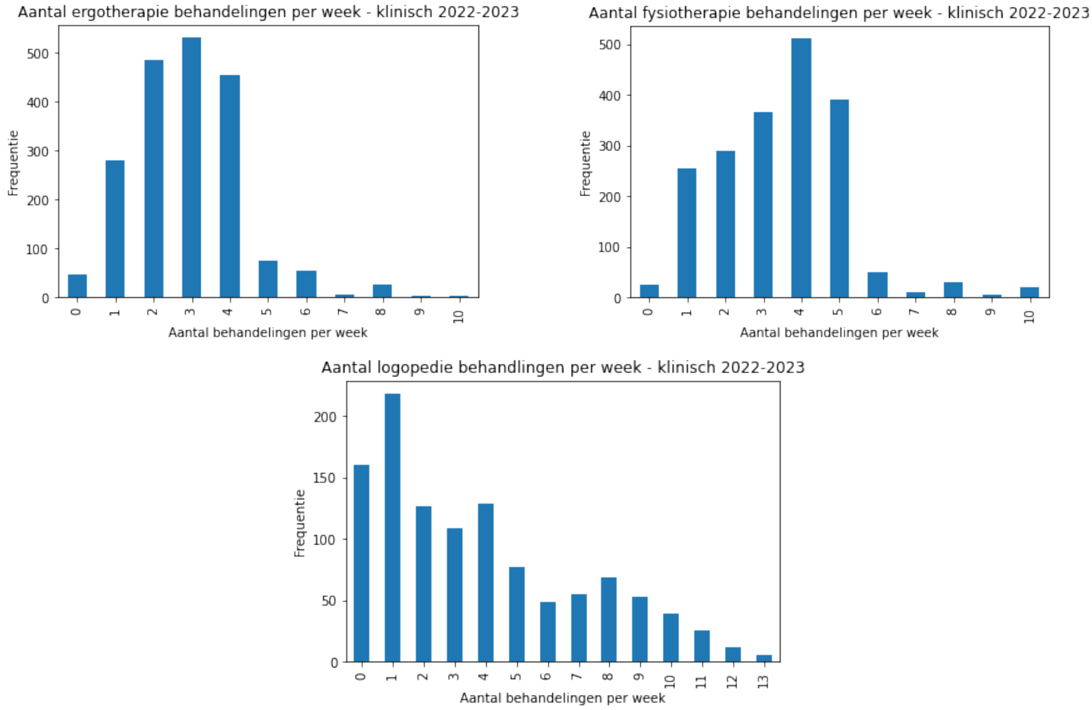
Figuur 1 geeft de lengte van de opnameduur (klinisch) of behandelduur (poliklinisch) weer, waarbij uitschieters naar boven niet zijn meegenomen. De bijbehorende boxplots zijn ook weergegeven. Voor klinische patiënten is de gemiddelde opnameduur 7 weken, waarbij de grootste groep, ongeveer 95%, tussen 2 en 13 weken valt. Voor poliklinische patiënten is de gemiddelde behandelduur 12 weken, waarbij de grootste groep tussen 2 en 35 weken valt.

In de huidige planning worden veel mutaties gedaan, maar de gemiddelde inzet van het aantal behandelingen per discipline is redelijk stabiel. Dit is uitgezet in figuur 2. Gemiddeld heeft een klinische patiënt drie keer ergotherapie en vier keer fysiotherapie per week, daarbij zijn groepslessen niet meegerekend. Er kunnen Voor de logopedie drie groepen worden onderscheiden, namelijk één, vier of acht behandelingen per week.

Poliklinisch, laten de data zien dat patiënten meestal één, en soms twee behandelingen per discipline per week hebben, groepslessen daarbuiten gelaten. Met behulp van deze data-analyse zouden zogenaamde zorgpaden kunnen worden opgesteld, die het plannen vergemakkelijken. Deze worden dan ook gebruikt in het model. Vanwege de wachtlijsten voor deze disciplines zijn de psycholoog en maatschappelijk werk buiten de data-analyse gehouden.

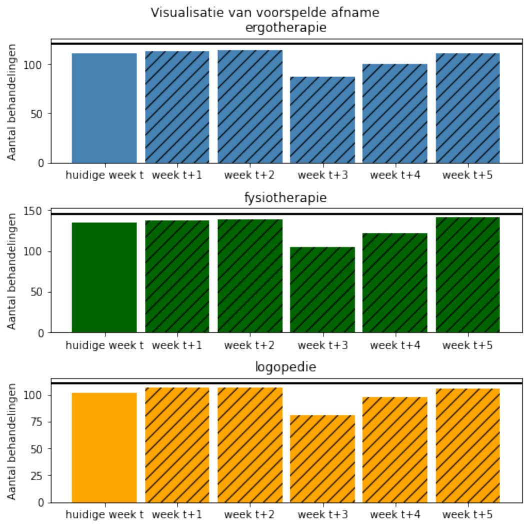
Resultaten van het model

Klinische patiënten worden direct opgenomen, mits er plek is. Poliklinische patiënten komen echter eerst op een wachtlijst. Het model adviseert op wekelijkse basis hoeveel en welke patiënten van deze wachtlijst het best kunnen starten met hun traject. Het maakt gebruik van de data-analyse, met de daarbij bepaalde zorgpaden. Het model maakt een voorspelling voor de komende maand, rekening houdend met nieuwe opnames en eventuele verlenging van de ontslagdatum voor de huidige patiënten. Figuur 3 visualiseert een voorspelling van het model voor een willekeurige week. Op basis hiervan geeft het model een advies voor de planning van opnames. Verdere theoretische details zijn te vinden in de bijbehorende scriptie.



Figuur 2: Distributie van het aantal behandelingen per week per patiënt, voor ergotherapie, fysiotherapie en logopedie.

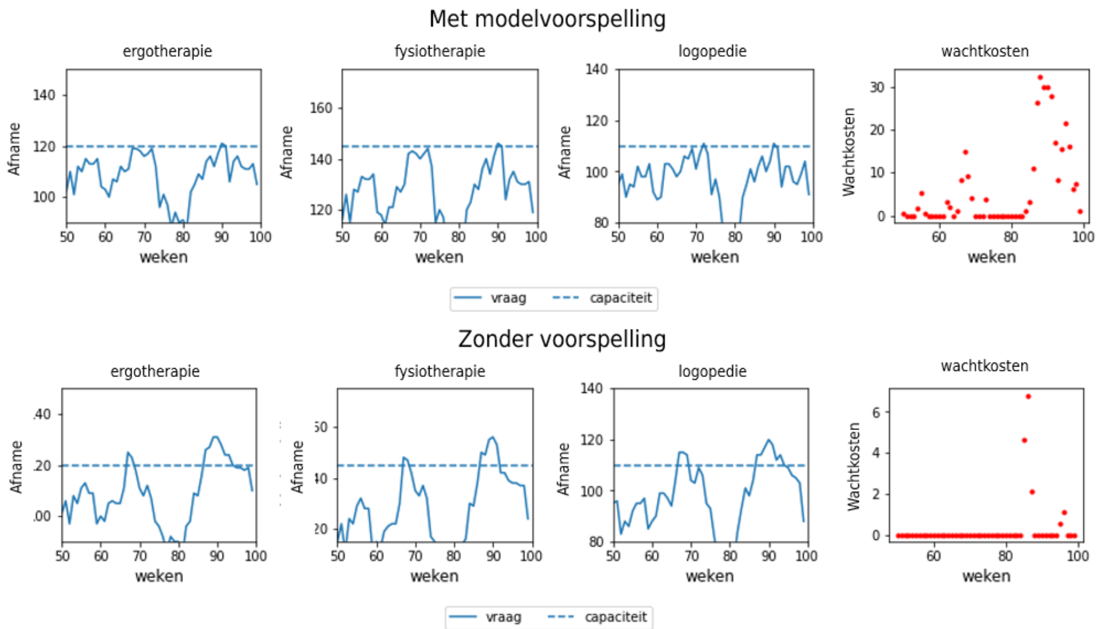
Momenteel is er weinig beleid met betrekking tot de poliklinische wachtlijst. Voor een planner is het beheer van de wachtlijst echter de enige mogelijkheid om de vraag van behandelingen te sturen. De opname van klinische patiënten en het bepalen van het aantal behandelingen ligt namelijk buiten de macht van de planner.



Figuur 3: Visualisatie van de voorspelde vraag voor ergotherapie, fysiotherapie en logopedie, voor een willekeurige week. Voor de huidige week is de vraag bekend, voor de weken daarna is de vraag (deels) voorspeld.

Figuur 4 visualiseert een simulatie van de vraag van behandelingen en de wachttijd van poliklinische patiënten. Hierbij wordt een vergelijking gemaakt tussen het model met een voorspelling en een simpelere versie die geen voorspelling gebruikt. De zogenaamde wachtkosten laten zich vertalen naar een gemiddelde wachttijd van negen dagen, ongeveer equivalent is aan de huidige wachttijd bij *Merem*. Te zien is dat met een goed poliklinisch opnamebeleid de vraag nauwelijks boven de vastgestelde capaciteit komt. Echter valt ook te zien dat de fluctuatie in vraag groot is. Hier valt (wiskundig) weinig aan te doen. Dit verklaart waarom een behandelaar af en toe meer of minder druk kan zijn.

Dus, antwoord gevend op de vraag of plannen beter kan met AI: wiskundige modellen kunnen zeker bijdragen aan de optimalisatie van de planning. Met het gebruik van zorgpaden en het minimaliseren van onverwachte capaciteitswisselingen, kan het model ervoor zorgen dat door slim te starten met poliklinische patiënten, de vraag van behandelingen binnen de capaciteit kan blijven. Het definiëren van zorgpaden is noodzakelijk om een dergelijk model te implementeren, iets wat behandelaren in samenspraak met management zullen moeten bepalen. Vanzelfsprekend betekent dit ook implementatie van dit model in de planningsapplicatie.



Figuur 4: Een simulatie van vraag van behandelingen en wachtkosten met het model met voorspelling (boven) en zonder de voorspelling die het model maakt (onder).

Contents

Chapter 1

Introduction

Operations Research (OR) in healthcare is an emerging field that holds significant importance. Enhancing logistics in this field improves patient care, increases satisfaction, and reduces costs. Many articles have been published on capacity management, patient admission planning, and appointment scheduling, see [1, 2, 3]. However, research frequently focuses on specific situations and is often not applicable to other areas of care. For example, hospitals usually consider only single-appointment and single-discipline planning. However, planning in different healthcare sectors, such as rehabilitation clinics or psychology wards, can be more complicated due to multi-appointment and multi-disciplinary needs. Access times for these sectors are often high, and capacity can usually not be increased easily. Access time refers to the duration from entering the waiting list to the first appointment. In this research, we consider an admission optimization problem of a rehabilitation clinic. A well-designed admission policy is critically important for optimizing access times while simultaneously ensuring that patient appointments are managed effectively by not exceeding the capacity. This is essential to prevent overcapacity and maintain the quality of care. Although this research is visioned for a rehabilitation clinic, its practical implications can be extended to other multi-disciplinary healthcare settings, such as psychology wards, or other specialty clinics, such as cancer clinics or epilepsy clinics. By implementing similar strategies, these sectors can enhance their operational efficiency and improve patient outcomes.

For this project, we are inspired by the combined neurology and surgical department of the rehabilitation clinic *Merem*. This department treats inpatients, patients who stay within the clinic, and outpatients, patients who visit the clinic for appointments. Patients are treated by various disciplines, such as occupational therapy, physiotherapy, and speech therapy, as defined by a given treatment plan. The clinic has established a formal agreement with the nearby hospital to facilitate all inpatient admissions, given that a bed is available. The inpatients take up a part of the capacity outside the influence of the planner. Outpatients, however, are placed on a waiting list first. The planner needs to determine every week which patients to admit. He should ideally fit all appointments within the set capacity. The admission of outpatients is the only influence a planner has on the asked demand. Currently, no defined policy is made on which patients to admit. This can lead to overuse of capacity or longer access times that could have been foreseen.

Consequently, we formulated the following research question:

How can outpatient admission planning within a multi-appointment, multi-disciplinary setting for a rehabilitation clinic be optimized, taking into account access time and overuse of capacity ?

That is to say, this thesis aims to develop an admission policy based on a mathematical model by finding a balance between access times and overuse of capacity. We organize this thesis as follows. In Chapter 2, we describe the current planning at *Merem* and perform a discrete event simulation to evaluate the demand-capacity ratio. A literature review is given in Chapter 3. The first part of the literature review describes the current state-of-the-art inpatient admission planning and patient appointment scheduling. The main goal of the second part of the literature review is to give a short overview of the theory of the models and methods used within this project. We formulate a Markov Decision Process (MDP) in the first part of Chapter 4. This MDP models the dynamics of patient admissions and the duration of their treatments. The decision to be made in each decision epoch is which patients to admit. We conclude that directly solving the best decision in each epoch for this MDP is computationally impossible due to the size of the state space. In Chapter 5, we describe techniques to approximate the solution of the MDP and choose a solution approach that consists of a combination of a Direct Lookahead and Cost Function Approximation policy, a DLA-CFA policy. We establish the framework for this policy application, in which a parametrized demand forecast is made for every decision epoch. This forecast can be done through a stochastic tuning or a newsvendor model. We tested the designed policy on single- and multi-disciplinary systems in Chapter 6. Finally, we apply the designed policy to a setting that resembles the rehabilitation clinic *Merem* at the end of this chapter. In Chapter 7, an analysis of our research will be presented, accompanied by suggestions for further research. We conclude this thesis in Chapter 8.

Chapter 2

Current situation

In this chapter, we give an overview of the rehabilitation clinic *Merem*, focusing on its combined neurology and surgical team. This chapter includes a general description of the clinic and explains the current planning strategy. Furthermore, a data analysis is performed to assess the performance of this department. This chapter is concluded by motivating the research question based on the current state of the scheduling procedure and expert knowledge.

2.1 The rehabilitation clinic

Merem is a rehabilitation clinic for both adults and children [4]. Its main location is in Hilversum, where inpatients and outpatients are served. The organization also has locations in Almere and Lelystad, serving only outpatients at these sites. The location in Hilversum is divided into multiple teams, each serving a different type of patient. Figure 2.1 shows an overview of the various departments. A large part of the clinic is specialized in children's treatment and rehabilitation, consisting of the mythyl school, the toddler group (TPG), and the 4 to 20-year-old (4-20Y) department. The pain department helps patients with severe chronic pain. The lung department specializes in improving the lives of patients with lung diseases, such as COPD, by treating physical and psychological complaints. The above departments will not be considered in this project, as these have a different planning strategy than the department considered.

The focus of this project is the neurological and surgical department. In this department, patients with neurological and/or surgical implications (e.g., CVA or paraplegia) are treated. The department's team serves both inpatients and outpatients.

Inpatients are patients who are staying in the clinic. These individuals typically enter the healthcare system through the emergency room of a local hospital (*Ter Gooi Ziekenhuis* or *Flevoziekenhuis*). After the initial hospitalization and treatment in the neurological/surgical department of the hospital, the patient continues to the rehabilitation clinic. These patients are considered emergency patients for the clinic, as upon enrollment, the clinic has four (working) days to admit the patient. This is an agreement with the nearby hospitals, following the nationally regulated 'treeknorm' [5]. The condition of this patient type can range from very severe (e.g., no ability to walk nor talk) to relatively well (e.g., only arm- and handmisfunction), and therefore, the length of stay per patient can also differ highly, see Section 2.3.1. After some time at the clinic, the patient can return home and continue in outpatient care by coming to the clinic a few times a week for treatments.

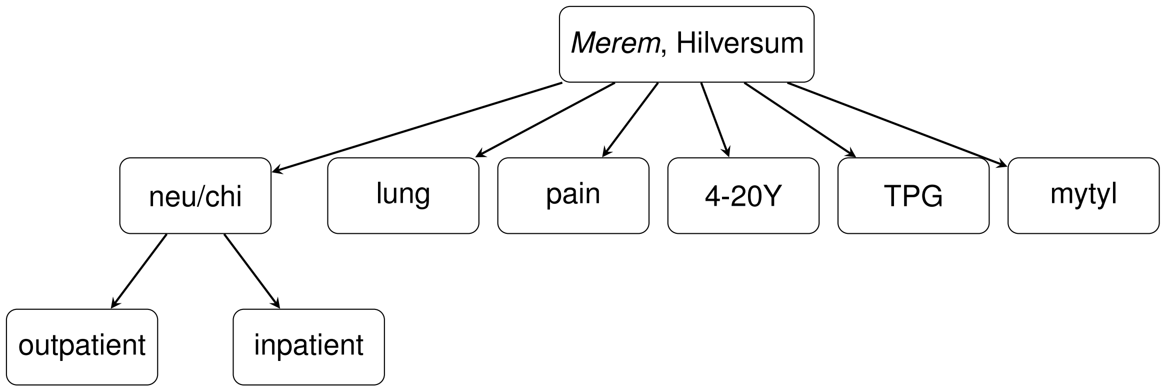


Figure 2.1: Visualization of the different departments of Merem at the location in Hilversum.

An outpatient is a patient who visits the clinic for care without staying overnight. An outpatient can be referred by the general practitioner or come from another rehabilitation clinic or hospital. These patients are placed on a waiting list, ranked by access time and priority. If space is available in the schedule, these patients can be admitted.

The neurological/surgical department team of *Merem* comprises about 30 doctors and health experts, divided into physiotherapists, occupational therapists, and speech therapists. In addition, the team includes psychologists and social workers. Apart from direct patient treatment, health experts also need time for administrative tasks and meetings with colleagues to, for example, discuss patients or organizational issues.

2.2 Current planning strategy

The planning department of *Merem* comprises 15 planners, of which one is fully dedicated to the neurological/surgical department's planning. The planning includes (planned) meetings between the practitioners and/or physicians, patient-practitioner appointments, administration time for practitioners, and resting and eating time for the inpatients.

This results in a weekly schedule for each patient and practitioner. This schedule is published and sent out to the outpatients every Friday for the next week. The practitioners and (inpatient) patients also get a printed schedule at the beginning of each day.

2.2.1 Weekly schedules

This subsection outlines a typical week for inpatients, outpatients, and practitioners to give a clearer picture of how their schedules look. It's important to note that planners can only view this schedule, they do not have access to additional metrics, such as available free periods or the number of patients admitted.

Patients

A simplified example week of an inpatient's schedule at the clinic is presented in Figure 2.2. This patient's prescription consists of five times physiotherapy, five times occupational therapy, and five times speech therapy this week. Furthermore, the doctors will conduct weekly clinical rounds on Tuesday afternoons, as can be seen. Also, a swimming session (hydrotherapy) and lunch and resting periods are planned. The neurological/surgical department planner does not plan the hydrotherapy, but a swimming instructor schedules the swimming sessions for all patients within *Merem*. Furthermore, this patient has an appointment in the hospital, so the speech therapy session on Thursday has been canceled.

	MON	TUES	WED	THU	FRI
8.30	REST	REST	REST	REST	REST
9.00					
9.30					
10.00	ET				
10.30					ET
11.00		ET	FT		
11.30				ET	
12.00	LUNCH	LUNCH	LUNCH	LUNCH	LUNCH
12.30	REST	REST	REST	REST	REST
13.00	REST	REST	REST	REST	REST
13.30					
14.00		Doctor's visit		FT	FT
14.30			FT	ET	
15.00	HT				
15.30					
16.00	LOGO	LOGO	LOGO	Hospital visit	LOGO
16.30					
17.00	FT				
17.30					
18.00	REST	REST	REST	REST	REST

Figure 2.2: Example week of appointment for an inpatient in the neurological ward in Merem. ET = occupational therapy, HT = hydrotherapy, LOGO = speech therapy, FT = physiotherapy.

This example week is, however, a simplified overview, as a patient's prescription can contain other types of therapy sessions and activities than shown. These include a fitness (group) session, breakfast support group (to help with breakfast activities), specialty physiotherapy group sessions (e.g., the arm/hand group where arm and hand exercises are practiced), et cetera. Group sessions complicate planning because they are scheduled at specific times and are not easily rescheduled.

In Figure 2.3, an example week of an outpatient is visualized. This patient visits *Merem* on Tuesdays, Thursdays, and Fridays. Typically, outpatients have fewer treatments per week. Note further that it is preferred that the appointments are planned back-to-back. This results in less access time for the outpatients. In comparison, for inpatients, at least one free time slot between consecutive appointments is preferred, such that the patient can get some rest, and appointments are more spread out over the day and week.

	MON	TUES	WED	THU	FRI
9.00					
9.30					
10.00					
10.30		ET			
11.00		FT			
11.30				FT	FT
12.00					ET
12.30				HT	
13.00				ET	
13.30					
14.00					
14.30					
15.00					
15.30					
16.00					
16.30					

Figure 2.3: Example week of appointments for an outpatient in Merem. ET = occupational therapy, HT = hydrotherapy, LOGO = speech therapy, FT = physiotherapy.

Practitioners

Each practitioner can see their schedule online. This schedule includes which patient the practitioner is seeing at which time, as well as the meetings with other practitioners, time to do administrative tasks, and time for lunch. An MDM (*multidisciplinairy meeting*) is a meeting with all practitioners where the progress of multiple patients is discussed. An example schedule for a Tuesday for a (physiotherapy) practitioner, including an MDM, is shown in Figure 2.4.

2.2.2 Planning development

The current planning strategy can be described as a 'pen-and-paper' solution and is performed manually by inserting the appointments one-by-one in a computer program. Such a strategy can result in rigid schedules and is time-inefficient, as free slots need to be found by just "looking and scrolling". However, the planner still applies some structure to maintain continuity, which will be outlined in this section.

The planning strategy is made with a structural basis and a date-based finalization. Each patient has a prescription that describes the frequency of appointments per discipline per week, e.g., six times physiotherapy, five times occupational therapy, and five times speech therapy. The admissions of inpatients are considered by the admission committee. The admission committee, consisting of the doctors and two administrative personnel, consults with the hospital to determine these admissions and the admission date. The planner does not have any control over when these admissions occur. New outpatients initially enter a waiting list. The planner is responsible for determining which patients to admit from this waiting list, considering factors such as access time and the prescribed treatment plan.

	Planning	Patient type
8.30		
9.00	Weekly	
9.30	meeting	
10.00	ADP	
10.30	RPB - Ms. Janssen	K - NEU
11.00	FT - Mr. Hesen	K - NEU
11.30	ADP	
12.00	FT - Mr. Dobbelsteen	R - NEU
12.30	LUNCH	
13.00	MDM	
13.30		
14.00	FT - Ms. de Boer - Blekers	R - CHI
14.30	FT - Mr. Van Aert	K - NEU
15.00	FT - Mr. Hesen	K - NEU
15.30	ADP	
16.00	Fitness group	
16.30		
17.00		

Figure 2.4: Simplified overview of a Tuesday schedule of one of the physiotherapists. The names of the patients are fictional. The patient type is defined as inpatient (K) or outpatient (R), neurological (NEU), or surgical (CHI). The FT means a normal physiotherapy session. ADP is the planned time for administration tasks. RPB is a meeting with the patient to discuss the progress.

When a new patient is admitted to the clinic, a structural base planning is created. This planning encompasses all prescribed therapy sessions. However, it excludes incidental appointments, such as hospital visits (orange boxes in Figure 2.2). The patient is paired with two practitioners per discipline to enhance continuity of care. In the event that one practitioner is unavailable due to illness or vacation, the other practitioner will maintain oversight of the patient's treatment. The planner should ensure that the patient is seen by both practitioners at least once a week.

At an MDM, the doctors and practitioners can change the frequency of the appointments of patients. Typically, each patient is discussed every other week. The number of modifications addressed by this meeting differs per week. According to the main planner of the department, the average amount of mutations per week is around 70-80. Implementation means that the structural base planning for the patient is adapted to the new prescription. The planned discharge date is also discussed during an MDM.

The planner handles the changes semi-online (see Section 3.1), meaning that a request for a change is implemented directly, not (necessarily) taking into account possible further changes. This process is inefficient. For example, consider a situation where a significant change to a patient's treatment plan is requested on Tuesday, suggesting an increase to eight physiotherapy sessions per week instead of five. However, by Thursday, the planner receives notification that the patient has been discharged. In this case, the planner has spent time on tasks that ultimately proved unnecessary.

On Tuesday afternoon, after the requests of the morning team meeting are handled, the structural base planning planning is copied to the date planning. This planning contains incidental appointments, holidays, et cetera. This merging will result in conflicts in the date planning, as incidental appointments overlap with structurally planned therapy sessions or even problems in the structural planning. It is then the task of the planner to overcome the conflicts by switching, moving, or canceling appointments. The final plan for next week is sent out to the patients on Friday.

2.3 Data analysis

In this section, we describe the data analysis performed, with the main goal of introducing us to the planning size and assessing the current strategy. Furthermore, the management notes the irregularity of demand throughout each week, which we want to evaluate. We perform a discrete event simulation to evaluate the overuse of capacity.

In Section 2.3.1, the available data is described, and the possible lack of it is explained. An analysis of the system parameters is provided in 2.3.2, including patient population, the length of stay, the number of therapy sessions, and mutations. Based on these parameters, a simple discrete event simulation is performed to analyze the system behavior over time, described and visualized in Section 2.3.3. Next, in Section 2.3.4, the current strategy is assessed, for which, among others, the access time and the number of cancellations are analyzed. Due to the incompleteness of the data set, expert knowledge is also used to verify and get a complete insight in the system.

2.3.1 Origin of the data

The data is retrieved from the data-warehouse of the clinic, Ecaris. The data contains all appointments of patients who received treatment at the neurological/surgical department in *Merem* in 2022 and 2023. These years were chosen as it excludes the COVID year, which could have led to a difference in population due to extra patients after IC admission with COVID-19. Two datasets were obtained: the first containing all planned appointments for in-patients and outpatients. The second dataset includes all appointments that were confirmed and documented by the practitioner. Most appointments are contained in both datasets. However, if an appointment is canceled at the last minute, for example, due to the sickness of the patient or practitioner, the appointment is not shown in the second data set. However, if an appointment is canceled at the last minute, such as due to the illness of either the patient or the practitioner, that appointment will not be reflected in the second data set.

The appointments in the second dataset are referred to as operations, which the clinic reimburses through the health assurance companies. The data consists of 276084 appointments in the first dataset and 210635 operations in the second data set for the years 2022 and 2023. Note that we do not have an overview of the requested appointments. Therefore, we are unsure when an overuse of capacity is asked for, as the planner might have canceled appointments in advance.

2.3.2 System parameters

Using the data received, we conduct a data analysis to evaluate various system parameters, including the type of patients, arrival rate, length of stay, and number of appointments.

Type of patients

A list of unique patients the clinic served during these years can be obtained from the dataset. Each patient in the dataset has a reference number, a diagnosis, and an ICD10 description. An ICD10 description is the official disease description according to the International Statistical Classification of Diseases and Related Health Problems 10th Revision. In total, 339 patients are recorded over two years. Table 2.1 provides an overview of the number of inpatients for each diagnosis in 2022 and 2023. Most patients admitted suffered from a Cerebro Vascular Accident (CVA).

Table 2.1: Amount of patients admitted in 2022 and 2023 per diagnosis.

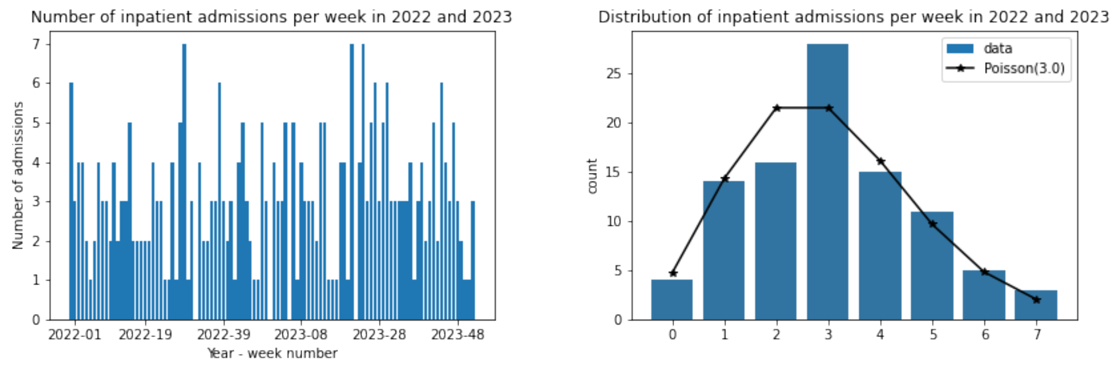
year	Description	Amount
2022	CVA	112
	Cerebrale functiestoornissen, incl. cong.	2
	Contusio Cerebri	4
	Infectieuze hersenaandoeningen	3
	Overige aandoeningen bewegingsapparaat	1
	Overige neurologische aandoeningen	1
	Overige orgaanaandoeningen	7
	Perifeer zenuwletsel, zenuwaandoeningen	35
2023	CVA	120
	Cerebrale functiestoornissen, incl. cong.	1
	Contusio Cerebri	5
	Overige aandoeningen bewegingsapparaat	1
	Overige neurologische aandoeningen	5
	Perifeer zenuwletsel, zenuwaandoeningen	16

Arrivals of inpatients

In Figure 2.5, the number of admissions inpatients per week during 2022 and 2023 can be seen. In 2022, the average number of admissions per week was 2.9. In 2023, it was 3.4 patients per week. However, as can be seen, the number of admissions differs significantly over the weeks.

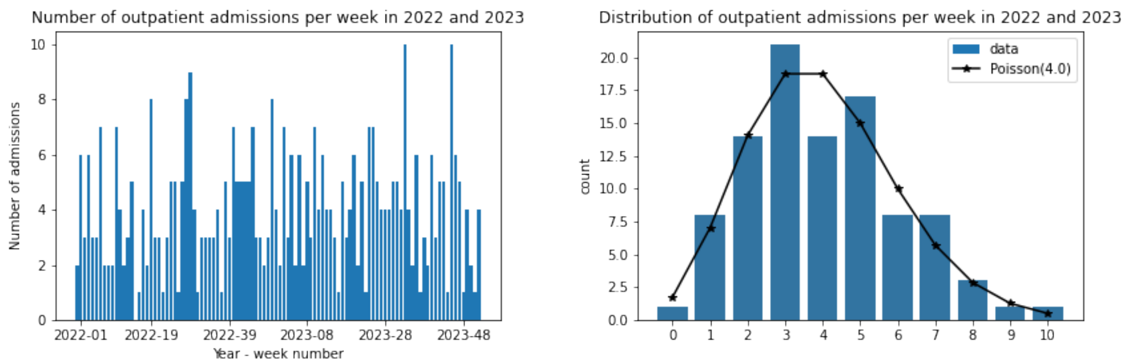
We think the arrivals occur following a Poisson distribution. If we perform a Chi-square Goodness of fit test on the arrival data and a Poisson distribution with an arrival rate of $\lambda_I = 3.0$, we obtain a p-value of 0.83. Therefore, we cannot reject the null hypothesis that the arrivals follow a Poisson distribution.

The number of admissions and discharges are closely related, as there is a maximum number of inpatients due to the number of beds available. No seasonal patterns were discovered regarding the length of inpatients' stay or the admission date. This aligns with the literature, stating that there is no seasonal variation in the occurrence of strokes [6].



(a) (b)

Figure 2.5: (a) Number of admissions per week during the years 2022 and 2023 for inpatients. (b) Distribution of the inpatient admissions per week over the years 2022 and 2023.



(a) (b)

Figure 2.6: (a) Number of admissions per week during the years 2022 and 2023 for outpatients. (b) Distribution of the outpatient admissions per week over the years 2022 and 2023.

Arrivals of outpatients

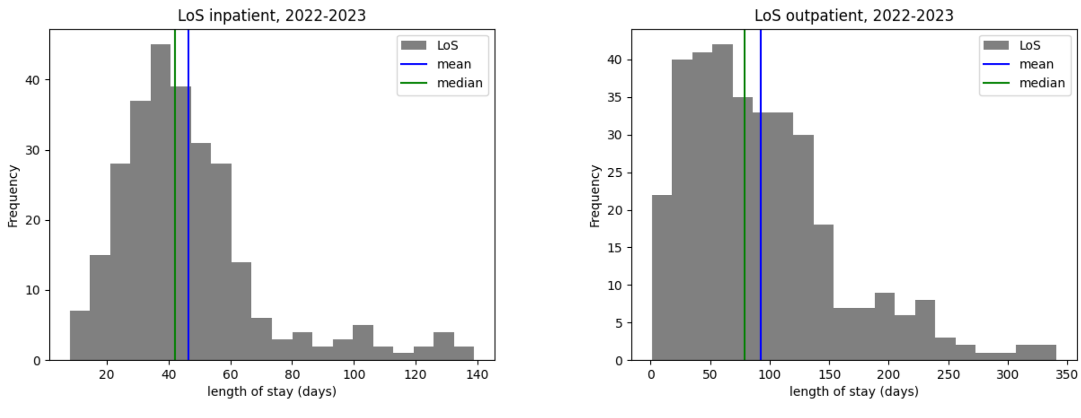
In Figure 2.6, the number of admissions of outpatients per week over 2022 and 2023 can be seen, respectively. In 2022, there was an average of 3.9 admissions, while in 2023, the average was 4.1 admissions per week. However, as can be seen, the number of admissions differs significantly over the weeks.

We think the arrivals occur following a Poisson distribution. If we perform a Chi-square Goodness of fit test on the arrival data and a Poisson distribution with an arrival rate of $\lambda_O = 4.0$, we obtain a p-value of 0.89. Therefore, if we take a 95% significance level, we cannot reject the null hypothesis that the arrivals follow a Poisson distribution. Consequently, we cannot reject the null hypothesis that the arrivals follow a Poisson distribution. Note that we performed this analysis using the appointment data, by which we actually determine the admission rate of the planner.

Length of stay of inpatients

The clinic's average length of stay (LoS) of inpatients is 46 days, with a standard deviation of 25 days, including weekends. The median is 42 days. Patients with a stay shorter than 14 days or longer than 140 days are excluded. The minimal admission length is two weeks unless other implications occur, such as hospitalization. According to the planners, staying longer than 20 weeks occurs only due to problems in the home situation, such as the need for home conversion. This excludes 15 patients out of the 339 patients in total.

In Figure 2.7, the length of stay for inpatients is visualized in a histogram. Most patients stay between 8 and 70 days in the clinic, with the first quartile at 31 days and the third quartile at 56 days.



(a) (b)

Figure 2.7: Histogram of the length of stay for inpatients (a) and outpatient (b) treated by neurological/surgical team with admission date in 2022 or 2023.

Length of stay of outpatients

The average treatment duration of outpatients is 92 days, with a standard deviation of 64 days and a median of 79 days. We disregard patients with a treatment duration of more than a year. These patients often have complex and multiple diseases and sometimes even unrelated and back-to-back admissions.

Number of therapy sessions

The number of therapy sessions per discipline can differ highly per patient and throughout the stay in the clinic or duration of treatments in an outpatient setting. In Figure 2.8, a given patient's therapy sessions are shown while inpatient. As can be seen, the number of therapy sessions changes weekly. Initially, the number of sessions was relatively high, after which it decreased slightly. Note that only individual patient sessions are taken into account here. Therefore, group sessions such as the fitness group are left out.

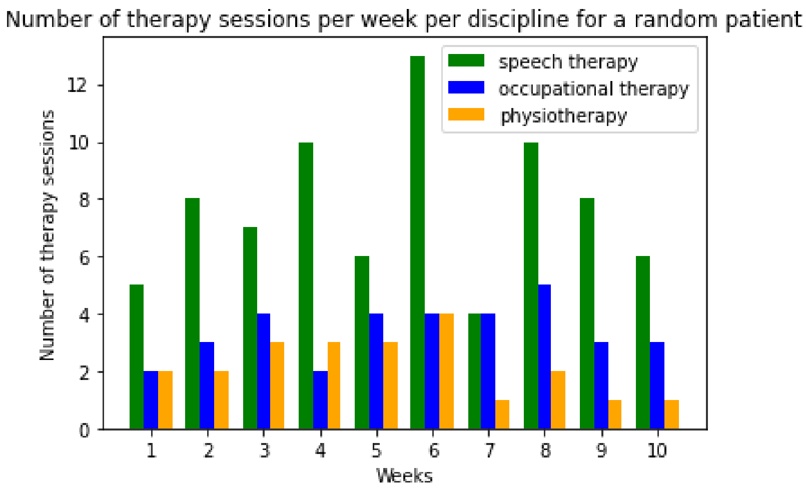


Figure 2.8: Number of therapy sessions per discipline (occupational therapy (ET), physiotherapy (FT), and speech therapy (LOGO)) for a given inpatient during its stay in the clinic.

There are multiple reasons that can be addressed as to why changes can be seen every week. Firstly, the practitioners can change the number of sessions per patient weekly. For example, practitioners would like to see the patient for extra sessions in the first and last week. However, in this case, other reasons also play a role. This patient stayed in the clinic during the national holidays of Easter, King's Day, and Ascension, which means fewer sessions can occur. Furthermore, one of the patient's main physiotherapy practitioners was sick in week 7, leading to fewer sessions.

Based on the data, we can compute the distribution of the number of therapy sessions per week per discipline and per patient type. In Figure 2.9, a histogram is made for each discipline and patient type, and a normal distribution is fitted. Note that three distributions are fitted within the speech therapy for inpatients. This can be explained by the fact that the patient can be divided into three groups with no, light or severe aphasia.

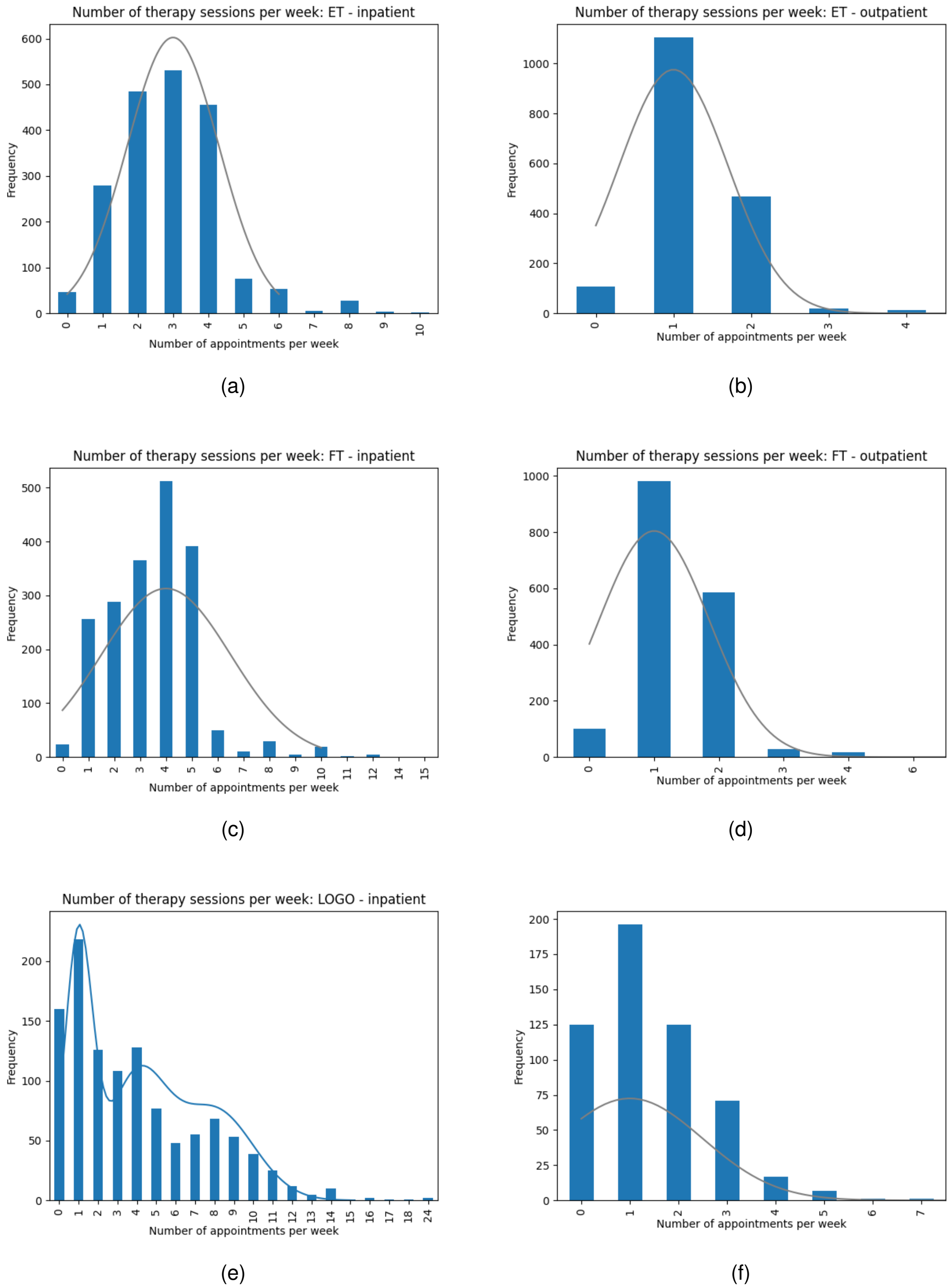


Figure 2.9: Number of therapy sessions per discipline (occupational therapy (ET), physiotherapy (FT), and speech therapy (LOGO)) and per patient type (inpatient and outpatient) determined on patients treated in 2022 and 2023.

Therapist utilization

The management of *Merem* strives for a 60% to 70% direct utilization, meaning sixty percent of a practitioner’s time should be dedicated to physical patient treatment. Furthermore, 20%

of the practitioner's time is dedicated to indirect patient time, such as administrative tasks and meetings. The remaining 20% of time is *conditional time*, which can be devoted to tasks not directly related to patient treatment, such as internship supervision, management tasks, et cetera.

In Table 2.2, we determined the fraction of time used for direct patient treatment for three different disciplines. We also determined the fraction of patient appointments used for individual patients, so we disregarded group sessions. This table also includes the total contracted working hours per discipline. Note that an appointment lasts 30 minutes.

Table 2.2: Summed working hours, the fraction of time dedicated direct patient treatment and the fraction of appointment dedicated to individual patient treatment per discipline, based upon the appointment of 2022 and 2023.

discipline	Working hours (per week)	direct patient treatment (time ratio)	individual patient treatment (number of appointments ratio)
ET	220	0.6	0.5
FT	260	0.6	0.6
LOGO	100	0.7	0.7

We conclude with a direct individual patient capacity of 132, 166, and 98 appointments per week for occupational therapy (ET), physiotherapy (FT), and speech therapy (LOGO).

2.3.3 Discrete event simulation

Based on the data analysis of the system parameters above, we made a discrete event simulation to evaluate the capacity-demand ratio. For this simulation, we used the determined distribution of the number of arrivals per week, the length of stay, and the number of therapy sessions per discipline per week, split out to inpatients and outpatients.

The simulation generates new inpatients and outpatients for each simulated week. We assume that each patient is admitted on a Monday and discharged on a Friday. A limit of 25 present inpatients is set, as this is the number of beds available in the clinic. The number of therapy sessions per discipline is drawn from the determined distributions every week, independently of the number of therapy sessions the week before. Figure 2.10 shows the outcome of one simulation run. The simulation is run for 50 weeks.

We can note that the demand changes highly over the weeks, which is also often emphasized by the planners. With the simulation, we can analyze the demand per discipline and relate this to the capacity. The capacity is based on the total working hours per week per discipline summed over all practitioners, the fraction of direct patient treatment, and the fraction of individual patient treatment, which is given in Table 2.2, in Section 2.3.2.

In Figure 2.11, the demand is visualized separately per discipline. The horizontal line in each figure resembles the capacity of the practitioners to be dedicated to direct individual patient treatment. We note that the ratio of capacity versus demand is most often exceeded for speech therapy. The planners recognize this as well, but note that the speech therapists are the most flexible with their schedules, such that capacity can easily be increased if needed.

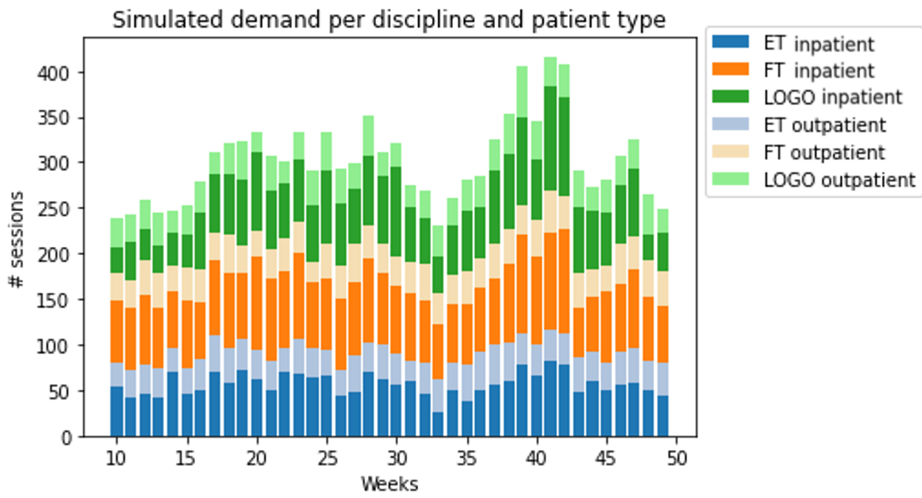
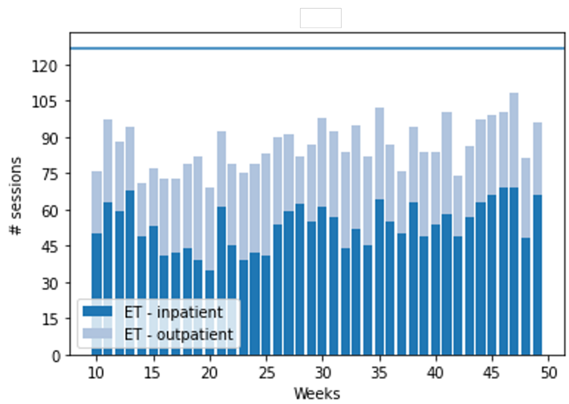
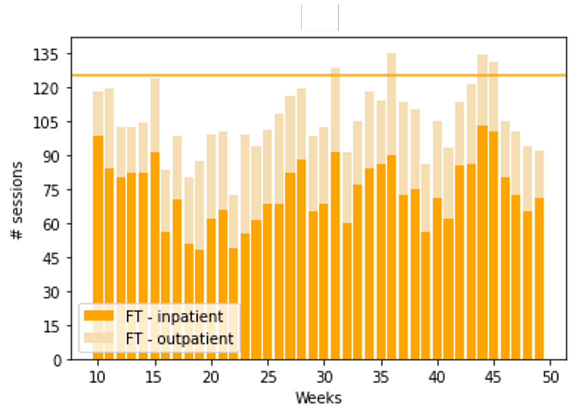


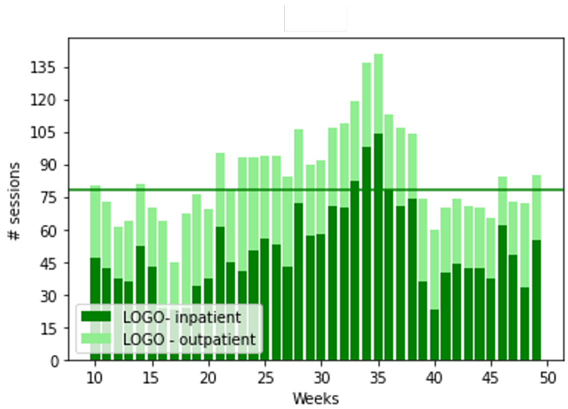
Figure 2.10: Total demand for therapy sessions in an example simulation run. ET = occupational therapy, FT = physiotherapy and LOGO = speech therapy



(a) occupational therapy



(b) physiotherapy



(c) speech therapy

Figure 2.11: Simulated demand per discipline. The horizontal line represents the weekly capacity.

2.3.4 Performance indicators

We define multiple performance indicators to assess the performance of the current planning strategy. Not all performance indicators have readily available data for comprehensive evaluation. In such instances, we seek the expertise of planners with specialized knowledge to provide informed estimations.

Waiting list

The waiting list for the neurological/surgical department of *Merem* consists of patients who want to start an outpatient trajectory at the clinic. The capacity does not always allow for another outpatient, so the planner adds these patient types to a waiting list.

A comprehensive analysis of the access time was conducted within the organization. However, the associated data is no longer available. It was concluded that the average access time for neurological patients in 2022 was 11.8 days, and for surgical patients, it was 15.0 days. In this case, the access time is from registration at the clinic to the first consultant appointment. It excludes patients with a high emergency (e.g., via hospital).

However, this does not reflect the actual situation. Patients are frequently registered, yet their condition may deteriorate to the point where they cannot be admitted to the clinic despite available capacity. This circumstance leads to prolonged access times. Excluding the patients outside the 'treeknorm' results in an access time of 9.2 days for neurological patients and 8.6 days for surgical patients. The waiting list for 2023 is not available.

Number of cancellations

The data only consists of appointments and operations in the finalized planning. Therefore, the number of cancellations cannot be determined using the given dataset. Cancellations may arise from various factors, including limited capacity, which necessitates prioritizing patients to ensure an effective scheduling process. Additionally, unexpected circumstances, such as the illness of a patient or the unavailability of a practitioner, can also lead to cancellations. The neurological department's planner addressed that weekly cancellations due to capacity occur and that the number of cancellations differs significantly from week to week. The exact number of cancellations cannot be determined due to a lack of available data.

Amount of time needed to plan

With the current planning strategy, one planner is a full-time (36h/w) planner dedicated to planning the neurological and surgical department. The expertise of the current planner makes it possible to finalize the planning each week. Without such expertise, more time would be needed. This expert knowledge consists of both what combinations of appointments are and are not allowed, but also already by using some policy by having a feeling what is and what is not possible when correcting the schedule, and is expected to perform better than a simple myopic policy.

2.4 Requirements, research direction and motivation

The analysis above shows that the current planning strategy is not inherently flawed. Still, given some observations, several suboptimal points can be addressed. The planning manager denotes the imbalances between capacity use over the different weeks. In one week, the clinic is working under capacity, while in other weeks, appointments need to be canceled due to overcapacity. We also note this in our discrete event simulation. This negatively impacts the satisfaction of both patients and employees. One reason that can be addressed for this is the high level of changes made by doctors and practitioners weekly. If we decrease this number of changes, could we design an admission policy that equals capacity use?

Another point to consider is the expert knowledge used to complete the planning on time. When asked why the planner is doing something, such as not admitting a patient from the waiting list while there is free capacity, the planner likely replies, ‘Then I have room for the emergency patient, who might come next week.’ A dedicated admission policy is not applied. Such knowledge is not written down but built upon years of experience. If a new planner would be appointed, this knowledge needs to be built up again. The admissions of outpatients from the waiting list are currently managed by planners, but there is no policy in place.

Thus, a well-designed admission policy could eventually be utilized alongside or even replace human planners, and reducing the necessity for such expert knowledge. A good admission policy should give better possibilities for better appointment scheduling, meaning less capacity overuse and fewer cancellations. Moreover, it will also provide insight into capacity needs and access times. In this project, we will design such an admission policy, which takes into account both access time and capacity overuse.

Chapter 3

Literature review

This chapter gives an overview of the literature relevant to our research. First, in Section 3.1, literature related to planning for health care, in general, is briefly described. This topic has received more and more attention in the last few years. The section is divided into different parts, describing different approaches for various healthcare settings. Hereafter, in Section 3.2, we provide an overview of literature related to the used models and methods for the project. Based upon this section, we describe to what extent our project could contribute in Section 3.3.

3.1 Health care planning in general and for rehabilitation clinics

For a general overview of Operations Reach and Management Science methods in Health Care, we refer to Hulshof et al. [7]. This article gives a taxonomy classification for different hierarchical levels in health care planning, dividing the planning into strategic, tactical, and operational levels. Where strategic planning addresses structural decision-making for the long term, this project mainly addresses both tactical planning and operational planning. Tactical planning denotes the planning ahead, being the organizational structure of the operations. Operational planning involves short-term decision-making, e.g., incoming mutations. This level can consist of online operational planning or offline operational planning. In online planning, the planner directly reacts to the changes prescribed. Offline planning progresses decisions up to a specific time, after which all changes are implemented simultaneously. In this project, the problem's operational part is mainly tackled offline, although the current planning strategy can be seen as a combination of the two. Generally, the choice between online and offline planning is a management decision. In literature, offline planning is more described, as it is easier to model than online planning, [8].

Most research in appointment planning in health care is done for hospitals. As hospitals are often regulated around one central bottleneck, e.g., an operating room, this can require a different planning strategy than for a rehabilitation clinic, where such a bottleneck often does not exist [8]. Much research in scheduling in hospitals is dedicated to what we would call acute patient scheduling, in the sense that the planning should take into account beforehand unknown arrivals of emergency patients without interfering with the regular schedule too much. This is evident in emergency care services, surgical services, and bed allocation issues. Section 3.1.1 provides a comprehensive overview of the literature related to acute (inpatient) patient planning and establishes a connection between this topic and admission planning for both acute and elective patients.

Of the many papers published on emergency appointment planning, early work already considered the scheduling in rehabilitation clinics such as Kolesab [9], but in a very general and non-technical manner. Recently, rehabilitation and, in particular, physiotherapy planning have become areas of increased research focus. Chien et al. [10] developed a genetic algorithm with scheduled patient appointments according to precedence constraints and using a hybrid shop model problem. This resulted in a decision support system to help the planner of the rehabilitation clinic. A hierarchical-built, integer optimization model to conduct weekly personnel scheduling for physical therapists is presented by Ogulata et al. [11]. The authors considered the urgency of the treatment of patients and a varying treatment time. The model is updated and recalculated on a weekly basis. Emergency patients are not considered necessary.

However, these last two cited papers only consider one discipline within the rehabilitation clinic. Most often, rehabilitation patients (and sometimes hospital patients) require multiple appointments with multiple disciplines at the same day or week. Section 3.1.2 gives an overview of research in multi-disciplinary and multi-appointment scheduling.

3.1.1 Acute patient planning

In this section, we do not consider actual emergency service planning. This department is often fully dedicated to emergency patients and, therefore, needs a different approach than if patients interfere with the regular schedule. Therefore, the term acute patients is introduced.

A typical acute patient planning problem is an operation room, where planned patients must be scheduled for a day, in a room, for a given time, et cetera, taking into account the possible arrival of emergency patients and minimizing waiting time, working overtime, et cetera. For example, Zhang et al. [12] designed a two-phase optimization model, combining a Markov Decision Process and stochastic programming for an advanced surgery scheduling problem.

On a tactical level, the available resource capacities are often divided among different defined patient groups by allocating time blocks to these specific patient groups. Such division can be based on various performance measures such as utilization of the operating rooms or patient waiting time as studied by Cardoen et al. [13]. On the (online) operational level, the incoming emergency patients should be inserted into the schedule. This can be done by delaying, canceling, or rescheduling other surgeries.

Another example of acute and inpatient scheduling is admission control and bed allocation in (hospital) wards. Samiedaluie et al. [14] find policies for this problem in the context of a neurology ward. Using an infinite horizon average cost dynamic program, the developed policy reduced the overall deterioration in patients' health status compared to several alternative policies.

Considering the Intensive Care Unit, Kim et al. [15] analyze the admission and discharge process through a queuing model and a simulation study to give insight into capacity utilization. Bai et al. [2] also consider the admission process of the ICU. The authors employ a Markov Decision Process to model the admissions procedure, thereby facilitating an analysis of the trade-offs between medical and financial objectives. They demonstrate that a substantial reduction in expected mortality can be attained with minimal increases in monetary expenditures.

Admission control is not limited to the Intensive Care Unit (ICU), it can a critical consideration across various departments within the healthcare system. A more generalized framework for admission scheduling and resource requirement forecasting is presented by Garg et al. [16]. Shen et al. [17] create a data-driven newsvendor model to study the elective-emergency control problem, such that complex factors such as uncertain bed capacity and unknown true probability distributions of patient arrivals and departures are included. A review of admission, discharge, and transfer control within different types of wards is provided by Zamani et al. [18], considering policies for ICUs and general wards.

We also want to note Zhu et al. [19], who, in contrast to the previously mentioned references, examine an admission policy within a multi-disciplinary context characterized by varying treatment patterns in a hospital setting. They employ a Monte Carlo tree search to solve the formulated MDP.

To summarize, the types of acute patient issues, particularly in the context of admission control, try to anticipate on future arrivals that influence the schedule or demand. They apply models such a Markov Decision process or queuing models. The problem we consider in this project also deals with such an unknown future.

3.1.2 Multi-appointment and multi-disciplinary scheduling

For a review of research related to multi-disciplinary appointment scheduling in health care, we refer to [20, 8]. Marynissen et al., [20], specifically review multi-appointment scheduling within hospitals, noting that this is most often performed in oncology and rehabilitation departments. Leeftink et al., [8], describe cross-relation between different applications that have been researched, also outside hospitals. In this review, a separate distinction for rehabilitation clinics is also made. It notes Braaksma et al. [3], who considers multi-disciplinary rehabilitation treatment week planning where the combination of appointments, therapist utilization, and simultaneous start of treatments are prioritized, and only outpatients are considered. The authors conclude that this approach can result in valuable decision-making support for the clinic where the study is performed. The following text presents more specific references aimed at enhancing our understanding of the various aspects of multi-disciplinary scheduling in healthcare.

Schimmelpfeng et al.[21] perform a similar job. Similar to Braaksma, it considers inpatients, distinguishes between individual and group sessions, and considers the time patients need to recover between sessions. Furthermore, it allocates appointment sessions to specific times during the day instead of only considering weekly capacity. The goal of becoming a decision-making support is noted.

The main goal of Saure et al. [22] is to show that computer-based scheduling is more effective than a 'pen-and-paper' solution. This is done by formulating the problem as a multi-objective combinatorial optimization problem. The approximated solution is sought through a three-stage local search-based approach.

With the disregard for emergency patients, all relevant factors become apparent, thereby the scheduling problem becomes discrete. Therefore, Raschendorf et al. [23] shows that the problem can be described as hierarchical edge coloring, where each color resembles one day of the week. By doing so, it can more easily be shown if a given problem is solvable or not. To find a solution of a multi-criteria model for daily therapy appointment scheduling, considering therapist satisfaction and priority of appointments, Kling et al. [24] develop a Greedy Randomized Adaptive Search Procedure (GRASP) in order to construct the most optimal schedule.

On the downside, we note that scheduling usually occurs only over a short period or when the variables remain unchanged after the scheduling occurs. As in rehabilitation clinics, the frequencies of treatments can still be changed in a short time frame, so this is difficult to apply here. Leeftink et al. [25] consider stochastic integer programming in multi-disciplinary appointment planning, where the arrival of patients to the oncology department for diagnosis is an uncertain variable, and a blueprint schedule is used to optimize patient waiting times over the day. The problem described is a flow-shop problem in which the appointment order matters.

A stochastic program that accounts for unknown patient arrivals and variable service times is outlined by Hur et al. [26]. The deterministic version of this model is formulated as a mixed-integer program. From there, a two-stage stochastic program is developed to address the inherent randomness. For practical reasons, the authors created appointment templates with fixed time slots to derive a solution to the program.

To summarize, the literature on multi-appointment multi-disciplinary scheduling typically employs deterministic models, which are often addressed using heuristic approaches. Additionally, an optimal policy is established through the development of a blueprint schedule that serves as a guiding framework for effective resource allocation and appointment management.

3.2 Models and methods

This section discusses some of the literature on the models we have used in this project as applicable solution methods. We revise relevant literature on Markov Decision Processes in Section 3.2.1, Stochastic Approximation in Section 3.2.2, and the newsvendor model in Section 3.2.3. This theory is used throughout our project. The details are explained in the relevant chapters of this report to keep this section short.

3.2.1 Markov Decision Processes

A Markov Decision Process (MDP) is an approach to modeling a decision-making process with uncertainty. In such a process, a decision-maker observes the system's state. Based on this, the decision-maker chooses an action, resulting in a certain cost/reward, and the system evolves to a new state at a subsequent moment in time, named an epoch.

We describe an MDP model by a tuple $(\mathcal{T}, \mathcal{S}, \mathcal{A}_s, p(\cdot|s,a), r(s,a))$. Here $\mathcal{T}$ is the set of decision epochs in the time horizon of the problem, $\mathcal{S}$ is the set of all possible states of the system, $\mathcal{A}_s$ is the set of all possible action for a state s of the system, $p(\cdot|s,a)$ is the probability matrix describing the transition from state s with action a to the next state, and $r(s,a)$ is the reward (or cost, in which case we typically use $c(s,a)$) for being in state s and taking action a.

In this project, we describe the arrival and treatment duration of patients at the rehabilitation clinic as an MDP. As the arrival and change in demand happen under uncertainty, and a decision for the admission of patients should be made at multiple time points, this is a suitable model.

Solving an MDP is to find the optimal policy that maximizes (or minimizes in the case of a cost function) the objective function. How this objective is defined precisely depends on the problem itself. For more details about MDPs, we refer to Puterman [27].

Only for a small set of MDPs can the solution be determined exactly by applying solution strategies like value iteration, policy iteration, or linear programming. Slightly more complex MDPs often suffer from *curse of dimensionality*, a term introduced by Richard E. Bellman [28]. This curse indicates that the problem becomes so large that finding an exact solution is computationally impossible. To find a near-optimal policy, we need to apply different solution methods. Powell [29] describes different techniques under the name of Approximate Dynamic Programming (ADP), which is the most researched strategy but not the only strategy.

In subsequent work, Powell [30] describes two strategies, each consulting two policies, to solve the problem, highlighting the underlying commonalities across various research domains. Additionally, in [31], he asserts that, in principle, you can use any policy for any problem, but specific policies often yield superior results for particular problems. The details of which we will return to in Section 5.

In this project, we use the proposed combined policy, which combines the Direct Lookahead and Cost Function Approximation policies (DLA-CFA). Powell was the first to name it as such specifically, as shows the use of such combined policy for an energy storage problem by Powell et al. [32].

However, other research, often on vehicle routing and employee distribution problems, has been done without explicitly naming the combined policy as such. For instance, Chen et al. [33] explores and resolves a similar problem and solves it by employing a forecast cost function approximation to enhance decision-making efficiency in technician allocation.

3.2.2 Stochastic approximation

To solve the model created in this project, we need to apply stochastic approximation (SA) methods. The stochastic approximation is very similar to general optimization or root-finding problems. However, the describing function is not deterministic. The basic goal of stochastic is to find:

$$\min_{\theta} f(\theta) = \mathbb{E}_{\xi}[F(\theta, \xi)],$$

where ξ is some random variable. A comprehensive introduction to SA is provided by Spall. [34], who handles multiple methods such as Stochastic Search, Simulated Annealing, and Genetic Algorithms.

3.2.3 Newsvendor model

The newsvendor model, or newsboy problem, is a classical inventory problem in operations research and operations management and was first described by Arrow et al. [35]. The model identifies the optimal number of newspapers, denoted as x, that a newsvendor should purchase. This determination is based on the known probability distribution of daily demand, represented by the random variable R, as well as the wholesale cost of purchasing each newspaper, represented by w, and the selling price per unit, represented by c. In this way, one only needs to solve the following formulation

$$\arg\max_{x} \mathbb{E}(s \min(x, R)) - wx.$$

Numerous studies have been conducted addressing this issue, encompassing a range of frameworks that include multi-item and multi-period analyses, both with and without known probability distributions, as well as involving multiple retailers. A comprehensive review of these research efforts is provided by Qin et al. [36].

Most studies related to the newsvendor model address an inventory management solved using the newsvendor approach. This is known as ”decision rule optimization” and is associated with the three main frameworks for learning policies identified by Sadana et al. [37]. An illustrative example of this type of study is presented by Halman et al. [38], who compare the newsvendor model with a single-item stochastic lot-sizing problem (SLS).

In this project, the newsvendor model is utilized to identify the parameters needed for the cost function approximation, with the same goal as in the use of SA. This approach is consistent with the Integrated Learning and Optimization (ILO) framework, another of the three frameworks as defined by Sadana et al. [37]. This methodology is also known by other names, such as “predict-then-optimize” or “end-to-end learning”. The models in this framework consist of a prediction model that provides a prediction and an optimization model that takes the prediction as input and returns a decision. An example of a study using this framework is the study of Chang et al. [39], where they first predict the localized demand for shared bikes, after which they use this prediction as an input for the relocation of the shared bike, taking into account broken bikes. The prediction is done using a neural network. Another example study related to health care is described by Jiang et al. [40]. They predict the spread of infectious diseases and use this information to optimize the allocation of resources for medicines.

The use of the newsvendor approach within this ILO framework is less common. We identified an article that comprises the newsvendor model within an MDP, although it does not totally fit within the ILO framework. The article by Amaruchkul [41] discusses the application of a newsvendor model to optimize harvest outcomes by deciding on the recruitment of seasonal workers. The newsvendor model is utilized to determine the expected reward based on the current yield distribution and the number of seasonal workers hired. Hiring seasonal workers earlier can result in lower costs. However, since the yield distribution can vary during the season due to factors such as weather, the author presents an MDP that allows for adjustments in the yield distribution and in the number of hired seasonal workers.

In the comparison of the application of the newsvendor model in this article with our research, we identify two significant distinctions between the two models. Firstly, Amaruchkul's model is characterized as a finite-horizon MDP, wherein revenue is realized at the conclusion of the season, specifically upon the completion of the harvest. Consequently, the newsvendor model is employed solely for forecasting this single final decision epoch. In contrast, our research utilizes the newsvendor model to forecast demand on a rolling basis, thereby encompassing multiple decision epochs.

Moreover, Amaruchkul employs the newsvendor model within MDP problem, utilizing it not as a tool for parameter valuation but as an integral component of the optimization process itself. Consequently, for larger systems, it is necessary for him to articulate a heuristic approach to address the MDP. In our research, we adopt the newsvendor model as an approximation method to solve the MDP, presenting it as an alternative to stochastic approximation. To the best of our knowledge, this application has not been previously documented in the literature.

3.3 Contribution

We note a gap between stochastic acute patient admission planning and deterministic patient appointment scheduling for multi-disciplinary and multi-appointment patient scheduling. Acute patient admission planning often does not consider multi-appointment and multi-period planning. In contrast, patient appointment scheduling does not consider the stochasticity of the demand and, therefore, not the long-term efficiency.

We also see that designing good appointment scheduling will be difficult without a sufficient admission policy. Therefore, we contribute a patient admission scheduling policy that considers multi-disciplinary and multi-appointment patient scheduling, where both the capacity-demand ratio and waiting time are included. When applying such a policy, one should be able to design a method for planning appointments based on the articles described above.

Additionally, when MDPs are discussed, VFA is often used to address the curse of dimensionality [32]. We design a policy for this project using the combined DLA-CFA policy.

Whereas the VFA for outsiders looks like a 'black-box' method, DLA-CFA allows for modeling like the planners themselves. This allows management and current planners to include the current way of planning, which allows for more straightforward implementation in practice. It is new to apply such a DLA-CFA policy from a healthcare point of view.

In conclusion, we advance the existing literature by employing a newsvendor model for demand estimation, which serves as an input for the DLA-CFA policy framework. While this aspect is not the initial focus of the current report, it illustrates the significant potential and applicability of this methodological approach in optimizing inventory management strategies.

Chapter 4

Model

4.1 The Markov Decision Process

In this section, we formulate a Markov Decision Process (MDP) for selecting patients from the waiting list for admission. Table 4.1 describes the sets and parameters we used. The parameters will be further explained in their corresponding sections.

We designed the model to resemble reality, but we made some assumptions. Firstly, we assume that a patient has one of the pre-described treatment plans and that the treatment plan does change during its treatment duration. Secondly, we assume that a patient has a predefined treatment duration, described by T^{start}, which can be increased for a given number of weeks. This can occur for a maximum number of times $m^{\max}$. Thirdly, we assume that the capacity for each discipline does not change over the weeks but is consistently defined by $\mathbf{r}$.

Table 4.1: Sets and parameters used in the MDP model.

Sets	**Description**
$\mathcal{T} = \{1, 2 \ldots\}$	Set of decision epochs, indexed by t.
$\mathcal{C} = \{1, \ldots, C\}$	Set of disciplines, indexed by c, with C the number of disciplines considered.
$\mathcal{P} = \{(\mathbf{p})\}$	Set of all possible treatment plans for an outpatient, consisting of the number of therapy sessions per discipline, indexed by $\mathbf{p}$, where $\mathbf{p} \in \mathbb{N}_0^C$.
$\mathcal{Q} = \{(\mathbf{q})\}$	Set of all possible treatment plans for an inpatient, consisting of the number of therapy sessions per discipline, indexed by $\mathbf{q}$, where $\mathbf{q} \in \mathbb{N}_0^C$.
$\mathcal{U} = \{1, \ldots, U\}$	Set of urgency levels, indexed by u, with U the highest level of urgency.
$\mathcal{L}_I = \{0, \ldots T_I^{\text{start}}\}$	Set of possible weeks until discharge for inpatients.
$\mathcal{L}_O = \{0, \ldots T_O^{\text{start}}\}$	Set of possible weeks until discharge for outpatients.
$\mathcal{M}_I = \{0, \ldots m_I^{\max}\}$	Set of the possible number of times that a discharge date is postponed for inpatients.

$\mathcal{M}_O = \{0, \ldots m_O^{\max}\}$	Set of the possible number of times that a discharge date is postponed for outpatients.
Parameters	**Description**
$T_O^{\text{start}}, T_I^{\text{start}}$	Initial number of weeks that a patient will be admitted for outpatients and inpatients, respectively.
$T_O^{\text{MDM}}, T_I^{\text{MDM}}$	Number of weeks between decision moment, during an MDM, and possible discharge date for outpatients and inpatients, respectively.
$p_{n,O}^{\text{cont}}, p_{n,I}^{\text{cont}}$	Describing the probability that a patient continues treatment where he already had n MDMs for outpatients and inpatients, respectively.
λ_O, λ_I	Mean number of arrivals per week for outpatients and inpatients, respectively.
$p_{\mathbf{p}}^{\text{plan}}, p_{\mathbf{q}}^{\text{plan}}$	Describing the probability that a patient enters with treatment plan $\mathbf{p}$, $\mathbf{q}$, for outpatients and inpatients, respectively.
$m_O^{\max}, m_I^{\max}$	Maximum number that a discharge date can be postponed, for outpatients and inpatients, respectively
$\mathbf{r}$	Capacity limit, described for each discipline , $\mathbf{r} \in \mathbb{N}_0^C$.
κ_c	Costs of exceeding the capacity, $\kappa_c \in \mathbb{N}_0^C$.
κ_u	Costs for not admitting a patient, $\kappa_u \in \mathbb{N}_0^U$.
K	Number of weeks that the policy forecasts the demand.

4.1.1 Decision epochs and booking horizon

Every (beginning of the) week, a decision epoch takes place. The decision describes which outpatients are to be admitted from the waiting list. We describe the current week by t. In the developed policy, we forecast demand for the upcoming K weeks. Therefore, if we decide on an action in the current week t, we already have some information about weeks $t+1$ up to $t+K$. Note that decisions of continuation for an admitted patient only take place at $T_O^{\text{MDM}}/T_I^{\text{MDM}}$ weeks before the current set discharge date.

4.1.2 State space

We denote the set of all possible states by $\mathcal{S}$. At time t, a state $S_t \in \mathcal{S}$, is a triple $\big(H_t, O_t, I_t\big)$, recording the patients on the waiting list, the inpatients admitted, and the outpatients that currently follow treatment. With t we indicate which week the state occurs. For this model, the state is not necessarily time-dependent. We describe each list in more detail below.

Patient waiting list

We describe the patient waiting list at time t by H_t:

$$H_t = ((h_{u\mathbf{p}})_{u \in \mathcal{U}, \mathbf{p} \in \mathcal{P}})_t, \tag{4.1}$$

where $h_{u\mathbf{p}}$ is the number of patients on the waiting list with urgency level u and treatment plan $\mathbf{p}$.

Outpatient list

We describe the outpatient list at time t by O_t:

$$O_t = ((o_{\mathbf{p}\ell m})_{\mathbf{p}\in\mathcal{P},\ell\in\mathcal{L}_{\mathcal{O}},m\in\mathcal{M}_O})_t, \tag{4.2}$$

where $o_{\mathbf{p}\ell m}$ describes the number of outpatients with treatment plan $\mathbf{p}$, the (currently set) number of weeks until discharge ℓ, and the number of times the discharge date is postponed m. Postponing the discharge date is determined during a Multi-Disciplinary Meeting (MDM). A new patient starts with $\ell = T_O^{\text{start}}$. Every week ℓ decreases by one for each patient, except when $\ell = T_O^{\text{MDM}}$. We assume that this week, the decision to postpone the discharge date will be made. If it is decided to postpone the discharge date, ℓ is set back to T_O^{start}, meaning that the patient will stay $T_O^{\text{start}} - T_O^{\text{MDM}}$ weeks longer in the clinic. Otherwise, the patient will leave after the initial planned T_O^{MDM} weeks. The evolution of O_t will be further discussed in Section 4.20.

Inpatient list

We describe the outpatient list at time t by I_t:

$$I_t = ((i_{\mathbf{q}\ell m})_{\mathbf{q}\in\mathcal{Q},\ell\in\mathcal{L}_{\mathcal{I}},m\in\mathcal{M}_I})_t, \tag{4.3}$$

where $i_{\mathbf{q}\ell m}$ describes the number of inpatients in the current week with treatment plan $\mathbf{q}$, the (currently set) number of weeks until discharge ℓ, and the number of times the discharge date is postponed m.

For this method, we assume that a patient's treatment plan does not change throughout their stay.

4.1.3 Action space

An action $\mathbf{a} = (a_{u\mathbf{p}})_{u\in\mathcal{U},\mathbf{p}\in\mathcal{P}}$ describes the number of patients with urgency level u and (initial) treatment plan $\mathbf{p}$ that are chosen to be admitted from the waiting list. These patients' first appointment is in week $t+1$ if we are currently in week t.

We cannot admit more patients with an urgency level u and (initial) treatment plan $\mathbf{p}$ than those who are on the waiting list. So, we define the action space $\mathcal{A}(S_t)$ for a state $S_t \in \mathcal{S}$, by (4.4).

$$\mathcal{A}(S_t) = \left\{ (\mathbf{a}) \;\middle|\; \begin{array}{ll} a_{u\mathbf{p}} \leq (h_{u\mathbf{p}})_t & \forall u \in \mathcal{U}, \mathbf{p} \in \mathcal{P} \\ a_{u\mathbf{p}} \in \mathbb{N}_0 & \forall u \in \mathcal{U}, \mathbf{p} \in \mathcal{P} \end{array} \right\} \tag{4.4}$$

4.1.4 Transition probabilities

After an action $\mathbf{a}$ is chosen in week t and the week starts, new information comes in about changes in the treatment process of the current patients and new patients' arrivals, inpatient and outpatient. We assume that patients arrive on Monday and are discharged on Friday. We describe the different transition parts in more detail below.

- *Admit chosen patients from waiting list*
  The action determines which patients from the waiting list are admitted. These patients are removed from the waiting list, and we can describe the post-decision waiting list as $H_t^{\mathbf{a}}$:

$$H_t^{\mathbf{a}} = H_t - \mathbf{a}(S_t). \tag{4.5}$$

  The patients are admitted the next week and are therefore added to the list of outpatients, with $m = 0$, and $\ell = T_O^{\text{start}}$.

$$((o_{\mathbf{p}\ell m})^{\mathbf{a}})_t = \begin{cases} a_{u\mathbf{p}'}, & \text{if } \ell = T_O^{\text{start}}, \mathbf{p} = \mathbf{p}' \text{ and } m = 0, \\ 0, & \text{otherwise,} \end{cases} \tag{4.6a}$$

$$O_t^{\mathbf{a}} = \left((o_{\mathbf{p}\ell m})^{\mathbf{a}}_{\mathbf{p}\in\mathcal{P},\ell\in\mathcal{L}_O,m\in\mathcal{M}_O}\right)_t. \tag{4.6b}$$

- *Arrival of new patients on waiting list*
  $B_{u\mathbf{p}} \sim \text{Pois}(\mu_{u\mathbf{p}})$ is a random variable describing the number of new arrivals of outpatients to the waiting list with urgency u and treatment plan $\mathbf{p}$, which is determined by a Poisson distribution with mean λ_O and distribution probability $p_{\mathbf{p}}^{\text{plan}}$. We describe the set of patients entering the waiting list in the next state by $\widehat{H}_{t+1}$:

$$\widehat{H}_{t+1} = (B_{up})_{u\in\mathcal{U},\mathbf{p}\in\mathcal{P}}. \tag{4.7}$$

  Note that this definition does not limit the number of patients on the waiting list.

- *Urgency of waiting patients increases*
  A patient's urgency is increased when the patient is not admitted. We increase a patient's urgency level by one if it is not admitted, if the patient is not yet at the highest urgency level, U. We describe the transition of the urgency level of patients on the waiting list who are not admitted by $\widetilde{H}_t$:

$$((\widetilde{h}_{u'\mathbf{p}})^{\mathbf{a}})_t = \begin{cases} (h^{\mathbf{a}}_{(u'-1)\mathbf{p}})_t, & \text{if } u' < U\ , \\ (h^{\mathbf{a}}_{u'\mathbf{p}})_t, & \text{if } u' = U\ , \end{cases} \tag{4.8a}$$

$$\widetilde{H}_t^{\mathbf{a}} = ((\tilde{h}_{u\mathbf{p}})^{\mathbf{a}}_{u\in\mathcal{U},\mathbf{p}\in\mathcal{P}})_t. \tag{4.8b}$$

  The waiting list in the next state is now defined combined the new arrivals, (4.7), and the transition to the next urgency level of the patients that were not admitted, (4.8):

$$H_{t+1} = \widetilde{H}_t^{\mathbf{a}} + \widehat{H}_{t+1}. \tag{4.9}$$

- *Discharge of outpatients*
When an outpatient is $\ell = T_O^{\text{MDM}}$ weeks from the set discharge moment, the practitioners either decide to discharge the patient T_O^{MDM} weeks later or to continue treatment for T_O^{start} weeks longer.

We describe the number of decisions that take place for one patient as a terminating Markov chain, which is a Markov chain where all states are transient except one which is absorbing. If the patient is discharged directly, the patient is admitted for T_O^{start} weeks in total. If the patient's length of stay is increased at the first MDM, the patient is admitted for $T_O^{\text{start}} + (T_O^{\text{start}} - T_O^{\text{MDM}})$ weeks, et cetera. This is visualized in Figure 4.1.

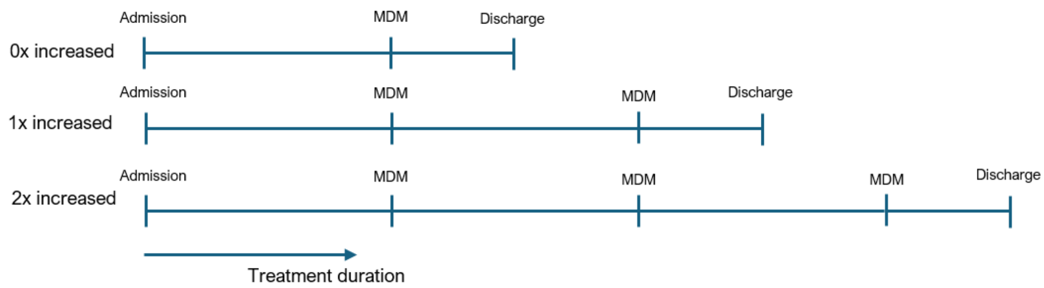


Figure 4.1: Visualization of the treatment duration of a patient whose treatment is not increased, once increased and twice increased.

Let $m_O^{\max}$ be the maximum number that the discharge date can be moved forward. The transition probability matrix of such a Markov chain is defined as

$$P^{\text{cont}} = \begin{bmatrix} \mathbf{T} & \mathbf{T_0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} := \begin{bmatrix} \mathbf{T} & (\mathbf{I}-\mathbf{T})\mathbf{1} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \tag{4.10}$$

where $\mathbf{T}$ is an $(m_I^{\max} - 1) \times (m_I^{\max} - 1)$ matrix whose only non-zero elements are on the super diagonal, such that $T_{ij} = 0$ if $i+1 \neq j$ and $T_{ij} = p_i^{\text{cont}}$ if $i+1 = j$, where p_i^{cont} is the probability that a patients' treatment is continued, where he already had i MDMs before.

Note that if we assume that a patient's treatment plan does not change, we can describe the transition function of the outpatient list. We denote $\tilde{O}_t$ as the outpatient list matrix after all transitions but without the new arrivals. For each element in the outpatients list matrix, the transition to the next state is defined by a binomial distribution.

$$\mathbb{P}((\widetilde{o}_{\mathbf{p}\ell' m})_{t+1} = x) = \begin{cases} f(x, (o_{\mathbf{p}\ell(m-1)})_t, T_{m-1,m}), & \text{if } \ell = T_O^{\text{MDM}}, \ell' = T_O^{\text{start}}, \\ & \quad m = 1, \dots m_O^{\max}, \\ 1 - f(x, (o_{\mathbf{p}\ell(m-1)})_t, T_{m-1,m}), & \text{if } \ell = T_O^{\text{MDM}}, \ell' = T_O^{\text{start}}, \\ & \quad m = 1, \dots m_O^{\max}, \\ 1 & \text{if } x = (o_{\mathbf{p}(\ell+1)m})_t, \\ & \quad \ell \neq T_O^{\text{MDM}}, \ell \neq 0 \\ 0 & \text{otherwise} \end{cases}, \forall \mathbf{p} \in \mathcal{P} \tag{4.11}$$

where $f(x, n, p)$ denotes the probability of exactly getting x successes in n independent Bernoulli trials, all with rate p, which is defined as follows:

$$f(x, n, p) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}. \tag{4.12}$$

We can denote the transition of the outpatient list, without the admission of new outpatients by $\widetilde{O}_{t+1}$:

$$\widetilde{O}_{t+1} = \left((o_{\mathbf{p}\ell m})_{\mathbf{p}\in\mathcal{P}, \ell\in\mathcal{L}_\mathcal{O}, m\in\mathcal{M}_O}\right)_{t+1}. \tag{4.13}$$

Now, we have the transition of the outpatient list, defined by O_{t+1}:

$$O_{t+1} = O_t^{\mathbf{a}} + \widetilde{O}_{t+1}. \tag{4.14}$$

Average length of stay

The initial state probabilities τ are defined as $\tau = \begin{bmatrix} 1 & \mathbf{0} \end{bmatrix}$, stating that a patient always starts in state 1, equal to $m = 0$.

Using the marginal state probabilities, where X_k denotes the state a time k

$$\begin{aligned} \begin{bmatrix} \mathbb{P}(X_k = 1) \dots \mathbb{P}(X_k = m_I^{\text{max}}) \end{bmatrix} &= \sum_{i=1}^{m_I^{\text{max}}} \mathbb{P}(X_k = j | X_0 = i)) \mathbb{P}(X_0 = i) \\ &= \begin{bmatrix} \tau & 1 - \tau \end{bmatrix} \begin{bmatrix} \mathbf{T}^k & (\mathbf{I} - \mathbf{T}^k)\mathbf{1} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \\ &= \begin{bmatrix} \tau \mathbf{T}^k & 1 - \tau \mathbf{T}^k \mathbf{1}. \end{bmatrix} \end{aligned}$$

We know that the probability that the patient is discharged after n MDMs is given by

$$\mathbb{P}(X_k = m_I^{\text{max}}) = 1 - \tau \mathbf{T}^k \mathbf{1}.$$

Let $K = \min\{k = 0, 1, 2, \dots | X_k = n\}$ and note that

$$X_k = n \Leftrightarrow K \leq k.$$

Then, we can determine the probability that a new patient will stay for n MDMs, after which it is discharged:

$$\begin{aligned}\mathbb{P}(K=k) &= \mathbb{P}(K \leq k) - \mathbb{P}(K \leq k-1) \\ &= \mathbb{P}(X_k = m_I^{\text{max}}) - \mathbb{P}(X_{k-1} = m_I^{\text{max}}) \\ &= (1 - \tau \mathbf{T}^k \mathbf{1}) - (1 - \tau \mathbf{T}^{k-1}\mathbf{1}) \\ &= \tau(\mathbf{T}^{k-1} - \mathbf{T}^k)\mathbf{1} \\ &= \tau \mathbf{T}^{k-1}(I - \mathbf{T})\mathbf{1} \\ &= \tau \mathbf{T}^{k-1}\mathbf{T}_0.\end{aligned} \tag{4.15}$$

And so, the average number of MDMs that one patient goes through is defined by:

$$\mathbb{E}(K) = \sum_{k=1}^{\infty} k\mathbb{P}(K=k). \tag{4.16}$$

- *Arrival and discharge of inpatients*
  Inpatients are handled similarly to outpatients. However, as inpatients are admitted directly, we do not have to take into account the action that was chosen. We define the number of decision moments an inpatient will have during its stay by an absorbing discrete-time Markov chain, whose transition matrix is defined as in (4.10). Then, the transition to the next state of the inpatient list is defined by $\tilde{\imath}_t$:

$$\mathbb{P}((\tilde{\imath}_{\mathbf{q}\ell' m})_t = x) = \begin{cases} f(x, (i_{\mathbf{q}\ell(m-1)})_t, T_{m-1,m}), & \text{if } \ell = T_I^{\text{MDM}}, \ell' = T_I^{\text{start}}, \\ & \quad m = 1, \dots m_I^{\max}, \\ 1 - f(x, (i_{\mathbf{q}\ell(m-1)})_t, T_{m-1,m}), & \text{if } \ell = T_I^{\text{MDM}}, \ell' = T_I^{\text{MDM}} - 1, \\ & \quad m = 1, \dots m_I^{\max}, \\ 1, & \text{if } x = (i_{\mathbf{q}(\ell+1)m})_t, \\ & \quad \ell \neq T_I^{\text{MDM}}, \ell \neq 0, \\ 1, & \text{if } x = 0, \\ & \quad \ell = T_I^{\text{start}}, m = 0, \\ 0, & \text{otherwise.} \end{cases} \quad , \forall \mathbf{q} \in \mathcal{Q} \tag{4.17}$$

where $f(x, n, p)$ is defined in (4.12), and denotes the probability of exactly getting x successes in n independent Bernoulli trials, all with rate p.

The transition of the inpatient list without new arrivals can be written as $\tilde{I}_t$:

$$\tilde{I}_t = \big((\tilde{\imath}_{\mathbf{q}\ell m})_{\mathbf{q}\in\mathcal{Q}, \ell\in\mathcal{L}_{\mathcal{I}}, m\in\mathcal{M}_I}\big)_t. \tag{4.18}$$

Inpatients arrive and are discharged a couple of weeks later. We define the list of arriving inpatients in week t by $\hat{I}_t$. The number of beds available is described by $n_I^{\max}$. We denote the current number of patients admitted by $||\widetilde{I}_t||$:

$$||\widetilde{I}_t|| := \sum_{\mathbf{p}\in\mathcal{C}} \sum_{\ell\in\mathcal{L}_I} \sum_{m\in\mathcal{M}_I} (\tilde{\imath}_{\mathbf{q}\ell m})_t. \tag{4.19}$$

The probability of y inpatient arrivals with a given treatment plan is given in (4.20). The arrivals follow a Poisson distribution. If no beds are available upon arrival, the

$$\begin{aligned}\mathbb{P}((\hat{\imath}_{\mathbf{q}T_I^{\text{start}}0})_{t+1} = y) = \\ &\sum_{d=y}^{n_I^{\max}-||\widetilde{I}_t||-1} \mathbb{P}\big(D=d\big) f\big(y,d,p_{\mathbf{q}}^{\text{plan}}\big) \\ &+ \mathbb{P}\big(D \geq n_I^{\max} - ||\widetilde{I}_t||\big) f\big(y, n_I^{\max} - ||\widetilde{I}_t||, p_{\mathbf{q}}^{\text{plan}}\big),\end{aligned} \tag{4.20}$$

where D is a Poisson random variable describing the number of new inpatient arrivals on the inpatient list, $p_{\mathbf{q}}^{\text{plan}}$ is the probability that a patient follows treatment plan $\mathbf{q}$ and $f(x,n,p)$ denotes the probability of exactly getting k successes in n independent Bernoulli trials, all with rate p, as defined in (4.12).

The updated inpatient list is given by $\hat{I}_t$:

$$(\hat{\imath}_{\mathbf{q}\ell m})_{t+1} = 0 \;\; \forall \ell \neq T_I^{\text{start}}, \tag{4.21a}$$

$$\hat{I}_{t+1} = \big((\hat{\imath}_{\mathbf{q}\ell m})_{\mathbf{q}\in\mathcal{Q},\ell\in\mathcal{L}_\mathcal{I},m\in\mathcal{M}_I}\big)_{t+1}. \tag{4.21b}$$

The inpatient list for week $t+1$ is described by the arrivals during week t, and the changes regarding the current inpatient. We describe the new inpatient list by I_{t+1}:

$$\mathbb{P}((i_{\mathbf{q}\ell m})_{t+1} = z) = \begin{cases} \mathbb{P}((\tilde{\imath}_{\mathbf{q}\ell m})_t = z), & \text{if } \ell \neq T_I^{\text{start}}, m \neq 0, \; \forall \mathbf{q} \in Q, \\ \sum_{x=0}^{z} \mathbb{P}\big((\tilde{\imath}_{\mathbf{q}\ell m})_t = x\big)\mathbb{P}\big((\hat{\imath}_{\mathbf{q}\ell m})_{t+1} = z-x\big), & \text{if } \ell = T_I^{\text{start}}, m = 0, \; \forall \mathbf{q} \in Q, \end{cases} \tag{4.22a}$$

$$I_{t+1} = \big((i_{\mathbf{q}\ell m})_{\mathbf{q}\in\mathcal{Q},\ell\in\mathcal{L}_\mathcal{I},m\in\mathcal{M}_I}\big)_{t+1}. \tag{4.22b}$$

With the above descriptions, we can derive the transition function from a state S_t to a state S_{t+1} after deciding on action $\mathbf{a}(S_t)$ by $\mathbb{P}(S_{t+1}|S_t,\mathbf{a})$:

$$\mathbb{P}(S_{t+1}|S_t,\mathbf{a}) = \mathbb{P}(H_{t+1}|H_t,\mathbf{a}_t)\mathbb{P}(I_{t+1}|I_t)\mathbb{P}(O_{t+1}|O_t,\mathbf{a}), \tag{4.23}$$

where

$$\mathbb{P}(H_{t+1}|H_t,\mathbf{a}) = \begin{cases} \prod_{p\in\mathcal{P},u\in\mathcal{U}} \mathbb{P}(B_{up}), & \text{if } H_{t+1} \text{ satisfies (4.9)}, \\ 0, & \text{otherwise}, \end{cases} \tag{4.24a}$$

$$\mathbb{P}(O_{t+1}|O_t, \mathbf{a}) = \begin{cases} \prod_{\substack{\mathbf{p}\in\mathcal{P},\ell\in\mathcal{L}_{\mathcal{O}},\\ m\in\mathcal{M}_O}} \mathbb{P}((\tilde{o}_{\mathbf{p}\ell m})_{t+1} = x), & \text{if } O_{t+1} \text{ satisfies (4.14)},\\ 0, & \text{otherwise,}\end{cases} \tag{4.24b}$$

$$\mathbb{P}(I_{t+1}|I_t) = \begin{cases} \prod_{\substack{\mathbf{q}\in\mathcal{Q},\ell\in\mathcal{L}_{\mathcal{I}},\\ m\in\mathcal{M}_I}} \mathbb{P}((i_{\mathbf{q}\ell m})_{t+1} = z), & \text{if } I_{t+1} \text{ satisfies (4.22b)},\\ 0, & \text{otherwise.}\end{cases} \tag{4.24c}$$

4.1.5 Cost function

The direct costs associated with a state-action pair $S_t \in S, \mathbf{a} \in \mathcal{A}(S_t)$ has two components: costs for patients on the waiting list and for exceeding the capacity of the combined demand from both in- and outpatients.

(i) The higher the urgency of a waiting patient, the higher the cost if a patient is not admitted. We define $f_u(S_t, \mathbf{a})$ to be the urgency cost function. We define this function in 4.25.

$$f_u(S_t, \mathbf{a}) = \kappa_u \sum_{\mathbf{p}\in\mathcal{P}} (h_{u\mathbf{p}})_t - a_{u\mathbf{p}}, \tag{4.25}$$

where κ_u is a $1 \times U$ vector with the u-the element defining the weight of not admitting a patient with urgency level u.

(ii) Exceeding capacity results in costs that are determined by the current state, which is influenced by previous actions. This is evaluated by counting the number of therapy sessions that go beyond the capacity, encompassing both inpatient and outpatient sessions. We denote the cost function of exceeding demand with $f_d(S_t)$. We determine the cost over the current week t.

We define the cost associated with overusing capacity as follows:

$$f_d(S_t) = \kappa_c \big[\sum_{\mathbf{q}\in\mathcal{Q}} \mathbf{q} \sum_{\substack{\ell\in\mathcal{L}_I,\\ m\in\mathcal{M}_I}} (i_{\mathbf{q}\ell m})_t + \sum_{\mathbf{p}\in\mathcal{P}} \mathbf{p} \sum_{\substack{\ell\in\mathcal{L}_O\\ m\in\mathcal{M}_O}} (o_{\mathbf{p}\ell m})_t - \mathbf{r} \big]^+, \tag{4.26}$$

where κ_c describes the factor of exceeding slots of discipline $c \in \mathcal{C}$, $\mathbf{r}$ is the set capacity, and $[a]^+ = \max(0, a)$.

We describe the full cost function in (4.27):

$$C(S_t, \mathbf{a}) = f_u(S_t, \mathbf{a}) + f_d(S_t). \tag{4.27}$$

4.1.6 The Bellman equations

We consider the cost incurred by the action taken this week more important than those incurred the weeks after. Therefore, we introduce a discount factor $\gamma \in [0, 1)$, to evaluate later costs at this week's level. The value function $v_\gamma^\pi(S)$ describes the total expected discounted cost (with discount factor γ) for a state $S \in \mathcal{S}$ over the infinite horizon under policy π, which tells the planner which action $\mathbf{a}$ to take in the given state. For an optimal policy π^*, we have

$$v_\gamma^{\pi^*}(S_t) \geq v_\gamma^\pi(S) \qquad\qquad \forall S \in \mathcal{S}, \pi \in \Pi$$

To find such an optimal policy, we need to solve the optimality equations (also known as the Bellman equations):

$$v(S) = \min_{\mathbf{a} \in \mathcal{A}(S)} \left\{ C(S, \mathbf{a}) + \gamma \sum_{S' \in S} \mathbb{P}(S'|S, \mathbf{a}) v(S') \right\}, \ \forall S \in \mathcal{S}. \tag{4.28}$$

4.2 Finding exact solutions for small instances

From the state space definition, we note that if we limit the maximum number of patients on the waiting list and in both the inpatient and outpatient clinic, the size of the state space is finite $\forall S \in \mathcal{S}$.

The action space, defined in (4.4), is finite $\forall S \in \mathcal{S}$, as the number of patients on the waiting list bounds the number of actions for a state. The transition function is stationary, e.g., they do not change between decision epochs, and the direct costs $c(S, \mathbf{a})$ are bounded ($c(S, \mathbf{a}) \leq Z < \infty$ for some Z).

With these conditions fulfilled, we can apply Theorem 4.2.1 to the model, so a unique solution exists to the optimality equations.

Theorem 4.2.1 (*Theorem 6.2.5 of [27]*)**.** *Suppose $0 \leq \gamma < 1$, $\mathcal{S}$ is finite and $c(S, \mathbf{a})$ is bounded for all $S \in \mathcal{S}, \mathbf{a} \in \mathcal{A}(S)$. Then there exists an unique solution $v^* = \{v^*(S)\}_{S \in \mathcal{S}}$ to the optimality equations* (4.28)*.*

Proof. See the proof of Theorem 6.2.5 on p.151 of [27]. □

Theorem 4.2.2 (*Theorem 6.2.10 of [27]*)**.** *Assume S is discrete, and either*

- *(a) $\mathcal{A}(S)$ is finite for each $S \in \mathcal{S}$, or*
- *(b) $\mathcal{A}(S)$ is compact, $c(S, \mathbf{a})$ is continuous in a for each $S \in \mathcal{S}$, and, for each $S' \in \mathcal{S}$ and $S \in \mathcal{S}$, $\mathbb{P}(S'|S, \mathbf{a})$ in continuous in $\mathbf{a}$, or*
- *(c) $\mathcal{A}(S)$ is compact, $c(S, \mathbf{a})$ is upper semicontinuous in $\mathbf{a}$ for each $S \in \mathcal{S}$, and for each $S' \in \mathcal{S}$ and $S \in \mathcal{S}$, $\mathbb{P}(S'|S, \mathbf{a})$ is lower semicontinuous in $\mathbf{a}$.*

Then, an optimal deterministic stationary policy exists.

Proof. See the proof of Theorem 6.2.10 on p.154-155 of [27]. □

If we assume that the model fulfills the conditions, for such a case as we explained above, it follows from Theorem 4.2.2 that there exists an optimal deterministic stationary policy π^*. The corresponding policy solves the optimality equations (4.28). Hence the optimal decision rule is given the function $d : \mathcal{S} \to \mathcal{A}(S)$:

$$d(S) = a_S^* \in \arg\min_{\mathbf{a} \in \mathcal{A}(S)} \left\{ c(S, \mathbf{a}) + \gamma \sum_{S' \in S} \mathbb{P}(S'|S, \mathbf{a}) v^*(S') \right\}. \tag{4.29}$$

Linear programming

To find this optimal policy, we can use a linear program (LP). The primal LP is formulated in (4.30). Note that this LP is based upon the formulation given on page 223 of [27], but as we minimize with a cost function instead of maximize with a reward function, the inequality sign is flipped and the objective function is maximized instead of minimized.

$$\max_{\mathbf{v}} \sum_{S\in\mathcal{S}} v(S)\alpha(S) \tag{4.30a}$$

$$\text{s.t.} \quad v(S) \leq c(S,\mathbf{a}) + \gamma \sum_{S'\in\mathcal{S}} \mathbb{P}(S'|S,\mathbf{a})v(S'), \qquad \forall S \in \mathcal{S}, \mathbf{a} \in \mathcal{A}(S). \tag{4.30b}$$

The value of $\alpha(S)$ represents the weight of state S in the objective function and should satisfy $\sum_{S\in\mathcal{S}} \alpha(S) = 1$ and $\alpha(S) > 0$. This LP has $|S|$ variables and $\sum_{S\in\mathcal{S}} |\mathcal{A}_s|$ constraints. Hence, we prefer to use the dual of this LP to find a solution to the LP more easily. The dual of the LP is given in (4.31).

$$\min_{\mathbf{x}} \sum_{S\in\mathcal{S}} \sum_{\mathbf{a}\in\mathcal{A}(S)} c(S,\mathbf{a})x(S,\mathbf{a}) \tag{4.31a}$$

$$\text{s.t.} \sum_{\mathbf{a}\in A(S')} x(S',\mathbf{a}) - \sum_{S\in\mathcal{S}} \sum_{\mathbf{a}\in\mathcal{A}(S)} \gamma\mathbb{P}(S'|S,\mathbf{a})x(S,\mathbf{a}) = \alpha(S') \quad \forall S' \in \mathcal{S}. \tag{4.31a}$$

The dual LP has $\sum_{S\in\mathcal{S}} |\mathcal{A}_s|$ variables and the same amount of constraints.

4.2.1 Determining the problem size

First, we note that by theorem 4.2.1, finding exact solutions for the defined problem is possible for a finite state space. If we set limits on the size of the waiting and inpatient lists and set a maximum for the number of MDMs a patient can endure, we note that the state space for our model is finite.

Using a linear program to solve the problem is very difficult when the size of state space and state-action space is large, which is described by Bellman as the *curse of dimensionality* [28]. In this section, we determine the size of the state space of our model and provide an example for a chosen system setting.

Consider the waiting list, where we set the maximum number of patients on the list to be $n_H^{\max}$. But if we assume this limit, the number of possible waiting lists is defined by

$$|H| = \sum_{i=0}^{n_H^{\max}} \binom{|\mathcal{P}|U}{i}. \tag{4.32}$$

Consider the list of inpatients. A maximum of $n_I^{\max}$ patients can be admitted to the clinic. The construction of the inpatients list is very similar to that of the waiting list. We have the following number of possible combinations:

$$|I| = \sum_{i=0}^{n_I^{\max}} \binom{|\mathcal{Q}||\mathcal{L}_I|m_I^{\max}}{i}. \tag{4.33}$$

Consider the list of outpatients. In theory, no maximum of outpatients can be admitted. However, in reality, often, one can determine a realistic maximum, which we define as $n_O^{\max}$. So, we have the following number of possible combinations possible:

$$|O| = \sum_{i=0}^{n_O^{\max}} \binom{|\mathcal{P}||\mathcal{L}_O|m_O^{\max}}{i}. \tag{4.34}$$

The total size of the state space will be

$$|\mathcal{S}| = |H||I||O|. \tag{4.35}$$

Furthermore, to use the linear program to solve the program, for every state, one needs to define and determine the value of all states that this state can transition to. The set of all possible states that a state S_t can transition to, given a feasible action $\mathbf{a}$ is defined by

$$\mathcal{S}_{t+1}(S_t, \mathbf{a}) = \left\{ (S_{t+1}) \;\middle|\; \begin{array}{l} S_{t+1} \in \mathcal{S} \\ \mathbb{P}(S_{t+1}|S_t, \mathbf{a}) > 0 \end{array} \right\}.$$

The size of this space can be determined if we assume a certain maximum number of arriving outpatients. We define these values as n_O^{new}, n_I^{new}. Let n_I^{dec} be the number of inpatients that have a decision moment in the current week t, similarly n_O^{dec} for outpatients. The size of the state action space can be calculated as follows:

$$|\mathcal{S}_{t+1}(S_t, \mathbf{a})| = \sum_{n=0}^{n_O^{\text{new}}} |\mathcal{P}|^n \sum_{n=0}^{n_I^{\text{new}}} |\mathcal{Q}|^n |\mathcal{P}|^{n_O^{\text{dec}}} |\mathcal{Q}|^{n_I^{\text{dec}}}. \tag{4.36}$$

This indicates that even for smaller configurations of the model, both the state space and state-action space are considerably large, as detailed in the following section.

Problem size for a small system

We define a relatively small system, similar to the system in Section 6.3.2. The corresponding parameters are defined in Table 4.2. We set its $n_H^{\max} = 5$, and $n_I^{\max} = 14$.

Table 4.2: Settings for a small system model, single-discipline with finite bed capacity.

$C = 1$	
$U = 3$	
$r = \{120\}$	
$\mathcal{P} = \{\{1\}, \{2\}, \{3\}\}$	$\mathcal{Q} = \{\{2\}, \{3\}\}$
$p_{\mathcal{P}}^{\text{plan}} = \{0.6, 0.2, 0.2\}$	$p_{\mathcal{Q}}^{\text{plan}} = \{0.25, 0.75\}$
$T_O^{\text{MDM}} = 2$	$T_I^{\text{MDM}} = 2$
$T_O^{\text{start}} = 8$	$T_I^{\text{start}} = 5$
$\lambda_O = 1$	$\lambda_I = 2$
$p_O^{\text{cont}} = \{0.8, 0.6, 0\}$	$p_I^{\text{cont}} = \{0.6, 0.4, 0\}$
	$n_I^{\max} = 14$

Then, the size of the state space of this system is

$$|\mathcal{S}| \approx 382 \times 4 \cdot 10^8 \times 4 \cdot 10^{14} = 6.1 \cdot 10^{25}.$$

Given a state $S \in \mathcal{S}$ in this system and action $\mathbf{a} \in \mathcal{A}(S)$, what is the number of possible states it can transition to? Note that for a random Poisson variable X with mean $\lambda = 1$, we have $\mathbb{P}(X > 4) < 0.01$. Similarly, $\mathbb{P}(X > 6) < 0.01$, for a random Poisson variable X with $\lambda = 2$.

Therefore, in this calculation, we set $n_O^{\text{new}} = 4, n_I^{\text{new}} = 6$. As an example, this state currently has two inpatients and one outpatient that have a decision moment, so $n_I^{\text{dec}} = 2, n_O^{\text{dec}} = 1$. Then, the size of the state action space is,

$$121 \cdot 127 \times 3^1 \times 2^2 = 184404.$$

In conclusion, we have determined that the state space for this relatively small system and the potential number of transitions for a probable state are both extensive. This complexity makes it difficult to use linear programming to solve the system computationally impractical.

Chapter 5

Solution approach

As we have shown in chapter 4, determining an optimal policy using the Linear Program as in (4.30) or (4.31) is only possible for very small instances of the problem. For this problem, even a small system results in large state and state-action spaces. Thus, we must employ approximation techniques to identify (near-)optimal policies for instances that more closely resemble reality. In Section 5.1, we describe Powell's framework of the four classes of policies in more detail. We provide a detailed overview of the policy we have developed in Section 5.2.

5.1 Four classes of policies

As already mentioned in Section 3.2, Powell describes two strategies for finding a policy to solve a sequential decision problem in [30]. Each strategy consists of two policies, resulting in the four classes of policies. These strategies are policy search, including Policy Function Approximation (PFA) and Cost Function Approximation (CFA), and lookahead approximations split into Value Function Approximations (VFA) and Direct Lookahead Approximations (DLA). All classes of policies boil down to approximating a part of the model to avoid the *curse of dimensionality*. Such that the MDP can solve while circumventing constraints imposed by computational limitations. VFA and DLA describe a more natural strategy to consider the downstream impact of a decision in the current decision epoch. In contrast, for PFA and CFA, you consider a tunable policy. Below, we briefly describe each class separately.

PFAs and CFAs

A PFA is any analytical function mapping a state to an action using lookup tables, parametric functions, or non-parametric functions. PFA is similar to machine learning, as both try to approximate a function replicating reality. The only difference is the objective function. For a linear policy approximation function, we can write:

$$A^{\mathsf{PFA}}(S_t|\theta) = \sum_{f\in\mathcal{F}} \theta_f \phi_f(S_t),$$

where $(\phi_f(S))_{f\in\mathcal{F}}$ are the features on which the approximation is based, and $(\theta_f)_{f\in\mathcal{F}}$ are the parameter that gives weight to each of the features. These features need to be humanly defined in advance and describe certain information about the system's current state. PFAs are the most straightforward class of policies and easiest to compute, but one needs to have a preliminary understanding of the most effective policy. Tuning can be done offline or online, with the latter having the advantage that the transition function does not need to be known.

CFAs are parametrically modified optimization problems. The modifications can be applied to the constraints of the action space and/or the objective function. We define the tunable parameters by θ. N and the policy can be written as in (5.1):

$$A^{\text{CFA}}(S_t|\theta) = \arg\min_{\mathbf{a}\in\bar{\mathcal{A}}^{\pi}(\theta)} \bar{C}^{\pi}(S_t, \mathbf{a}|\theta), \tag{5.1}$$

where $\bar{\mathcal{A}}^{\pi}(\theta)$ is the modified set of constraints determined by policy π with tunable parameters θ, and $\bar{C}_t^{\pi}(S_t, \mathbf{a}(S_t)|\theta)$ is the modified objective function as determined by policy π with tunable parameters θ. Note that θ can describe multiple parameters, depending on the detailed policy description. There are several options available for modifying both the feasible region and the objective function. An example of the modified objective function with a linear correction factor defined by a set of features $\mathcal{F}$, described by both that state as the action chosen, is:

$$\bar{C}_t^{\pi}(S_t, \mathbf{a}|\theta) = C(S_t, \mathbf{a}) + \sum_{f\in\mathcal{F}} \theta_f \phi_f(S_t, \mathbf{a}).$$

where $\theta = (\theta_f)_{f\in\mathcal{F}}$ are parameters describing the weight for each of the features.

PFA and CFA policies both involve tuning of the parameters, which are often described by θ. To tune the parameters θ for the policy search policies, we can use a stochastic search method:

$$\min_{\theta} \mathbb{E}\big\{ \sum_{t=0}^{T} C(S_t, \mathbf{A}^{\pi}(S_t|\theta))|S_0 \big\}, \tag{5.2}$$

where S_0 is the initial state at $t = 0$, and T is the length of the considered timeinterval. Extensive literature on stochastic search methods can be found. We will return to this topic later on in section 5.3.1.

VFAs and DLAs

An optimal policy for a lookahead approximation strategy is generally written as

$$\mathbf{A}^*(S_t) = \arg\min_{\mathbf{a}\in\mathcal{A}(S_t)} \left(C_t(S_t, \mathbf{a}) + \mathbb{E}\left\{ \min_{\pi\in\Pi} \mathbb{E}\left\{ \sum_{t'=t+1}^{T} C_{t'}\big(S_{t'}, \mathbf{A}^{\pi}(S_{t'})\big)\big|S_{t+1} \right\} \big| S_t, \mathbf{a} \right\} \right), \tag{5.3}$$

which is computationally intractable for any realistic problem as one needs to track all possible options of action and arriving exogenous information to compute the expectation. Therefore, these classes of policies try to approximate the value in the expectation, and so the name 'lookahead' approximation.

In the first class, VFA, the expectation is substituted by a value approximating the value of being in a post-decision state after taking the action $\mathbf{a}$ in state S_t, and the policy is defined as:

$$\mathbf{A}^{VFA}(S_t) = \arg\min_{\mathbf{a}\in\mathcal{A}(S_t)} \big(C_t(S_t, \mathbf{a}) + \bar{V}^{\mathbf{a}}(S_t^{\mathbf{a}})\big), \tag{5.4}$$

where $\bar{V}^{\mathbf{a}}(S_t^{\mathbf{a}})$ is the approximation of the value function around the post-decision state. This approximation can be done by a linear model, separable piecewise linear functions, or Benders cuts. Although VFAs are widely studied in the literature, for example, Powell [29] and Sutton et al. [42], they can only be applied to a small set of problems. It works well for relatively simple inventory problems, 'single layer inventory problems', such as in Philpott et al. [43]. Often some structure needs to be exploited in order to use a VFA.

More complex time-dependent problems are often solved by an approximate lookahead model on a rolling basis. The general idea is to approximate the expectation by defining another (sub)problem, typically for an interval $[t, t+K], K > 0$. The challenge for this class is handling uncertainty as we plan over this shorter horizon. Also, for this policy class, one can consider multiple options, including stochastic programming and deterministic lookaheads, of which the general structures are given in (5.5) and (5.6).

$$A^{\text{S-DLA}}(S_t) = \arg\min_{\substack{\forall\omega\in\Omega_{[t,t+K]}, \widetilde{\mathbf{a}}_t\in\mathcal{A}(S_t),\\ \left(\widetilde{\mathbf{a}}_{t'}(\omega)\in\mathcal{A}(\widetilde{S}_{tt'}(\omega))\right)_{t'=t+1}^{t+K}}} \Big(C(\widetilde{S}_{tt}, \widetilde{\mathbf{a}}_t) + \sum_{\omega\in\Omega_{[t,t+K]}} C(\widetilde{S}_{tt'}(\omega), \widetilde{\mathbf{a}}_{t'}(\omega))\Big), \tag{5.5}$$

where ω is one of the scenarios in the set Ω_t of all possible scenarios of arriving exogenous information in the interval $[t, t+K]$, with $K > 0$, and $\widetilde{S}_{tt'}(\omega)$ describes the state forecast for week t', if it is currently week t, and $\mathbf{a}_{t'}$ describes the actions that is taken at week t'.

Note that the S-DLA is typically $|\Omega_t|$ times bigger than the D-DLA as one needs to iterate over all possible scenarios of the future in the upcoming (small) time interval.

Hybrid policies: parametric cost function approximation with deterministic lookahead

Powell outlines four distinct strategies for solving complex MDPs. However, hybrid strategies that merge two of these approaches may also be employed. Below, we describe the combination of Deterministic Lookaheads and parametrically modified CFA (DLA - CFA), a combination introduced by Powell and Ghadimi in [32]. We used this article as the primary inspiration for the policy we developed. This hybrid policy is helpful as it can handle complex, high-dimensional state variables and avoids VFA, which is more typically used.

Consider a deterministic lookahead policy:

$$\mathbf{A}^{\text{D-DLA}}(S_t) = \arg\min_{\mathbf{a}\in\mathcal{A}(S_t)} \Big(C(S_t, \mathbf{a}) + \min_{\left(\widetilde{\mathbf{a}}_{t'}\in\mathcal{A}(\widetilde{S}_{tt'})\right)_{t'=t+1}^{t+K}} \sum_{t'=t+1}^{t+K} C(\widetilde{S}_{tt'}, \widetilde{\mathbf{a}}_{t'})\Big), \tag{5.6}$$

In the above equation, $\widetilde{S}_{tt'}$ describes the state forecast for week t' if it is currently week t, and $\mathbf{a}_{t'}$ describes the action to be taken in a week t'. The tilde denotes that the states within the time interval are fixed beforehand. Therefore, this is just a linear program, as all exogenous information is disregarded.

Now, we can modify the cost function in the lookahead part in (5.6) to include a cost function approximation. This should contribute to the uncertainty that can happen in the (near) future of the short interval. As described in the description of the CFA, the parametric terms, defined by θ, can be both in the cost function and in the constraints.

$$\mathbf{A}^{\text{DLA - CFA}}(S_t) = \underset{\mathbf{a}\in\mathcal{A}(S_t)}{\arg\min}\Big(C(S_t,\mathbf{a}) + \min_{\left(\widetilde{\mathbf{a}}_{t'}\in\bar{A}(S_t)^{\theta}\right)_{t'=t+1}^{t+K}} \sum_{t'=t+1}^{t+K} \overline{C}_{t'}(S_t,\widetilde{\mathbf{a}}_{t'}|\theta)\Big), \tag{5.7}$$

and $\overline{C}(S_t,\widetilde{\mathbf{a}}_{t'}|\theta)$ is the cost function approximation and $\bar{A}(S_t)^{\theta}$ is the refined action space for the lookahead model. When employing this model, actions for the time interval $[t, t+K]$ are defined. Only the solution for the current week t, as described by $\mathbf{a}$, is kept and performed. All $\widetilde{\mathbf{a}}_{tt'}$ for $t' \neq t$ is discarded, and we move forward in time.

Let $\bar{F}^{\pi}(\theta,\omega) = \sum_{t=0}^{T} C_t\Big(S_t(\omega), A_t^{\pi}(S_t(\omega)|\theta)\Big)|S_0$. $\bar{F}^{\pi}(\theta,\omega)$ describes the summed costs over a timeinterval of length T. To use this hybrid policy, we now need to tune parameters θ by solving:

$$\min_{\theta}\Big\{F^{\pi}(\theta) := \mathbb{E}_{\omega}[\bar{F}^{\pi}(\theta,\omega)] = \mathbb{E}\Big[\sum_{t=0}^{T} C_t\Big(S_t(\omega), A_t^{\pi}(S_t(\omega)|\theta)\Big)|S_0\Big]\Big\}, \tag{5.8}$$

where ω denotes the arrival of random exogenous information, and S_0 is the initial state. This can be done with different stochastic search methods, depending on the function $\bar{F}(\theta,\omega)$. A derivative-based and a derivative-free method are shown in [44].

How to choose a policy?

Powell's framework allows for structure in choosing a suitable solution approach for sequential decision problems. However, as Powell shows in [31], every policy class could, in theory, be applied to every problem. The small details of the problem and the solution approach determine which class is best to use. With the quote "designing policies is rather an art than a science" in his book [32], Powell addresses the difficulty and human thinking needed to achieve a sufficient result. We will further discuss our solution approach in the next section.

5.2 Detailed policy design

Our problem is not fitted for VFA, as the state space is very large, and the number of transition options for a given state-action pair is also prominent for larger instances of the model. Without state aggregation, this would be computationally intractable. Furthermore, a clear policy is difficult to pinpoint, which leaves PFA out of the options. Therefore, we continue with the combined DLA-CFA policy.

To define possible parametrizations for the DLA-CFA policy, we need to look into features that give insight into how available actions will impact further states without knowing how the states developed. We do this by evaluating the known and unknown demand for the coming weeks and the (un)certainty of arrivals on the waiting list. This results in a forecast of the demand for a timeinterval $[t, t+K]$, where the current state is week t and a forecast timeinterval of length K.

We describe the designed DLA-CFA policy in equation 5.17 in Section 5.2.2. The forecast that this policy design considers includes the arrival of inpatients and the outcomes of decision moments for both inpatients and outpatients. However, the arrival of outpatients is not taken into account as the action is dependent upon these. Therefore, we first introduce a recourse model to account for these.

5.2.1 Recourse model

We introduce a recourse model to address the uncertainty of outpatient arrivals to the waiting list within the forecast interval $[t, t+K]$. In this context, we can consider the arrivals to the waiting list within the designated forecast interval as known for the DLA-CFA policy.

The outpatient arrivals to the waiting list do not fit with the parametrization in the DLA-CFA policy, which we described in more detail in the next section, as the possible actions depend on this.

The recourse model should comprehensively evaluate all potential sample paths of arrivals to the waiting list in order to identify the most suitable action:

$$A^{\pi^{\text{recourse}}}(S_t, \theta) = \underset{\mathbf{a} \in \mathcal{A}(S_t)}{\arg\min}\, \mathbb{E}_{\omega_H \in \Omega_H} \bar{Q}^{\pi^{\text{DLA-CFA}}}(S_t, \theta, \omega_H, \mathbf{a}), \tag{5.9}$$

where $\bar{Q}^{\pi^{\text{DLA-CFA}}}(S_t, \theta, \omega_H, \mathbf{a})$ describes the costs for taking action $\mathbf{a}$ in state S_t, given parameters θ and sample path ω_H as determined by the parametrization of the costs for the forecast timeinterval. These costs are described by the costs provided by equation (5.17), which is the formulation of the DLA-CFA policy. The recourse model is an overlaying model used within the DLA-CFA policy. This implies that at each decision epoch, the recourse model is implemented, wherein every iteration of the recourse model adheres to the DLA-CFA policy.

The expectation in (5.9) can often not be determined directly as it needs to evaluate over all possible sample paths, $|\Omega_H|$, and all possible actions $|\mathcal{A}(S_t)|$. Furthermore, evaluating actions described by an integer array is challenging, making the optimization function non-convex.

We use a Sample Average Approximation (SAA) method to find the best action based on N_{batch} draws of the possible sample paths. Such a solution approach is often performed in literature, for example, by Krãtschmer [45].

To find the best action, we can draw a batch of N_{batch} sample paths and conclude the best action based on only these sample paths by selecting the most frequently occurring action.

Algorithm 1 Recourse model algorithm, using most frequency chosen action

Given a state S_t, a batch size N_{batch}
Initialize possible actions list $A_{\text{pos}} = \{\}$
for n in $1, \dots, N_{\text{batch}}$ **do**
    Generate arrivals to the waiting list in time-interval $[t, t+K]$, $\omega_n \in \Omega_H$
    Determine action based upon the generated arrivals, $\mathbf{a} = A^{\pi^{\text{DLA-CFA}}}(S_t, \theta, \omega_n)$ according to (5.17)
    Add resulting action to action list $A_{\text{pos}} = A_{\text{pos}} \cup \mathbf{a}$
end for
Return the most occurring action $\mathbf{a}^*$ in list A_{pos},

$$\mathbf{a}^* = \arg\max_{\mathbf{a}' \in A_{\text{pos}}} \sum_{\mathbf{a} \in A_{\text{pos}}} 1_{\mathbf{a}=\mathbf{a}'}$$

5.2.2 Demand estimation

Utilizing a recourse model, we have organized the arrivals to the waiting list. Now we consider the external information related to new inpatients and the outcomes of the decision moments for both patient groups.

Given a state, we note that a considerable part of the demand for the next few weeks is known. Figure 5.1 gives an example of the known demand for some model settings. To manage capacity effectively, we must consider the uncertain future demand resulting from the remaining unknown transitions.

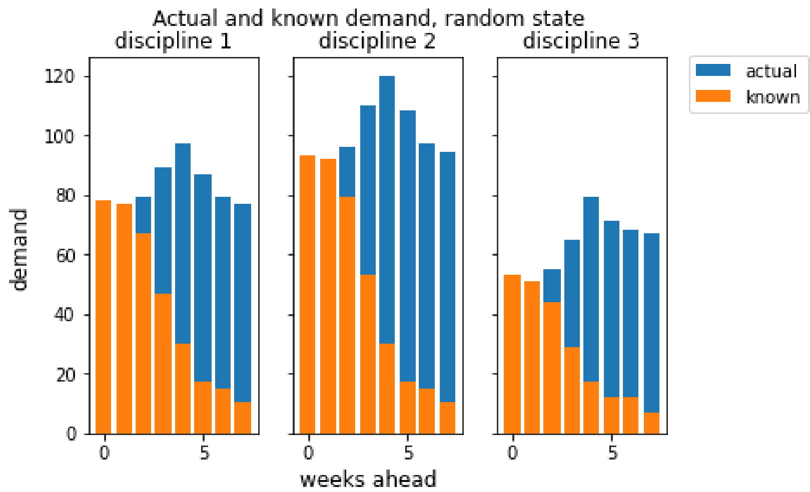


Figure 5.1: Simulated example of the actual and known demand for the next eight weeks, visualized for a random state.

To determine this (orange) demand, we assume that every decision moment results in a ‘discharge’, and do not consider any new inpatients. We define the function $d(S_t, \tau)$ that describes this demand for week $t+\tau$ and current state S_t, resulting in a $\mathbb{N}_0^C$ vector describing the known demand in week τ from the current week t.

$$d(S_t,\tau) = \sum_{\substack{\mathbf{q}\in\mathcal{Q},\\ m\in\mathcal{M}_I}} \mathbf{q}(\sum_{\ell=\tau}^{T_I^{start}} i_{\mathbf{q}\ell m})_t + \sum_{\substack{\mathbf{p}\in\mathcal{P}\\ m\in\mathcal{M}_O}} \mathbf{p} \sum_{\ell=\tau}^{T_O^{start}} (o_{\mathbf{p}\ell m})_t \tag{5.10}$$

Typically, this demand does not increase because we know the patients who will leave in the next week, but we do not know the arrival of new inpatients and the possible extensions of a patient's stays.

The arrival of inpatients and potential extensions of their stay are uncontrollable factors. The admission of outpatients is the only aspect that a planner can influence in the system. Based on our knowledge of the model's transition function, but without actually using this function, we can 'forecast' the extra demand that will occur in the next couple of weeks to result in the estimation of the demand for week τ from the current week t, defined by $\tilde{d}(S_t,\theta_\tau,\tau)$. This estimation can be incorporated into a DLA-CFA policy. One potential approach is to introduce a parameter $\theta \in \mathbb{R}^C$ to describe the extra unknown demand, without the performed action, at time $t+\tau$. This can be done via various definitions:

$$\tilde{d}(S_t,\theta_\tau,\tau) = \theta_\tau + d(S_t,\tau),$$

or even

$$\tilde{d}(S_t,\theta,\tau) = \tau\theta + d(S_t,\tau),$$

For the first, θ differs per τ, while the latter only uses one θ parameter, which is increased the further the estimation. This allows for less tuning, with the downside of less precision than the first option.

The effectiveness of this approach may vary based on system parameters, given that it does not incorporate key system insights, such as decision moments and the maximum number of inpatients. We have established three parameters that are dependent on the external arriving information relevant to the designed MDP framework.

Use of decision moment

For a given state S_t, we can determine the number of patients with a decision moment within t and $t+\tau$. Remind, in the MDP, the system parameters p_n^{cont} determine the probability that a patient's discharge date will be postponed if this has already occurred n times.

We define the feature functions $\phi_O^{\text{dec}}(S_t,\tau)$ and $\phi_I^{\text{dec}}(S_t,\tau)$ for outpatients and inpatients, which describe the demand of patients that will have a decision moment in the upcoming weeks, and possibly result in extra demand at time τ:

$$\phi_I^{\text{dec}}(S_t,\tau) = \begin{cases} 0, & \text{if } \tau \le T_I^{\text{MDM}}, \\ \sum_{\mathbf{q}\in\mathcal{Q},m\in\mathcal{M}_I} \mathbf{q} \sum_{\ell=T_I^{\text{MDM}}}^{\tau} (i_{\mathbf{q}\ell m})_t, & \text{otherwise,} \end{cases} \tag{5.11}$$

$$\phi_O^{\text{dec}}(S_t,\tau) = \begin{cases} 0, & \text{if } \tau \leq T_O^{\text{MDM}}, \\ \sum_{\mathbf{p}\in\mathcal{P},m\in\mathcal{M}_O} \mathbf{p} \sum_{\ell=T_O^{\text{MDM}}}^{\tau} (o_{\mathbf{p}\ell m})_t, & \text{otherwise}, \end{cases} \tag{5.12}$$

With the feature functions defined above, we can define an estimation of the demand at week $t+\tau$ being in week t for the new demand as a result of decision moments.

$$\theta^{\text{dec},I}\phi_I^{\text{dec}}(S_t,\tau) + \theta^{\text{dec},O}\phi_O^{\text{dec}}(S_t,\tau),$$

where $\theta^{\text{dec},I}$ and $\theta^{\text{dec},O}$ describe the ratio of the total demand describing the demand forecast related to the decision moment. Suppose θ is defined as less than average. In that case, the model induces a risk by lowering the estimated demand and allowing for more (outpatient) admissions or vice versa in case of a higher θ.

New inpatients

New inpatients can arrive every week. Formulating an estimation of these new patients' demand can be challenging, especially if the possible treatment plans are diverse. To keep it simple, we define a parameter θ^{new}, multiplied by τ if estimating τ weeks ahead.

Demand forecast

Combining the forecast of demand for the new arrivals with the forecast related to the decision moments, we update the estimation of the demand forecast to be used at week τ from now, which we describe by $\tilde{d}(S_t,\theta,\tau)$:

$$\tilde{d}(S_t,\theta,\tau) = d(S_t,\tau) + \theta^{\text{dec},I}\phi_I^{\text{dec}}(S_t,\tau) + \theta^{\text{dec},O}\phi_O^{\text{dec}}(S_t,\tau) + \tau\theta^{\text{new}}. \tag{5.13}$$

Demand from chosen action

A policy should determine the optimal action to take. This action should not only be ideal for the current week but needs to consider its influence on the demand in the next couple of weeks. An action $\mathbf{a}$ induces a demand when the decision is made in week t. We define this demand as $\hat{d}(\mathbf{a})$:

$$\hat{d}(\mathbf{a}) = \sum_{\mathbf{p}\in\mathcal{P}}\sum_{u\in\mathcal{U}} \mathbf{p}a_{u\mathbf{p}}, \tag{5.14}$$

where the set of possible actions is defined by (4.4).

5.2.3 A DLA - CFA policy

We use a DLA-CFA policy to include all known information and the costs of not admitting patients directly. A policy deterministically outlines the actions to take for the next K weeks within the defined time interval $[t,t+K]$ for forecasting. To determine which action to take, we set up an LP. Next to the action for the current week t, this LP also determines the actions within the forecast period $[t, t+K]$. With $\mathbf{a}_{t'}$, we indicate the action resulting from the LP to take at time t'. Note that all actions are disregarded after solving the LP, except $\mathbf{a}_t$.

First, we define two CFA functions, describing costs for not admitting patients and exceeding costs. $\bar{f}_u\big(\hat{H}_{t'}\big(H_t, (\mathbf{a})_{t,..,t'-1}, \omega_H\big), \mathbf{a}_{t'}\big)$ represents the cost associated with expected patients on the waiting list at time t' who are not admitted.

$$\bar{f}_u\bigg(\hat{H}_{t'}\big(H_t, (\mathbf{a})_{t,..,t'-1}, \omega_H\big), \mathbf{a}_{t'}\bigg) = \kappa_u \sum_{\mathbf{p}\in\mathcal{P}} (\hat{h}_{u\mathbf{p}})_{t'} - (a_{u\mathbf{p}})_{t'}, \tag{5.15}$$

where $\hat{H}_{t'}\big(H_t, (\mathbf{a})_{t,..,t'-1}, \omega_H\big)$ describes the waiting list at time t', with all chosen actions up to time t' and arrivals to the waiting list up to t' as described by sample path ω_H of the arrivals of the waiting list.

$\bar{f}_d(S_t, (\mathbf{a})_{t,..,t'-1}, \theta)$ represents the cost associated with overusing capacity using the demand forecast and the action determined up to time t'.

$$\bar{f}_d\big(S_t, (\mathbf{a})_{t,..,t'-1}, \theta\big) = \kappa_c\big[\tilde{d}(S_t, \theta, t') + \sum_{\hat{t}=t}^{t'-1} \hat{d}(\mathbf{a}_{\hat{t}}) - \mathbf{r}\big]^+ \tag{5.16}$$

where $\tilde{d}(S_t, \theta, t')$ is defined in (5.13) and $\hat{d}(\mathbf{a}_{\hat{t}})$ is defined in (5.14).

The formulation of the LP is now as follows

$$\begin{aligned} A^{\pi\text{DLA-CFA}}(S_t, \theta, \omega_H) = {} & \operatorname*{arg\,min}_{\mathbf{a}_t\in\mathcal{A}(S_t)} C(S_t, \mathbf{a}_t) \\ & + \min_{\left(\mathbf{a}_{t'}\in\hat{\mathcal{A}}(S_t,(\mathbf{a})_{t,.,t'-1},\omega_H)\right)_{t'=t+1}^{t+K}} \sum_{t'=t+1}^{t+K} \bar{f}_u(\hat{H}_{t'}((\mathbf{a})_{t,..,t'-1}, \omega_H), \mathbf{a}_{t'}) \\ & \qquad + \bar{f}_d(S_t, (\mathbf{a})_{t,..,t'-1}, \theta), \end{aligned} \tag{5.17}$$

where $\hat{\mathcal{A}}(S_t, (\mathbf{a})_{t,.,t'-1}, \omega_H)$ describes the state-action space given current state S_t, chosen actions $(\mathbf{a})_{t,.,t'-1}$ at week t up to and including week $t'-1$, and arrivals to the waiting list according to sample path ω_H. This formulation is dependent on θ. The process of determining these parameters will be further detailed in the next Section 5.3.

5.3 Tuning the parameters

The optimization of the parameters θ, as outlined in equation (5.13), can be effectively conducted through the application of stochastic search methods. Numerous algorithms are available for this purpose. An overview of different types of aspects is described by Spall [34]. According to the "No Free Lunch Theorem", as articulated by Spall [34], along with Powell's "Tuning is Hard" [32], it becomes evident that there are significant challenges involved in both identifying an appropriate algorithm for a specific problem and in selecting the correct hyperparameters for the corresponding algorithm to achieve convergence. These factors contribute to the complexity and difficulty of the optimization process. In this section, we will outline two distinct methodologies applicable to our problem for identifying an optimal parameter θ and evaluating the resulting outcomes. We describe a stochastic approximation (SA) algorithm in Section 5.3.1, and we introduce the use of a newsvendor model for tuning in Section 5.3.2.

Problem definition

We define a function $\bar{F}(\theta, \omega)$ to be the total costs of the model for a timeinterval of length T, given parameter θ, and during which the state transitions according to the sample path ω between each decision epoch, and actions chosen according to DLA-CFA policy as defined in (5.17):

$$\bar{F}(\theta, \omega) = \sum_{t=0}^{T} C\Big(S_t(\omega), A^{\pi}\big(S_t(\omega), \theta\big)\Big).$$

To find the optimal values for θ, we need to solve the following:

$$\begin{aligned} \arg\min_{\theta} F(\theta) &:= \arg\min_{\theta} \mathbb{E}_{\omega} \bar{F}(\theta, \omega | S_0) \\ &= \arg\min_{\theta} \mathbb{E}_{\omega} \left[\sum_{t=1}^{T} C\Big(S_t(\omega), A^{\pi}\big(S_t(\omega), \theta\big)\Big) | S_0 \right], \end{aligned} \tag{5.18}$$

where S_0 is the initial state.

5.3.1 Stochastic approximation

A class of algorithms designed to solve (5.18) is stochastic approximation (SA), which requires computing stochastic subgradients of the objective function iterative. The general form of an iteration, which resembles the steepest descent method, is described as the following equation,

$$\hat{\theta}_{n+1} = \hat{\theta}_n - \beta_n \hat{\mathbf{g}}_n(\hat{\theta}_n),$$

where $\hat{\mathbf{g}}_n$ is an estimation of the gradient of the objective function concerning θ_n, and a_n is the step size. The gradient, $\frac{\delta F}{\delta \theta}$, cannot be computed directly. Therefore, we need to design an estimation.

Stochastic gradient descent

One of the most known and oldest methods to approximate $\hat{\mathbf{g}}_n$ is the finite difference method (FD) [34]:

$$\hat{\mathbf{g}}_n(\theta_n, \omega^n) = \begin{bmatrix} \frac{\bar{F}(\theta_n + c_n \xi_1, \omega^n) - \bar{F}(\theta_n - c_n \xi_1, \omega^n)}{2c_k} \\ \vdots \\ \frac{\bar{F}(\theta_n + c_n \xi_p, \omega^n) - \bar{F}(\theta_n - c_n \xi_p, \omega^n)}{2c_k} \end{bmatrix} \tag{5.19}$$

where ξ_i denotes a vector with a 1 in the *i*-th place and 0's elsewhere, c_n defines the difference magnitude and $\omega^n \in \Omega$ is a randomly drawn sample path.

The construction of the above gradient approximation requires twice the dimension of θ of measurements of the objective function. To lower the number of measurements, we can define the perturbation to be simultaneous. This is called the simultaneous perturbation (SP) method [34]:

$$\hat{\mathbf{g}}_n(\theta_n, \omega^n) = \begin{bmatrix} \frac{\bar{F}(\theta_n + c_n\Delta_n, \omega^n) - \bar{F}(\theta_n - c_n\Delta_n, \omega^n)}{2c_k\Delta_{n1}} \\ \vdots \\ \frac{\bar{F}(\theta_n + c_n\Delta_n, \omega^n) - \bar{F}(\theta_n - c_n\Delta_n, \omega^n)}{2c_k\Delta_{np}} \end{bmatrix} \tag{5.20}$$

where Δ_k is a mean-zero p-dimensional random perturbation vector. The examples shown above are two-sided versions. One can also define a one-sided version where the result with the random input, $\theta_n + c_n\Delta_n$, is compared with the result of θ_n.

Note that if we directly use basic Stochastic Gradient Descent (SGD), such as equations (5.19) and (5.20), the values of θ_n will fluctuate highly because of the stochasticity of the function $\bar{F}(\theta, \omega)$. And so, it will be very unlikely that such an algorithm will converge. Therefore, we adapt the sampling, estimation of gradient, and/or step size by using methods such as mini-batch sampling, momentum, and different step size policies. We will explain this method below.

Mini batch sampling

The results of $\bar{F}(\theta, \omega)$ and $\bar{F}(\theta + \eta\zeta, \omega)$ can vary for different ω. In many cases, a single measurement of both variables may yield an inadequate gradient estimation. To enhance the accuracy of our results, we generate i sample paths, which allows for a more robust estimation of the gradient. We then average the gradient estimations obtained from these sample paths to produce a more precise estimate.

$$\mathbf{g}_n(\theta_n, \omega^n) = \frac{1}{m_n} \sum_{i=1}^{m_n} \hat{\mathbf{g}}_n(\theta_n, \omega^{n,i}) \tag{5.21}$$

An algorithm with mini-batch sampling will increase the computational costs, as each iteration asks for $2i$ times the dimension of θ function evaluations. However, mini-batch sampling often requires fewer iterations [46].

Momentum

Momentum in an SA algorithm implies that in iteration n, not only the current gradient estimation accounts for the update of the search variable but also the momentum of all previously determined gradient estimation accounts.

$$G_n = (1 - \alpha_n)G_{n-1} + \alpha\mathbf{g}_n$$

where, G_n is used to update the search variable, with $G_0 = 0$ and $\mathbf{g}_n$ is the current gradient estimation.

The algorithm Ghadmini and Powell propose in their article [44] on the DLA-CFA policy tuning uses a momentum method version. This algorithm is presented in Algorithm 2. Here α_n is the momentum hyperparameter. When $\alpha_n = 1$, this is just an SGD. β_n determines the step size. In this article, they use the RMSprop step size policies to update β_n. They determine the hyperparameter α_n as follows:

$$\alpha_n = \frac{a}{(\delta+4)N}, \tag{5.22}$$

where δ is the dimension of the tuning parameter, N is the total number of iterations, and a is a hyperparameter determining the rate of momentum.

Step size policies

RMSprop, introduced by Geoffrey Hinton in [47], adapts the step sizes to improve the performance and speed of convergence. Typical with stochastic approximation, the magnitude of the gradient can differ. This step size policy tackles this by using a moving average of the squared gradient and adjusting the weight updates by this magnitude, by hyperparameter γ.

$$\mathbb{E}[\mathbf{g}^2]_n = \gamma\mathbb{E}[\mathbf{g}^2]_{n-1} + (1-\gamma)\mathbf{g}_n^2 \tag{5.23}$$

$$\theta_n = \theta_{n-1} - \frac{b}{\sqrt{\mathbb{E}[\mathbf{g}^2]_n + \epsilon}}\mathbf{g}_n \tag{5.24}$$

where b is a tunable parameter, and ϵ is a small constant added for numerical stability.

Algorithm 2 The Stochastic Averaging Numerical Gradient method (SANG) based on [44].

1: **Input:** Initialize $\theta^0 \in \mathbb{R}$, $\bar{G}^0 = 0$, an iteration limit N and positive sequences $\{\alpha_n\}_{n\geq 1}$, $\{\beta_n\}_{n\geq 1}$, $\{\eta_n\}_{n\geq 1}$ supported on $(0,1)$.
2: **for** $n = 1$ to N **do**
3: Update policy parameters as

$$\theta_n^y = \theta_{n-1} - \beta_n\bar{G}^{n-1}. \tag{5.25}$$

$$\theta_n = (1-\alpha_n)\theta_{n-1} + \alpha_n\theta_n^y. \tag{5.26}$$

4: Generate trajectory ω^n and a random Gaussian vector ν_n to compute the gradient estimator as

$$G_{\eta_n}(\theta_n, \omega^n) = \frac{\bar{F}(\theta_n + \eta_n\nu_n, \omega^n) - \bar{F}(\theta_n, \omega^n)}{\eta_n}\nu_n. \tag{5.27}$$

5: Update the overall estimator as

$$\bar{G}^n = (1-\alpha_n)\bar{G}^{n-1} + \alpha_n G_{\eta_n}(\theta_n, \omega^n). \tag{5.28}$$

6: **end for**

Adam, which stands for Adaptive Moment Estimation, is a step-size policy well-suited for a range of non-convex problems with non-stationary objectives. Although unpublished, see [48], this algorithm is widely applied in various stochastic optimization problems. It uses estimations of the first and second (raw) moments, and by doing so it can adjust the learning rate. The algorithm is described in Algorithm 3.

Algorithm 3 Parameter tuning using an Adam algorithm for stochastic search, based on [48]

Input: Initialize $\theta_0 \in \mathbb{R}^D$, $\mathbf{m}_0 = 0$ and $\mathbf{v}_0 = 0$. Set N as the number of iterations. Define β_1, β_2 as exponential decay rates, and β_3 that determines the learning rate. ϵ is a small number that accounts for stability.
for $n = 1$ to N **do**
    Generate trajectory ω^n and obtain gradient estimator by SGD, $\mathbf{g}_n(\theta_n, \omega_n)$ as in (5.27).
    Update biased first moment estimate:

$$\mathbf{m}_n = \beta_1 \mathbf{m}_{n-1} + (1 - \beta_1)\mathbf{g}_n \tag{5.29}$$

    Update biased second raw moment estimate:

$$\mathbf{v}_n = \beta_2 \mathbf{v}_{n-1} + (1 - \beta_2)\mathbf{g}_n^2 \tag{5.30}$$

    Compute bias-corrected first moment estimate: $\hat{\mathbf{m}}_n = \frac{\mathbf{m}_n}{1-(\beta_1)^n}$
    Compute bias-corrected second raw moment estimate: $\hat{\mathbf{v}}_n = \frac{\mathbf{v}_n}{1-(\beta_{})2^n}$
    Update parameters

$$\theta_n = \theta_{n-1} - \frac{\beta_3 \hat{\mathbf{m}}_n}{\sqrt{\hat{\mathbf{v}}_n} + \epsilon} \tag{5.31}$$

end for

Implementation

The *No Free Lunch Theorem*, [49], tells us that, when considering the performance of optimization algorithms across the entirety of problem instances, no single algorithm can be deemed superior universally. Instead, the theorem establishes that the effectiveness of an algorithm is inherently tied to the specific problem it is deployed to solve. Hence, we cannot determine the best-performing algorithm without testing it on our specific problem. Both algorithms, RMSprop and Adam, ask for tuning of the (gain) coefficients.

But first, we note that we do not necessarily need to simultaneously tune all separate θ entries when tuning with a simulation. Within the simulation, we already have the perfect forecast of the separate θ demand influences. In this way, $\theta^{\text{dec},O}$, $\theta^{\text{dec},I}$ and θ^{new} can be tuned separately. We highlight this further in Chapter 6.

5.3.2 Direct approximation

The downsides of stochastic approximation include the uncertainty surrounding convergence and the substantial computational costs associated with the method. Establishing a formulation that directly approximates the parameters could significantly enhance efficiency and reduce the computational time. In response to this, we propose the newsvendor model as a suitable formulation. However, a primary limitation of the newsvendor model is the necessity of having a certain level of knowledge regarding the distribution of the incoming exogenous information.

Implementing a newsvendor model for parameter tuning for policy search has been relatively unexplored in the current literature, as we mentioned earlier in Section 3.2.3.

Newsvendor model

We reshape the model in (5.18) to a state-dependent problem. Instead of determining a θ that fits all states, we develop a formulation that fits the θ to the current state. Let Ω be the set of all possible sample paths of new arriving patients combined with the demand of patients with an upcoming decision moment. $p(\omega)$ is the probability of drawing such a sample path. We define the cost function, $Q(S_t, \omega, \theta)$. Then we solve

$$\theta^* = \arg\min_{\theta} \mathbb{E}_\omega Q^{\text{news}}(S_t, \omega, \theta). \tag{5.32}$$

Single period

We first consider finding an estimate for the next week.

Define $d(S_t)$ as the demand of the current patients. Define $D(S_t, \omega) = d(S_t) + A(S_t) + \omega$ and $\hat{D}(S_t, \theta) = d(S_t) + A(S_t) + \theta$, where $A(S_t)$ describes the demand of all patients in the waiting list in state S_t.

In a typical newsvendor model, one defines this cost function such that one accounts for overestimation of the demand, $\theta > \omega$, and underestimation $\omega > \theta$ separately.

$$Q^{\text{news}}(S_t, \omega(S_t), \theta) = o(S_t)\big[\theta - \omega(S_t)\big]^+ + u(S_t)\big[\omega(S_t) - \theta\big]^+ \tag{5.33}$$

where $o(S_t)$ defines the costs for an overestimation of the demand and $u(S_t)$ defines the costs of an underestimation, which both can be state-dependent, and $[x]^+ := \max(0, x)$. Note that the possible sample paths can be dependent on the current state S_t.

Theorem 5.3.1. *The optimal solution, θ^*, for* (5.33) *is given by*

$$\mathbb{P}(\omega(S_t) \leq \theta^*) = \frac{u(S_t)}{u(S_t) + o(S_t)} \tag{5.34}$$

The proof of theorem 5.3.1 is provided by most literature related to the newsvendor model, for example, see the introduction of [36].

But now consider our problem and the cost function determination in a newsvendor-type model. Note that 'costs' for under- or overestimation of the demand result from too many or too few admissions from the waiting list, resulting in possible overcapacity use or waiting list costs. These costs only occur if the capacity limit is met, therefore we rewrite the cost function such that the costs are accounted for when exceeding the capacity. If this is not the case, the costs are only accounted for by a factor $\alpha \in (0, 1]$. We define the objective function $Q(S_t, \omega, \theta)$:

$$\begin{aligned} Q^{\text{news}}(S_t, \omega(S_t), \theta) =& \mathbb{1}_{\hat{D}(S_t,\theta) > f}\, o(S_t)\big([\theta - \omega(S_t)]^+ - [f - D(S_t, \omega(S_t))]^+\big) \\ &+ \mathbb{1}_{D(S_t,\omega(S_t)) > f}\, u(S_t)\big([\omega(S_t) - \theta]^+ - [f - \hat{D}(S_t, \theta)]^+\big) \\ &+ \alpha \mathbb{1}_{\hat{D}(S_t,\theta) \leq f} \mathbb{1}_{D(S_t,\omega(S_t)) \leq f}\big(o(S_t)\big[\theta - \omega\big]^+ + u(S_t)\big[\omega(S_t) - \theta\big]^+\big), \end{aligned} \tag{5.35}$$

where

$$\mathbb{1}_{D(S_t,\omega(S_t))>f} = \begin{cases} 1, & \text{if } D(S_t,\omega(S_t)) > f, \\ 0, & \text{otherwise.} \end{cases}$$

Theorem 5.3.2. *The optimal solution,* θ^**, for* (5.35) *is given by*

$$\mathbb{P}(\omega(S_t) \le \theta^*) = \begin{cases} \frac{u(S_t)(1-\mathbb{P}(D(S_t,\omega(S_t))<f))}{o(S_t)+u(S_t)(1-\mathbb{P}(D(S_t,\omega(S_t))<f))} & \text{if } \hat{D}(S_t,\theta^*) > f \\ \frac{\alpha u(S_t)\mathbb{P}(D(S_t,\omega(S_t))\le f)}{\mathbb{P}(D(S_t,\omega(S_t))\le f)(\alpha(o(S_t)+u(S_t))-u(S_t))+u(S_t)} & \text{if } \hat{D}(S_t,\theta^*) \le f \end{cases} \tag{5.36}$$

Proof. First, we rewrite (5.35) as follows:

$$\begin{aligned} Q(S_t,&\omega(S_t),\theta) = \\ &\mathbb{1}_{\hat{D}(S_t,\theta)>f} o(S_t)\Big(- f + \hat{D}(S_t,\theta) - \min\big(\omega(S_t),\theta\big) + \min\big(f - d(S_t) - A(S_t), \omega(S_t)\big)\Big) \\ &+\mathbb{1}_{D(S_t,\omega(S_t))>f} u(S_t)\Big(\omega(S_t) - \min(\omega(S_t),\theta) - \mathbb{1}_{\hat{D}(S_t,\theta)\le f}\big(f - \hat{D}(S_t,\theta)\big)\Big) \\ &+\alpha \mathbb{1}_{\hat{D}(S_t,\theta)\le f}\mathbb{1}_{D(S_t,\omega(S_t))\le f}\Big(o(S_t)\theta + u(S_t)\omega(S_t) + (-o(S_t) - u(S_t))\min\big(\omega(S_t),\theta\big)\Big). \end{aligned} \tag{5.37}$$

Then we take the expectation with respect to ω:

$$\begin{aligned} &\mathbb{E}_{\omega(S_t)} Q(S_t,\omega(S_t),\theta) = \\ &\mathbb{1}_{\hat{D}(S_t,\theta)>f} o(S_t)\Big(- f + \hat{D}(S_t,\theta) - \mathbb{E}\min\big(\omega(S_t),\theta\big) + \mathbb{E}\min\big(f - d(S_t) - A(S_t), \omega(S_t)\big)\Big) \\ &+\mathbb{P}\big(D(S_t,\omega(S_t)) > f\big)u(S_t)\Big(\mathbb{E}\omega(S_t) - \mathbb{E}\min\big(\omega(S_t),\theta\big) - \mathbb{1}_{\hat{D}(S_t,\theta)\le f}\big(f - \hat{D}(S_t,\theta)\big)\Big) \\ &+\alpha\mathbb{1}_{\hat{D}(S_t,\theta)\le f}\mathbb{P}\big(D(S_t,\omega(S_t)) \le f\big)\Big(o(S_t)\theta + u(S_t)\mathbb{E}\omega(S_t) + (-o(S_t) - u(S_t))\mathbb{E}\min\big(\omega(S_t),\theta\big)\Big) \end{aligned} \tag{5.38}$$

Define $p(\omega(S_t) == i)$ the probability that the resulting demand arrivals is equal to i:

$$\mathbb{E}\min(\omega(S_t),\theta) = \sum_{i=0}^{\theta} ip(\omega(S_t) = i) + \theta \sum_{i=\theta+1}^{\infty} p(\omega(S_t)), \tag{5.39}$$

$$\frac{d}{d\theta}\mathbb{E}\min(\omega(S_t),\theta) = \theta \sum_{i=\theta+1}^{\infty} p(\omega(S_t)) = \mathbb{P}(\omega(S_t) > \theta). \tag{5.40}$$

Now we take the derivative of (5.38), with respect to θ:

$$\begin{aligned} &\frac{d}{d\theta} Q(S_t,\omega(S_t),\theta) = \\ &\begin{cases} o(S_t)\big(1 - \mathbb{P}(\omega(S_t) > \theta) + 0\big) & \\ \quad +\mathbb{P}(D(S_t,\omega(S_t)) > f)u(S_t)\big(0 - \mathbb{P}(\omega(S_t) > \theta)\big), & \text{if } \hat{D}(S_t,\theta) > f, \\ \mathbb{P}(D(S_t,\omega(S_t)) > f)u(S_t)\big(0 - \mathbb{P}(\omega(S_t) > \theta) + 1\big) & \\ \quad +\alpha\mathbb{P}(D(S_t,\omega(S_t)) > f)\big(o(S_t) + (-o(S_t) - u(S_t))\mathbb{P}(\omega(S_t) > \theta)\big), & \text{if } \hat{D}(S_t,\theta) \le f. \end{cases} \end{aligned} \tag{5.41}$$

When we set the above derivative to zero, and use $1 - \mathbb{P}(\omega(S_t) \le \theta) = \mathbb{P}(\omega(S_t) > \theta)$, we can solve for θ, we find the optimal value θ^* accordingly. □

Defining costs functions

For the undershoot costs, we use the capacity costs described in Table 4.1. The urgency costs need to be translated into demand for the overshoot costs. We use the ratio of the total urgency costs to the total demand of patients currently on the waiting list.

$$o(S_t) = \kappa_c. \tag{5.42}$$

$$u(S_t) = \begin{cases} \frac{\sum_{u\in\mathcal{U}} \kappa_u \sum_{\mathbf{p}\in\mathcal{P}} h_{u\mathbf{p}}}{\sum_{\mathbf{p}\in\mathcal{P}} \mathbf{p} \sum_{u\in\mathcal{U}} h_{u\mathbf{p}}}, & \text{if } \sum_{u\in\mathcal{U},\mathbf{p}\in\mathcal{P}} (h_{u\mathbf{p}})_t > 0, \\ \kappa_c, & \text{otherwise.} \end{cases} \tag{5.43}$$

Multi-item & multi-period

We define $d_\tau(S_t)$ to be the demand of the current patients in a week τ from now. Define the following demand estimations:

$$D_\tau(S_t, \omega(S_t)) = d_\tau(S_t) + A_\tau(S_t) + \sum_{t'=t}^{t+\tau-1} \omega_{t'}(S_t) + \omega_\tau(S_t), \tag{5.44a}$$

$$\hat{D}_\tau(S_t, \theta) = d_\tau(S_t) + A_\tau(S_t) + \sum_{t'=t}^{t+\tau-1} \theta_{t'} + \theta_\tau. \tag{5.44b}$$

Note that for week τ the values of $(\omega_\tau)_{\tau=t,\ldots,t+\tau-1}$ and$(\theta_\tau)_{\tau=t,\ldots,t+\tau-1}$ are fixed.

We rewrite the objective function to address the multi-disciplinary and multi-period forecasting.

$$\begin{aligned} &Q^{\text{news, mp}}(S_t,\omega(S_t),\theta) = \\ &\quad \sum_{\tau=t}^{t+K} \sum_{c\in\mathcal{C}} \mathbb{1}_{\hat{D}_\tau(S_t,\theta)_c > f_c} o(S_\tau)\big([\theta_{\tau c} - \omega_{\tau c}(S_t)]^+ - [f_c - D(S_t,\omega(S_t))_c]^+\big) \\ &\quad + \mathbb{1}_{D_\tau(S_t,\omega(S_t))_c > f_c} u(S_\tau)\big([\omega_{\tau c}(S_t) - \theta_{\tau c}]^+ - [f_c - \hat{D}_\tau(S_t,\theta)_c]^+\big) \\ &\quad + \alpha \mathbb{1}_{\hat{D}_\tau(S_t,\theta)_c \le f_c} \mathbb{1}_{D(S_t,\omega(S_t))_c \le f_c} \big(o(S_\tau)\big[\theta_{\tau c} - \omega_{\tau c}(S_t)\big]^+ + u(S_\tau)\big[\omega_{\tau c}(S_t) - \theta_{\tau c}\big]^+\big) \end{aligned} \tag{5.45}$$

If ω_{tc} would be independent across time, then we solve similarly as above. This independence holds if n_I^{max} is very large, as there is no limit set on the number of inpatients admitted, and therefore, the probability distribution is not dependent on previous estimations.

Then optimal value, θ^*, of the objective function in (5.45) is given by:

$$\begin{aligned} &\mathbb{P}(\omega_{\tau c}(S_t) \le \theta^*_{\tau c}) = \\ &\quad \begin{cases} \frac{u(S_t)(1-\mathbb{P}(D_\tau(S_t,\omega(S_t))_c < f_c))}{o(S_t)+u(S_t)(1-\mathbb{P}(D(S_t,\omega(S_t))_c < f_c))}, & \text{if } \hat{D}_\tau(S_t,\theta^*)_c > f_c, \\ \frac{\alpha u(S_t)\mathbb{P}(D_\tau(S_t,\omega(S_t))_c < f_c)}{\mathbb{P}(D_\tau(S_t,\omega(S_t))_c < f_c)(\alpha(o(S_t)+u(S_t))-u(S_t))+u(S_t)}, & \text{if } \hat{D}_\tau(S_t,\theta^*)_c \le f_c, \end{cases} \\ &\quad \forall \tau = t,\ldots,t+K, \forall c \in \mathcal{C}. \end{aligned} \tag{5.46}$$

If the number of currently admitted patients does influence the probability distribution, we can redefine it likewise. Let $\omega_{\tau c}(n)$ be the probability distribution of the extra demand in week τ for discipline c, given that there are already n inpatients admitted in the forecast of week τ. And let $\eta_\tau(n)$ be the probability distribution of new arrivals in a week τ given that there are n inpatients admitted in the forecast of week τ. Then the optimal solution for (5.45) is also dependent on the forecast of the number of inpatients up to the given time moment:

$$\mathbb{P}(\omega_{\tau c}(S_t, n_\tau) \leq \theta^*_{\tau c}) = \begin{cases} \frac{u(S_t)(1-\mathbb{P}(D_\tau(S_t,\omega(S_t,n_\tau))_c<f_c))}{o(S_t)+u(S_t)(1-\mathbb{P}(D(S_t,\omega(S_t,n_\tau))_c<f_c))}, & \textit{if } \hat{D}_\tau(S_t,\theta^*)_c > f_c, \\ \frac{\alpha u(S_t)\mathbb{P}(D_\tau(S_t,\omega(S_t,n_\tau))<f_c)}{\mathbb{P}(D_\tau(S_t,\omega(S_t,n_\tau))<f_C)(\alpha(o(S_t)+u(S_t))-u(S_t))+u(S_t)}, & \textit{if } \hat{D}_\tau(S_t,\theta^*)_c \leq f_c, \end{cases} \quad \forall \tau = t,\dots,t+K, \forall c \in \mathcal{C}. \tag{5.47a}$$

We can describe the estimated number of patient arrivals similarly to the previously defined optimal solution:

$$\mathbb{P}(\eta_{\tau c}(n_\tau) \leq n^*_{\tau c}) = \begin{cases} \frac{u(S_t)(1-\mathbb{P}(D_\tau(S_t,\eta(n_\tau))_c<f_c))}{o(S_t)+u(S_t)(1-\mathbb{P}(D(S_t,\eta(n_\tau))_c<f_c))}, & \textit{if } \hat{D}_\tau(S_t,\theta^*)_c > f_c, \\ \frac{\alpha u(S_t)\mathbb{P}(D_\tau(S_t,\eta(n_\tau))_c<f_c)}{\mathbb{P}(D_\tau(S_t,\eta(n_\tau))_c<f_c)(\alpha(o(S_t)+u(S_t))-u(S_t))+u(S_t)}, & \textit{if } \hat{D}_\tau(S_t,\theta^*)_c \leq f_c, \end{cases} \quad \forall \tau = t,\dots,t+K, \forall c \in \mathcal{C}. \tag{5.47b}$$

The forecast of the number of inpatients for week τ may vary across different disciplines $\mathcal{C}$. To determine the forecast for the following week, we use the rounded average of these numbers. The full algorithm used is described in Algorithm 4.

Algorithm 4 Parameter tuning using a newsvendor model

Given current state S_t, number of inpatient n_τ, and timeinterval K,
Initialize $\hat{n}_0 = 0$
for τ = 1 to K **do**
    for c = 1 to C **do**
        Determine $\theta^*_{\tau c}$ using (5.47a) given $\hat{n}_\tau$.
        Determine $n^*_{\tau c}$ using (5.47b) given $\hat{n}_\tau$.
    end for
    Set $\hat{n}_{\tau+1} = \hat{n}_\tau + \frac{1}{C}\sum_{c\in\mathcal{C}} n^*_{\tau c}$
end for
return (θ^*)

Chapter 6

Numerical results

In this chapter, we present the numerical results of an analysis of the developed policy. In Section 6.1, we evaluate the algorithms for tuning the parameters. In Section 6.3, we look into the result of a single-discipline system of the proposed model and evaluate the result. Finally, we use the data from the rehabilitation clinic to conclude a policy for this system in Section 6.4.

We perform our analysis of the results on an external server with 32 CPU cores and 64 GB memory. The LP of the DLA-CFA policy is solved using Gurobi 11.0.3 within a Python 3.10.11 environment. A decision for each week can be made in 0.02 seconds. The simulation of various timerange, number of iterations, and batch size scales accordingly. For example, a tuning algorithm of 10000 iterations with a timerange of 50 weeks and batch size of 5 takes about 34 h.

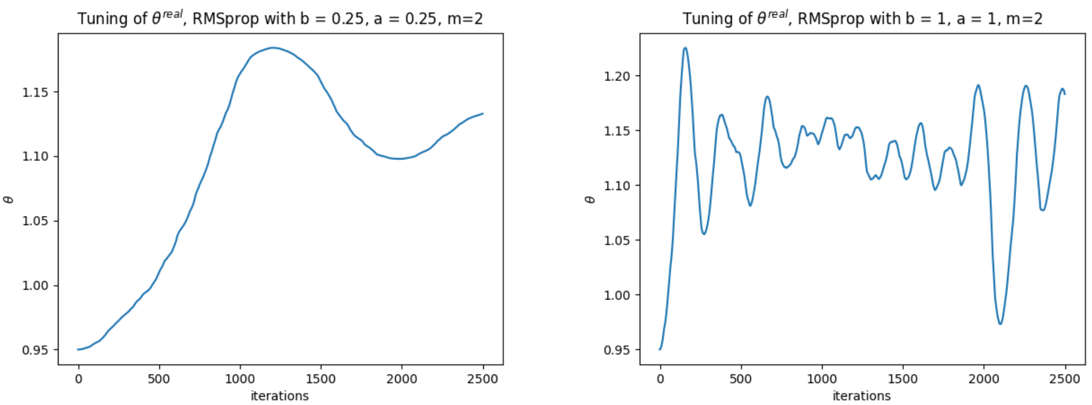
6.1 RMSprop versus Adam

In Section 5.3, we proposed two algorithms to tune the parameters in (5.13). In this section, we consider the two algorithms proposed, using RMSprop or Adam.

To use Algorithm 2, the RMSprop algorithm, we should choose the values of the algorithm's hyper-parameters. As these parameters significantly influence the algorithm's working and eventually possible convergence, we should carefully consider choosing the parameters. Figure 6.1 illustrates the impact of varying values for parameters a and b, as defined in (5.22) and (5.24). We analyze this within a policy with complete knowledge of forthcoming exogenous arriving demand, multiplied by the parameter θ^{real}. In Figure Figure 6.1(a), the observed convergence is notably slow, whereas Figure (b) depicts an excessively rapid response that leads to overshooting.

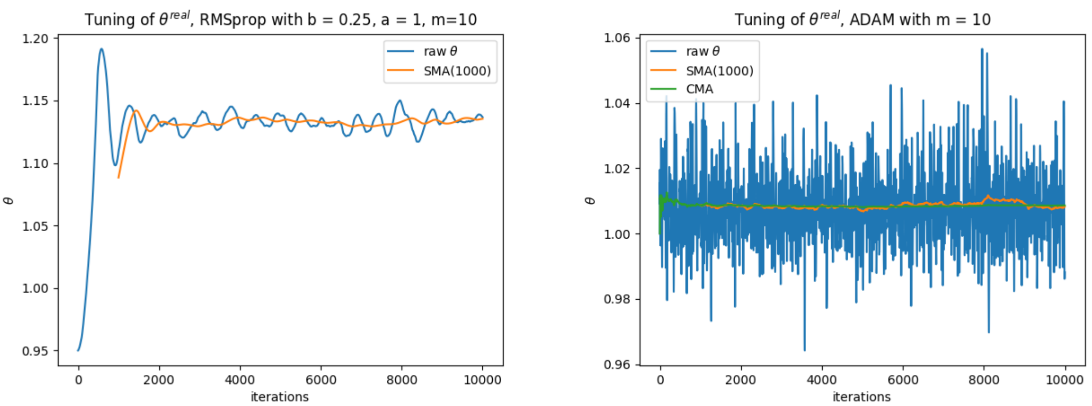
Algorithm 3, the Adam algorithm, also uses such hyper-parameters, but these have less influence on the step size. As advised in [49], we set $\beta_1 = 0.9$, $\beta_2 = 0.99$ and $\beta_3 = 0.001$. Figure 6.2 compares the two algorithms for 10000 iterations, both with a batch size of 10. The Adam algorithm demonstrates a more rapid convergence to obtain solutions than RMSprop while maintaining stability throughout the process. In contrast, the RMSprop algorithm exhibits a more gradual response to the acquired results, resulting in slower convergence rates. This difference in responsiveness can impact the efficiency of training the parameter, where Adam may be preferred for its quicker adaptation. When we tested the RMSprop algorithm for tuning the parameters of our developed model, we found that identifying suitable hyperparameters was particularly challenging. Despite conducting various tests, we encountered difficulties in attaining convergence within our results.. In contrast, our initial test with the Adam algorithm showed more promising results. Consequently, we decided to continue using Algorithm 3 in our analysis.

We introduce the Simple Moving Average (SMA(n)) for a given integer n, which denotes the moving average over the last n iterations given the current iteration. The Cumulative Moving Average (CMA) describes the average from the start of the tuning up to the current iteration. These moving averages are introduced to evaluate the value of that parameter tuning better.



(a) (b)

Figure 6.1: θ^{real} tuning for 2500 iterations with batch size 2 and (a) $a = 0.25$, $b = 0.25$ and (b) $a = 1$, $b = 1$



(a) (b)

Figure 6.2: Tuning with RMSprop (a) and ADAM (b), with settings $a = 0.25$, $b = 1$, and learning rate 0.001, respectively. Both use a batch size of 10.

6.2 Myopic policies

To evaluate the performance of the developed method, we introduce three myopic policies. We introduce the 'cheapest direct cost' policy, which admits all outpatients directly upon entering the waiting list. This effectively eliminates access time-related costs. By adopting this policy, we are able to analyze demand fluctuations and peak periods based on patient arrivals.

The second myopic policy is referred to as the 'no forecast policy.' This policy incorporates a strategic approach by extending its focus to a one-week horizon, utilizing the existing demand and considering patients whose discharge dates are this week. By analyzing the known demand for the upcoming week, this policy aims to optimize admissions to align more closely with the current state. We can also describe this by setting $K = 1$ and $\theta = 0$. This policy does not require the establishment of parameters. It presents the results without engaging in forecasting and does not take into account the forthcoming weeks, with the exception of the immediate subsequent week.

The final myopic policy possesses complete knowledge of the exogenous information that arrives within the forecasting time interval $[t, t + K]$. This policy is referred to as the 'full knowledge' policy. We introduce this policy as it eliminates the need for approximation of the parameters of the developed policy while still considering the upcoming weeks, such that the influence of the forecasting can be identified.

The rehabilitation clinic currently has no clear defined policy, making a direct comparison impossible. However, it is important to note that the developed policy mathematically describes what the planner currently does intuitively: forecasting the demand. We discuss the results with the planners in further detail in Section 6.4.7.

6.3 Testing on single-discipline systems

We first test our developed policy on single-disciplinary systems to evaluate the performance of various settings. We test the model for both an infinite and a finite bed capacity in Sections 6.3.1 and 6.3.2, respectively. In Section 6.4, we apply the policy design to the system parameters of the rehabilitation clinic.

6.3.1 Single-discipline and infinite bed capacity

The system parameters are given in Table 6.5. Note that the bed capacity is unlimited, simplifying the probability distribution for new demand and the use of the newsvendor model.

We use both the tuning version with Algorithm 3 and the newsvendor-like version, as in (5.36), to determine the values of the parameters. Initially, we will assess values through the application of the tuning algorithm. Subsequently, we will analyze the newsvendor model with regard to parameter estimation. Finally, we will evaluate the performance of the developed methodology.

In terms of computation time, a decision according to the DLA-CFA policy can be made in 0.02 seconds, given the CPU use. The simulation of various timeranges, number of iterations, and batch size scales accordingly. For example, a tuning algorithm of 10000 iterations with a timerange of 50 weeks and batch size of 5 takes about 34 h.

Table 6.1: Setting for a system model with single discipline and infinite bed capacity.

$C = 1$	$\kappa_c = \{1.5\}$
$U = 3$	$\kappa_u = \{0.55, 1.05, 1.55\}$
$\mathbf{r} = \{95\}$	
$\mathcal{P} = \{\{1\}, \{2\}, \{3\}\}$	$\mathcal{Q} = \{\{2\}, \{3\}\}$
$p_{\mathcal{P}}^{\text{plan}} = \{0.6, 0.2, 0.2\}$	$p_{\mathcal{Q}}^{\text{plan}} = \{0.25, 0.75\}$
$T_O^{\text{MDM}} = 2$	$T_I^{\text{MDM}} = 2$
$T_O^{\text{start}} = 5$	$T_I^{\text{start}} = 5$
$\lambda_O = 1$	$\lambda_I = 2$
$p_O^{\text{cont}} = \{0.8, 0.6, 0\}$	$p_I^{\text{cont}} = \{0.6, 0.4, 0\}$
	$n_I^{\max} = \infty$

Parameter estimation: stochastic approximation

In this subsection, we show the result of using Algorithm 3, the Adam algorithm to tune the θ^{new}, $\theta^{\text{dec},O}$ and $\theta^{\text{dec},I}$ parameters. When tuning one of the three parameters, we give the algorithm full knowledge of the other two parameters. If we tune and adjust multiple values for the parameters simultaneously, the algorithm may struggle to distinguish the individual effects of each parameter. We also observed this phenomenon empirical ly upon testing.

It is important to emphasize that the goal of tuning is not obtaining the most accurate estimation of the exogenous information. Instead, we seek to identify a value that yields the optimal policy decision. Given the stochastic nature of the model, perfect tuning results are difficult to achieve, as will be shown in the tuning figures below. We assume that convergence is deemed sufficient if the value demonstrates stability or exhibits fluctuations around a particular value throughout approximately 1000 iterations.

We can determine the expectation of the new inpatient demand: $2 \times (0.25 \times 2 + 0.75 \times 3) = 5.5$. However, to averse risk, it might be better to forecast more capacity to decrease the risk of overcapacity, which is what we also observe. In Figure 6.3, the result of tuning of θ^{new} is shown for 10000 iterations. The tuning indicates an optimal value of about 5.5 demand. To identify a value from the tuning process, we plot a Simple Moving Average, considering the average over the last 1000 iterations.

In equation (5.13), we see that the possible extra demand from the inpatient decision moment is state-dependent. We want to find value $\theta^{\text{dec},I}$ as a multiplier for $\phi_I^{\text{dec}}(S_t, \tau)$ to forecast the demand coming from these patients. Figure 6.4 displays the results of a tuning run comprising 10000 iterations, with a batch sample size of 5 for each iteration. The value for $\theta^{\text{dec},I}$ can be concluded in the range of 0.78-0.8. Similarly, we can tune $\theta^{\text{dec},O}$ for outpatients. Figure 6.5 shows the corresponding tuning, resulting in a value of about 0.96.

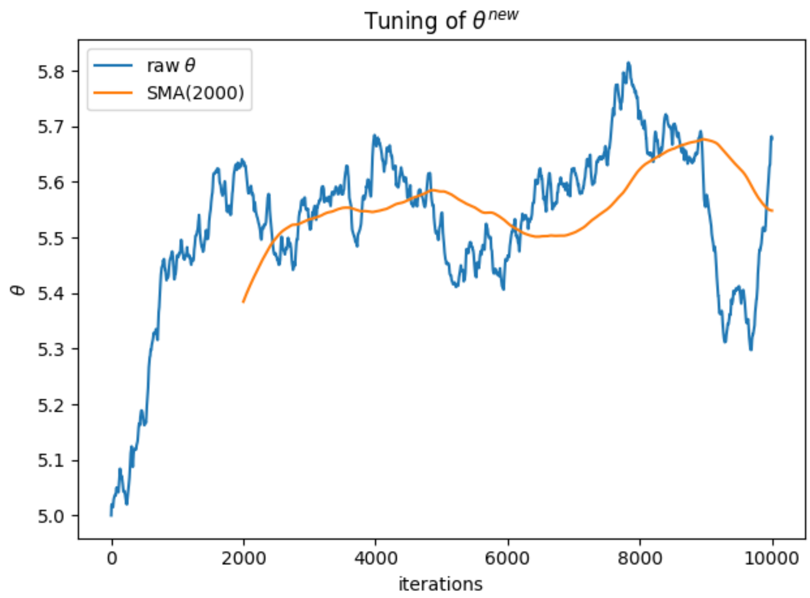


Figure 6.3: Tuning of θ^{new} using Adam algorithm, with m = 5 and learning rate 0.001.

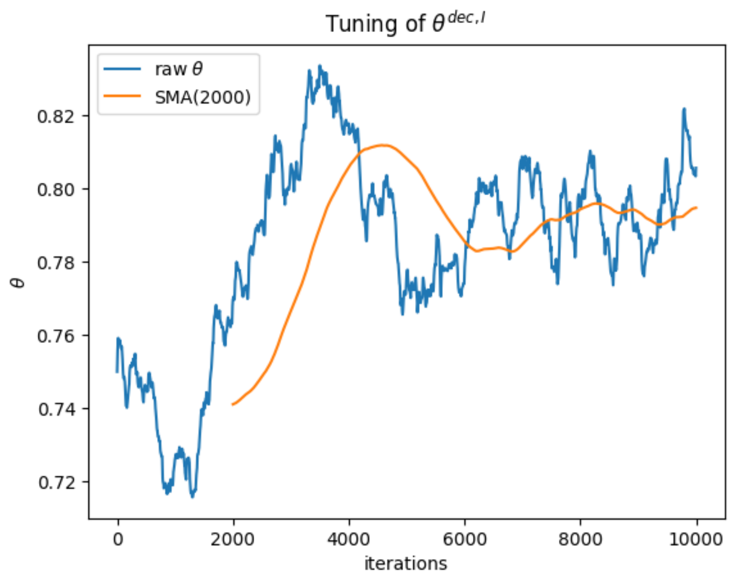


Figure 6.4: Tuning of $\theta^{dec,I}$ using Adam algorithm, with m = 5 and learning rate 0.001.

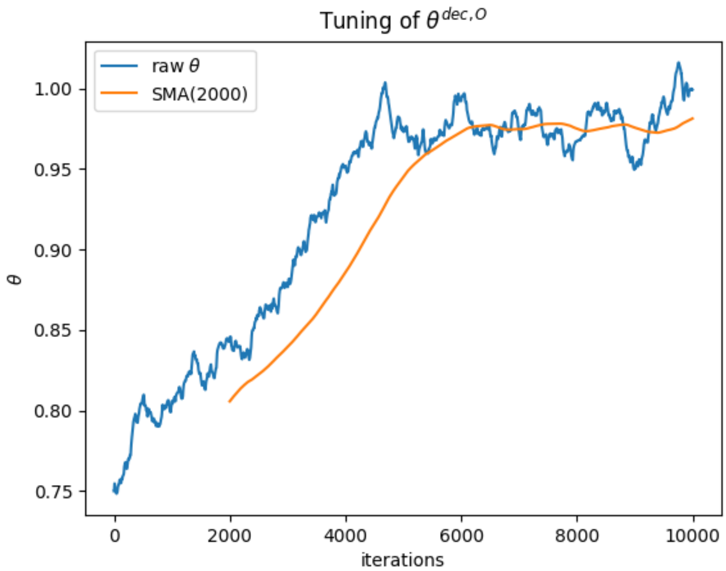


Figure 6.5: Tuning of $\theta^{dec,O}$ using Adam algorithm, with m = 5 and learning rate 0.001.

Parameter estimation: newsvendor model

One significant benefit of employing a newsvendor-type model to determine the parameter θ is eliminating the need for parameter tuning. However, one needs to determine the cost functions, $o(S_t)$ and $u(S_t)$, that we defined as in (5.36). Recall that the overshoot costs occur in case of an overestimation of demand, which results in fewer patients admitted and more waiting costs. In contrast, an undershoot estimation of demand can result in too many patients admitted and, therefore, in overuse of capacity.

The only parameter that needs to be chosen to use the newsvendor model is the value of α. In Table 6.2, we compare the value of a newsvendor model with all parameters with the 'full knowledge' policy. We also added the performance measures of We note that there is no significant difference between the various α values. We have decided to proceed with $\alpha = 0.1$, although this choice is arbitrary.

Comparison with other policies, with waiting list knowledge

We aim to evaluate how the proposed policy performs compared to alternative myopic policies, as proposed in Section 6.2. Figure 6.7 shows simulation runs with a random sample path with determined values for θ, using both the newsvendor model and the tuned version, and with separate tuning parameters. The values of the other parameters are, in this case, given to the model. In Figure 6.8, the same sample path is applied to myopic policies: 'cheapest direct cost', 'no forecast', and 'full knowledge'.

We can note that the DLA-CFA policy does lower the overcapacity peaks. For example, around week 95 in Figure 6.8(a), a high peak above the dotted line indicates overuse of capacity. In all sub-figures of Figure 6.7, the peak is lowered at the cost of not admitting

Table 6.2: Performance of different value of α for the newsvendor model applied to all three θ parameters at once, with 95% confidence intervals.

α	mean relative difference 'full knowledge' policy	mean absolute difference 'full knowledge' policy	average access time per urgency level	maximum average capacity exceeded nr of weeks
0.00001	$-0.299[-0.327, -0.270]$	$-15.99[-16.60, -15.40]$	0.46 [0.43, 0.49] 0.34 [0.31, 0.36] 0.21 [0.19, 0.24]	0.533 [0.497, 0.570]
0.01	$-0.299[-0.327, -0.270]$	$-15.99[-16.58, -15.40]$	0.46 [0.43, 0.49] 0.34 [0.31, 0.36] 0.21 [0.19, 0.24]	0.533 [0.497, 0.571]
0.1	$-0.299[-0.327, -0.270]$	$-15.98[-16.58, -15.38]$	0.46 [0.43, 0.49] 0.34 [0.31, 0.36] 0.21 [0.19, 0.24]	0.533 [0.497, 0.570]
0.25	$-0.300[-0.328, -0.269]$	$-15.98[-16.59, -15.39]$	0.46 [0.43, 0.49] 0.34 [0.31, 0.36] 0.22 [0.19, 0.24]	0.532 [0.597, 0.570]
0.5	$-0.301[-0.330, -0.272]$	$-16.06[-16.66, -15.46]$	0.47 [0.43, 0.50] 0.34 [0.31, 0.36] 0.22 [0.19, 0.24]	0.531 [0.495, 0.568]
1	$-0.302[-0.331, -0.273]$	$-16.09[-16, 69, -15.49]$	0.47 [0.43, 0.49] 0.34 [0.31, 0.36] 0.22 [0.19, 0.24]	0.530 [0.494, 0.567]

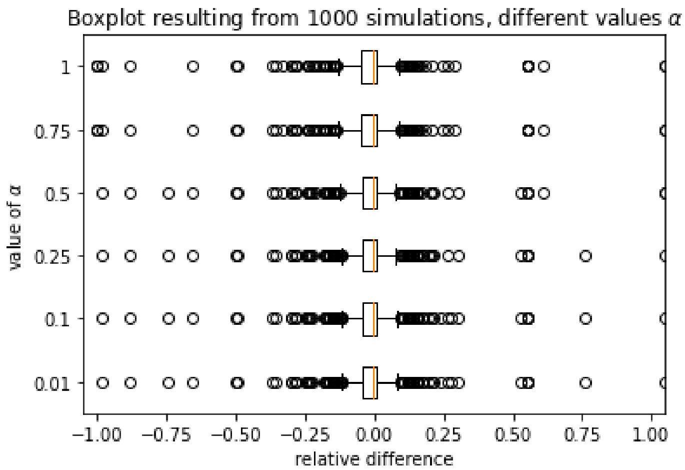
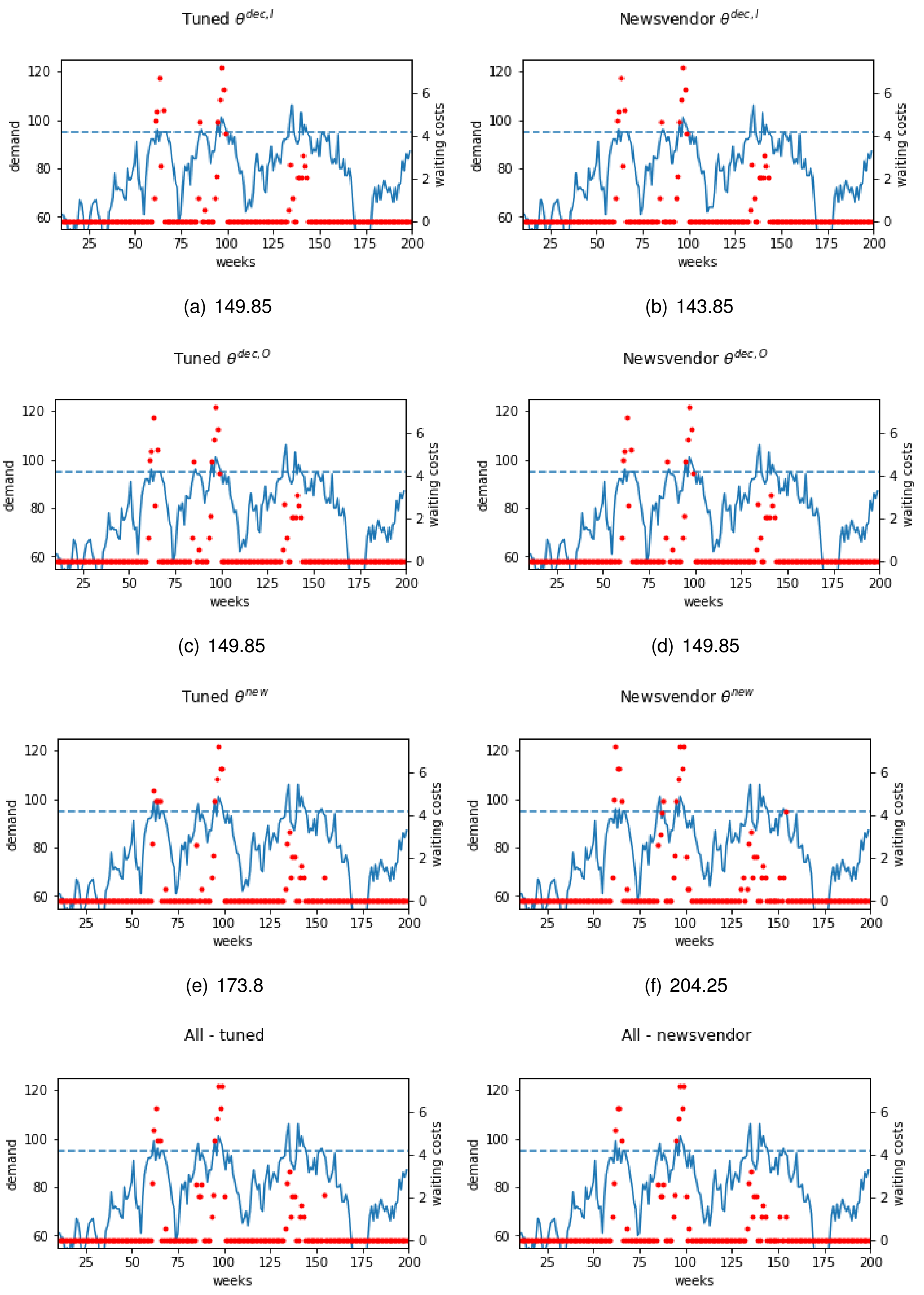


Figure 6.6: Boxplot of 1000 simulation runs with θ^{new}-newsvendor policy at different values for α with relative difference to a θ^{new}-tuned polict.

patients. We note that the total costs over the 200 weeks for all figures of the DLA-CFA policy are lower than for the 'cheapest cost' and 'no forecast' policies. The 'full knowledge' policy logically has the lowest costs, as it anticipates upcoming new arrivals and decision points.



(a) 149.85

(b) 143.85

(c) 149.85

(d) 149.85

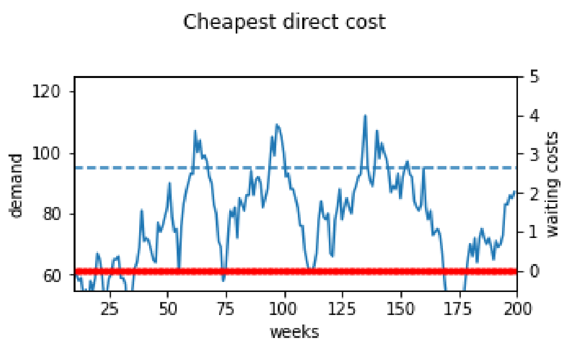
(e) 173.8

(f) 204.25

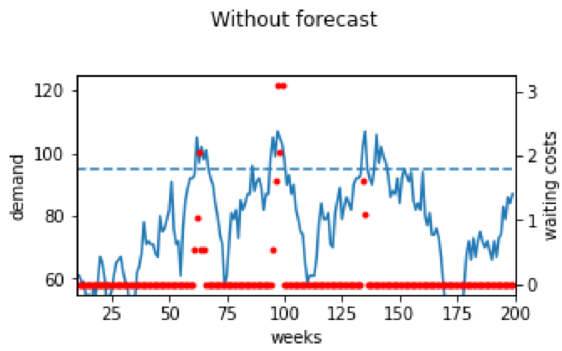
(g) 182.6

(h) 185.7

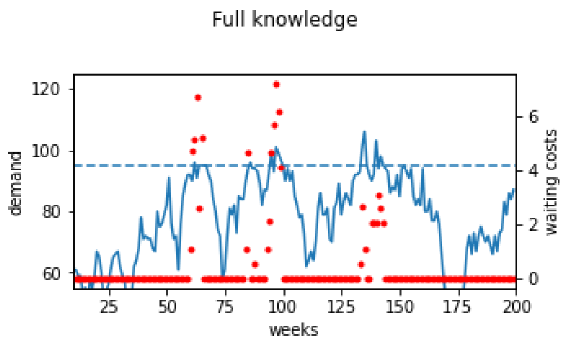
Figure 6.7: A simulation run of 200 weeks with different policies, the blue line is demand throughout the week, the dotted line is the capacity, and the red dots are waiting costs made in a given week. Costs are given for each simulation run of the given policy, with given tuned parameters. Figures (g) and (h) show the policies for which all three parameters are used.



(a) 223.5



(b) 205.35



(c) 149.85

Figure 6.8: A simulation run of 200 weeks with different policies, the blue line is demand throughout the week, the dotted line is the capacity, and the red dots are waiting costs made in a given week. Costs are given for each simulation run of the given policy, with given tuned parameters.

We would like to conclude that the proposed policy demonstrates superiority over myopic policies. However, further evaluation through increased simulation runs is necessary to substantiate this claim. To do so, we run the above simulations many more times. The resulting values of the simulation runs are unfair to compare, as they are highly correlated with the arriving demand. Hence, we will compare in two ways. First, we will use the 'cheapest direct cost' and 'full knowledge' policy to represent the worst and best performances. Second, we will compare the access times and capacity exceeded for our developed policy with those of other policies.

Cost comparison

For the first comparison, we compare each simulation run's resulting value to the corresponding value of the 'cheapest direct cost' and 'full knowledge' policies. We set the comparison so that the result is given a value between 0 and 1. A 1 tells us that the result is as good or better than the best policy, and a 0 tells us that the result is as bad or worse than the worst-performing policy.

The resulting average and median values for the different θ parameters separately are shown in Table 6.3. We note that not knowing the decision moments for either inpatient or outpatient, $\theta^{\text{dec},I}$ and $\theta^{\text{dec},O}$, results in a cost close to the full knowledge policy compared to

Table 6.3: Results of performance of 500 simulation runs for various policy settings, with mean and median values, both with 95% confidence interval.

policy	mean [CI]	median [CI]
θ^{new}-tuned	0.41 [0.38, 0.44]	0.42 [0.38, 0.48]
$\theta^{\text{dec},I}$-tuned	0.88 [0.85, 0.90]	1.0 [0.95, 1.0]
$\theta^{\text{dec},O}$-tuned	0.92 [0.91, 0.95]	1.0 [0.96, 1.0]

the demand for new arrivals, θ^{new}. This can be explained as the influence of the decision moment on the estimated demand is only for weeks T^{MDM} to K, after the current week t. On the other hand, new arrivals demand accounts for the full K weeks. Therefore, we note the influence of θ^{new} is larger than the influence of $\theta^{\text{dec},O}$ and $\theta^{\text{dec},I}$.

We also note, empirically, that the policy always performs as well as with full knowledge for high demand (high costs) simulation runs. In less demanding runs, the policy can perform worse but is often at the higher end in comparison with the 'cheapest direct cost' policy. We can explain this as for simulation runs with higher demand, the time of admission is more critical as capacity is sooner met or exceeded earlier if not chosen carefully.

Access times and exceeding capacity

For the second comparison, we do not only compare the costs but also the access time and overuse of capacity, as these are the assessment factors. Table 6.4 gives different policies' average access times and capacity exceeding, including 95% confidence intervals. We use a Wilcoxon signed rank test with 10000 samples to determine the confidence intervals. The average is determined over 3000 runs. For each run, we simulated 52 weeks to define the assessment factors on a yearly basis.

The 'full knowledge' policy performs best, cost-wise, and the 'cheapest direct cost' policy the worst, which we also expected. We note that the DLA-CFA policy, in both the tuned and the newsvendor versions, has similar access times as the 'full knowledge' policy but slightly worse performance in terms of exceeding demand. It has slightly better performance cost-wise than the 'no forecast' policy. However, there are different values for the other assessment factors. On average, the tuned and newsvendor version has 5.9 weeks yearly that the capacity is exceeded. This is lower than the 7.6 and 7.5 weeks for 'no forecast' and 'cheapest direct cost.' For such weeks, the overuse is, on average, 5.8 and 5.9 for the DLA-CFA policies but higher for the myopic policies, with 6.7 and 8.2, respectively. Furthermore, the demand peaks are also lower, with 7.7 and 7.8 average demand exceeding 9.4 and 10.8, respectively.

The trade-off for these benefits is longer access times. The lowest urgency class with the cheapest direct cost and relatively short forecast has zero access times on average. The admission policy for this class requires an average wait of 0.42 weeks, with a maximum average access time of 3.5 and 3.4 weeks. We conclude that the developed model's tuned and newsvendor versions behave similarly for these system settings.

Table 6.4: Performance of DLA-CFA policy for both a tuned as anewsvendor version, and 'full knowledge', 'no forecast' and 'cheapest direct cost' policies based on 3000 simulations of 52 weeks with knowledge of upcoming arrivals. (1) average cost, (2) median cost, (3) average number of weeks waiting per urgency level, (4) average maximum number of weeks waiting per urgency level (5) average number of weeks capacity exceeded per year per discipline, (6) average capacity exceeded per discipline, (7) average maximum peak overcapacity per discipline. Each includes a 95% confidence interval.

	Cost assessment	*Access time assessment*	
	(1) mean	(2) mean	(3) maximum
Full knowledge	70.6 [68.7, 72.5]	0.42 [0.31, 0.57] 0.34 [0.23, 0.47] 0.29 [0.18, 0.41]	3.9 [3.8, 4.0] 2.7 [2.6, 2.8] 1.8 [1.7, 1.9]
No forecast	85.6 [83.5, 87.8]	0.15 [0.09, 0.22] 0.12 [0.07, 0.16] 0.09 [0.04, 0.14]	1.4 [1.3, 1.5] 0.9 [0.8, 1.0] 0.5 [0.4, 0.6]
Cheapest direct cost	92.6 [90.3, 94.9]	0 0 0	0 0 0
DLA-CFA - Tuned	81.1 [79.0, 83.2]	0.40 [0.29, 0.52] 0.32 [0.22, 0.44] 0.24 [0.16, 0.33]	3.5 [3.4, 3.6] 2.5 [2.4, 2.6] 1.7 [1.6, 1.7]
DLA-CFA - Newsvendor	81.2 [79.1, 83.3]	0.42 [0.31, 0.53] 0.35 [0.24, 0.48] 0.24 [0.17, 0.33]	3.4 [3.3, 3.5] 2.4 [2.3, 2.5] 1.5 [1.5, 1.6]

	Demand assessment		
	(4) # weeks	(5) mean	(6) maximum
Full knowledge	4.7 [4.6, 4.8]	5.3 [5.2, 5.3]	6.4 [6.3, 6.5]
No forecast	7.6 [7.5, 7.8]	6.7 [6.6, 6.7]	9.4 [9.2, 9.5]
Cheapest direct cost	7.5 [7.4, 7.6]	8.2 [8.1, 8.3]	10.8 [10.6, 10.9]
DLA-CFA - Tuned	5.9 [5.8, 6.0]	5.8 [5.8, 5.9]	7.7 [7.5, 7.8]
DLA-CFA - Newsvendor	5.9 [5.8, 6.0]	5.9 [5.9, 6.0]	7.8 [7.6, 7.9]

6.3.2 Single-discipline and finite bed capacity

We now add a maximum number of inpatients admitted to the clinic. Remind that this complicates the transition function. Furthermore, the capacity is lowered, as fewer patients will be admitted.

Table 6.5: Settings for a system model, single-discipline with finite bed capacity.

$C = 1$	$\kappa_c = \{0.55, 1.05, 2.05\}$
$U = 3$	$\kappa_u = \{1.5\}$
$r = \{75\}$	
$\mathcal{P} = \{\{1\},\{2\},\{3\}\}$	$\mathcal{Q} = \{\{2\},\{3\}\}$
$p_{\mathcal{P}}^{\text{plan}} = \{0.6, 0.2, 0.2\}$	$p_{\mathcal{Q}}^{\text{plan}} = \{0.25, 0.75\}$
$T_O^{\text{MDM}} = 2$	$T_I^{\text{MDM}} = 2$
$T_O^{\text{start}} = 8$	$T_I^{\text{start}} = 5$
$\lambda_O = 1$	$\lambda_I = 2$
$p_O^{\text{cont}} = \{0.8, 0.6, 0\}$	$p_I^{\text{cont}} = \{0.6, 0.4, 0\}$
	$n_I^{\max} = 15$

Parameter estimation: stochastic approximations

Due to a limited number of beds, we now expect the value for θ^{new} to be lower. In Figure 6.9, the tuning results of θ^{new} are shown for 10000 iterations. The tuning indicates an optimal value of about 2.5.

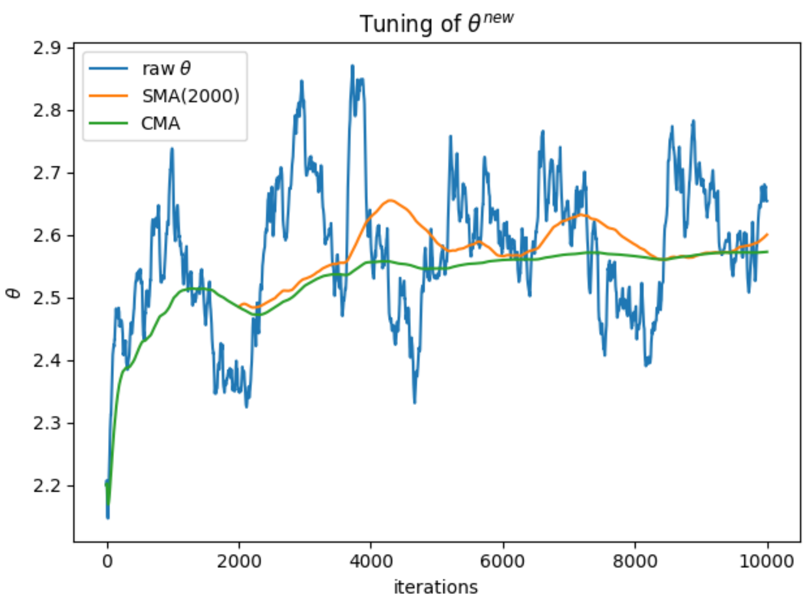


Figure 6.9: Tuning of θ^{new} using Adam algorithm, with m = 5 and learning rate 0.001.

In (5.13), we see that the possible extra demand from the inpatient decision moment is state-dependent. $\phi_I^{\text{dec}}(S_t, \tau)$ is the summed demand of all patients with a decision moment in week $t+\tau$. We want to find value $\theta^{\text{dec},I}$ as a multiplier to forecast the demand coming from these patients. As we cannot have more patients suddenly, we need to have $\theta^{\text{dec},I} \in (0, 1)$. In Figure 6.10, we find a value of $\theta^{\text{dec},I} = 0.52$. Similarly, for the tuning of $\theta^{\text{dec},O}$, we find a value of 0.74, as can be seen in Figure 6.15.

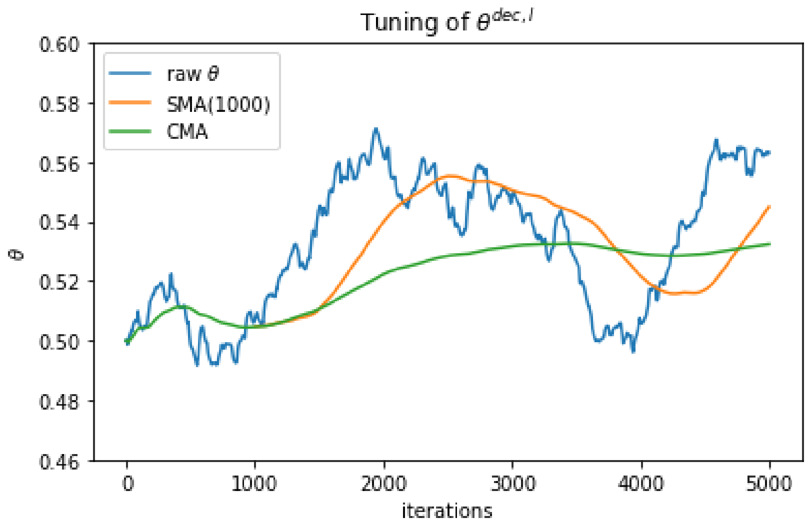


Figure 6.10: Tuning of $\theta^{dec,I}$ using Adam algorithm, with m = 5 and learning rate 0.001.

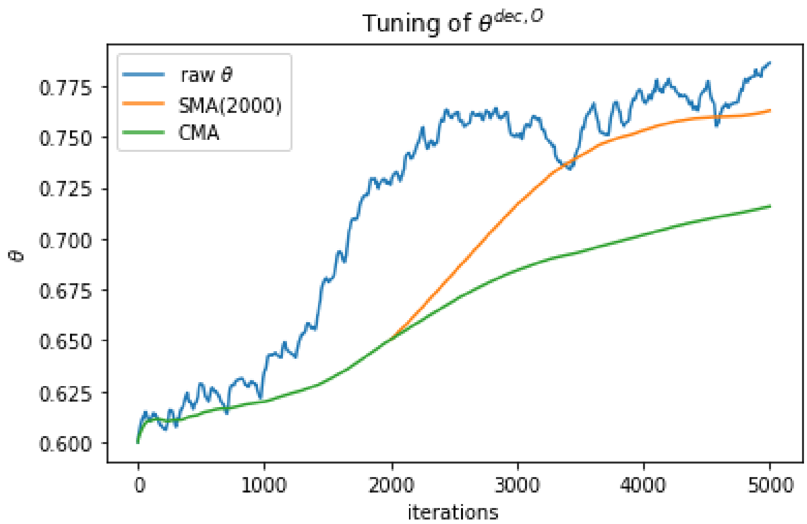


Figure 6.11: Tuning of $\theta^{dec,O}$ using Adam algorithm, with m = 5 and learning rate 0.001..

Parameter estimation: newsvendor model

The only parameter that needs to be defined for the newsvendor version of the DLA-CFA policy is the value of α, which determines the ratio at which costs under the capacity limit are applied. In Table 6.6, we compare the average results of various α values in comparison with a 'full knowledge' policy. We based this on 1,500 simulation runs, where for each run, the same sample path is used for the various α. It seems that the lower α, the better the performance, as it seems that with higher alpha, more priority is given to the exceeding of capacity as this performance measure is slightly higher there. However, we should note this difference is insignificant, and therefore we cannot conclude on the best setting for α. We choose $\alpha = 0.1$ for the remaining of the model, but we note that better research should be done to establish this hyperparameter more carefully.

Table 6.6: Performance of different value of α for the newsvendor model applied to all three θ parameters at once, based on 1500 simulation runs.

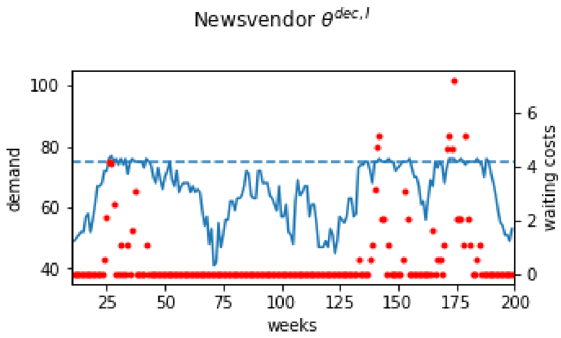
α	mean relative difference 'full knowledge' policy	mean absolute difference 'full knowledge' policy	average access time per urgency level	average capacity exceeded nr of weeks
0.00001	$-0.410[-0.445, -0.382]$	$-59.8[-65.1, -54.8]$	0.54 [0.50, 0.58] 0.39 [0.36, 0.43] 0.27 [0.24, 0.31]	0.104 [0.096, 0.113]
0.01	$-0.410[-0.444, -0.381]$	$-59.8[-65.1, -54.8]$	0.54 [0.50, 0.58] 0.39 [0.36, 0.43] 0.27 [0.24, 0.31]	0.104 [0.096, 0.113]
0.1	$-0.410[-0.445, -0.381]$	$-59.7[-65.7, -54.8]$	0.54 [0.50, 0.58] 0.39 [0.36, 0.43] 0.28 [0.24, 0.31]	0.104 [0.095, 0.112]
0.25	$-0.413[-0.448, -0.385]$	$-60.7[-66.0, -55.6]$	0.54 [0.50, 0.58] 0.39 [0.36, 0.43] 0.27 [0.24, 0.31]	0.102 [0.094, 0.111]
0.5	$-0.415[-0.449, -0.386]$	$-60.8[-66.1, -55.6]$	0.54 [0.51, 0.58] 0.40 [0.36, 0.43] 0.28 [0.24, 0.31]	0.101 [0.093, 0.109]
1	$-0.416[-0.451, -0.387]$	$-61.2[-66.9, -55.9]$	0.55 [0.51, 0.59] 0.40 [0.36, 0.43] 0.27 [0.24, 0.31]	0.099 [0.091, 0.108]

Comparison with other policies, with waiting list knowledge

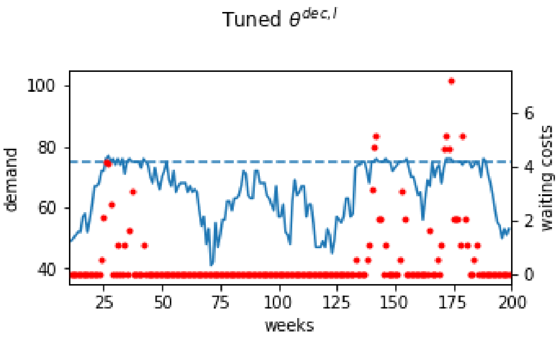
We compare the DLA-CFA policy with tuned values and newsvendor model determined values with a policy that uses the 'cheapest direct cost' (= admit all patients directly), a 'no forecast' policy ($\theta = 0$), and a 'full knowledge' policy, which are further detailed in 6.2. We aim to evaluate the efficiency of the proposed policy in comparison to the proposed myopic policies.

In Figure 6.12, the resulting demand and costs for not admitting patients for one sample path of 200 weeks are shown for the different θ, as well as combined and for both tuned and newsvendor versions of the model. In Figure 6.13, we show the same sample path under the 'cheapest direct cost', 'no forecasting' and 'full knowledge' policies.

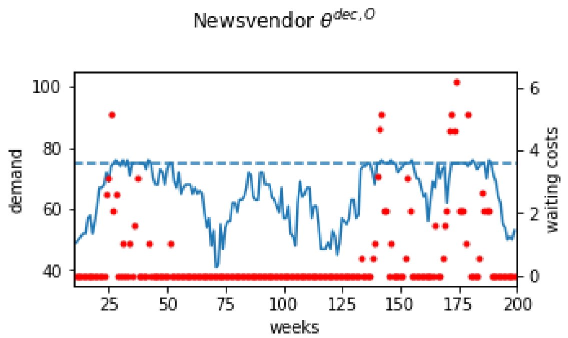
The designed policy tries to keep exceeding demand as low as possible. The newsvendor model especially seems to perform well. The tuning model for θ^{new} appears to be performing worse, as due to the limited bed availability, the demand for new arrivals is state-dependent, and it is not easily described by one number.



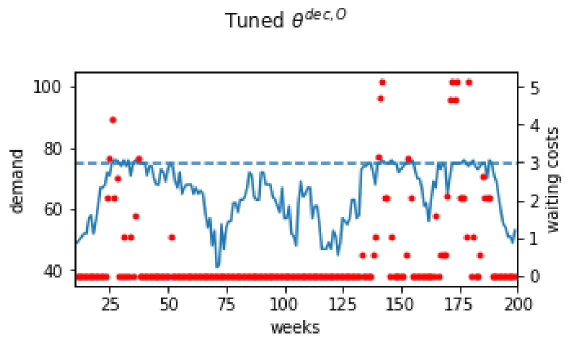
(a) 115.3



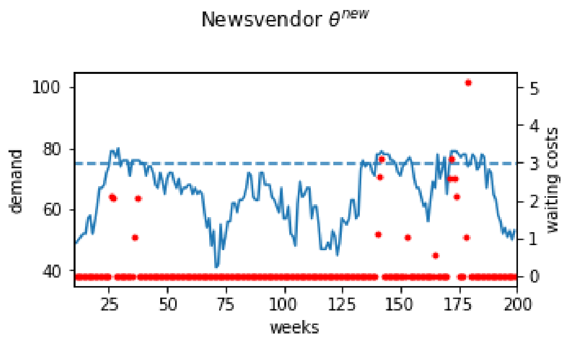
(b) 112.7



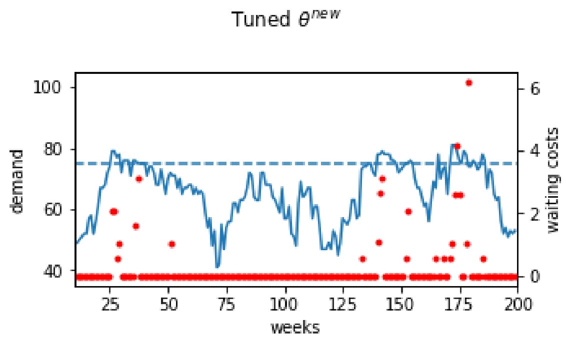
(c) 112.1



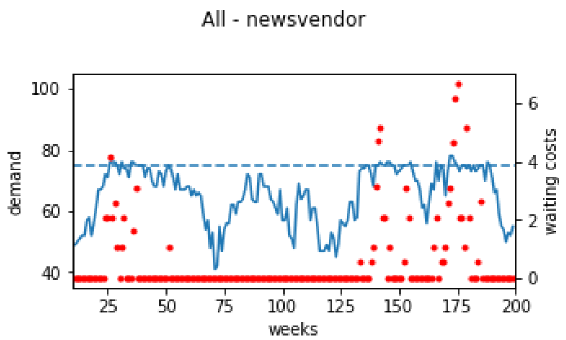
(d) 112.0



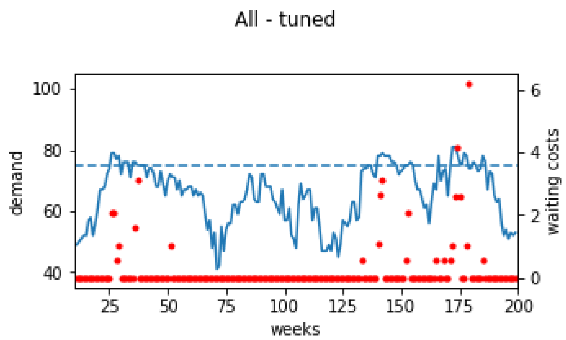
(e) 149.4



(f) 159.8

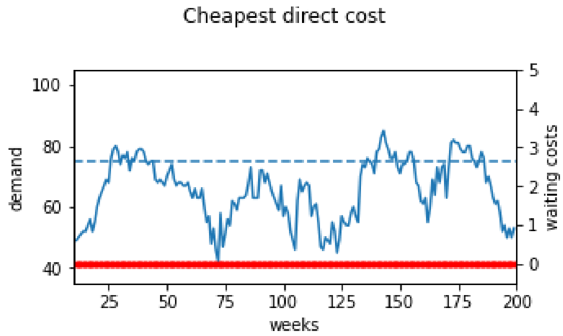


(g) 121.3

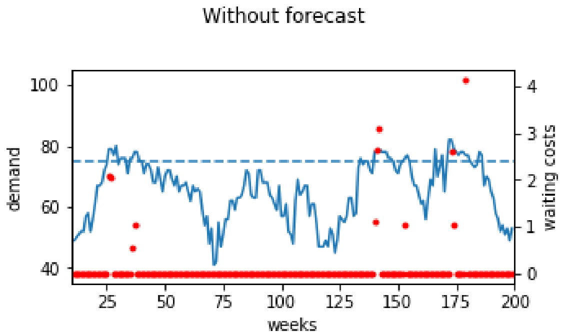


(h) 149.4

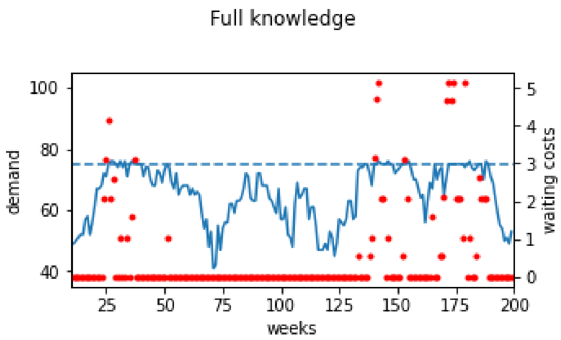
Figure 6.12: Simulation run of 200 weeks with different policies, the blue line is demand throughout the week, dotted line in capacity, red dots are waiting costs made in a given week. Costs are given for each policy.



(a) 196.5



(b) 164.0



(c) 110.0

Figure 6.13: Simulation run of 200 weeks with different policies with an estimation of new inpatient demand, the blue line is demand throughout the week, red dots are waiting costs made in a given week. Costs are given for each policy.

We further analyze this by running 3000 simulation runs. For each run, we simulated 52 weeks to define the assessment factors on a yearly basis. Table 6.8 includes the average cost as the values for other assessment factors, including 95% confidence bounds. We added the respective resulting boxplots in Figure 6.16 and 6.17. The cost values do not follow a normal distribution but are highly skewed; see Figure 6.14. So, we use a one-sample Wilcoxon signed rank test with 10000 samples to determine the confidence bounds.

Note that the median cost is significantly lower. The median suppresses the high-cost results, of which only a few exist. Regarding the median costs, the 'no forecast' does not necessarily perform worse than the developed method. We conclude that our method performs better in especially high-demand simulation runs, similar to the infinite bed capacity system.

Again, the results show a cost difference in performance between the DLA-CFA and myopic policies, of which the 95% confidence bounds do not overlap. There is a slight average cost difference between the tuned version and the newsvendor version of the DLA-CFA policy, where the newsvendor model performs slightly better, but this is not significant. However, the difference in access time and capacity use assessment factors is more interesting. Whereas the tuned version has a lower access time along all urgency levels at the cost of more capacity that is exceeded, the newsvendor version behaves the opposite with longer

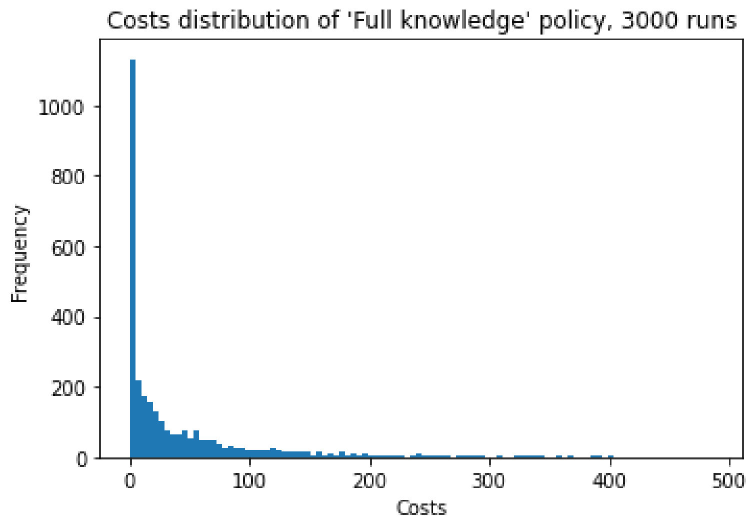


Figure 6.14: Costs distribution of resulting costs for a 'full knowledge' policy for 3000 simulation runs of 50 weeks.

access times and less exceeded demand. See the average waiting weeks, [0.23, 0.16, 0.10] versus [0.41, 0.27, 0.19], and the average number of weeks that the capacity is exceeded, 6.0 weeks versus 2.9 weeks, and the apparent difference in Figures 6.16, and 6.17.

We did not see the difference in assessment factors at the infinite bed capacity, so we expect the behavior to occur due to a different arrival distribution for each state. The newsvendor model does consider this but it is not directly included in the tuned version of the model, as this value in not state-dependent.

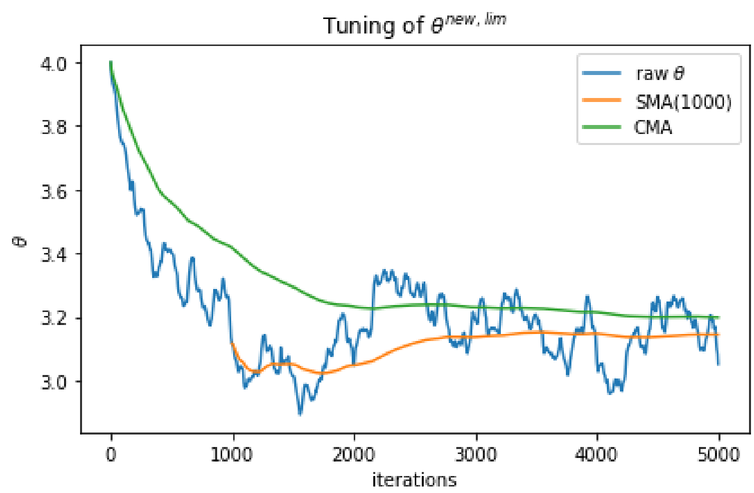


Figure 6.15: Tuning of $\theta^{new,lim}$ using Adam algorithm, with m = 5 and learning rate 0.001. SMA(1000) indicates a simple moving average of the last 1000 iterations. CMU indicates the cumulative moving average.

To determine if this is the case, we introduce $\theta_{\tau}^{\text{new, lim}}$ which is defined as follows:

$$\theta^{\text{new, lim}}(S_t, \tau) = \begin{cases} 0, & \text{if} \sum_{\substack{\mathbf{q}\in\mathcal{Q},\ell\in\mathcal{L}_I,\\ m\in\mathcal{M}_I}} (\hat{\imath}_{\mathbf{p}\ell m})_{\tau} = n_I^{\max}, \\ \theta^{\text{new}}, & \text{otherwise.} \end{cases} \tag{6.1}$$

In Tables 6.8 and 6.9, we add the performance of this new parameter. However, we note that the average costs are not lower, and the other assessment factor is performed similarly as with the previous θ^{new}. Although $\theta^{\text{new, lim}}$ does take into account the impossibility of a new arrival in case all beds are occupied. The parameter fails to differentiate between states with, for instance, one available bed and those with ten available beds, while these scenarios may exhibit distinct arrival probability distributions theoretically. Due to limited bed availability, the distribution functions differ between these states. To solve this, one can define θ^{new} for every number of admitted patients, but in this way, every separate parameter option also needs to be tuned.

Compare with other policies, without waiting list knowledge

In the comparison above, we designed the model with complete information about future outpatient arrivals, which is not feasible in real life. Therefore, we developed a recourse model in Section 5.2.1 to determine the best action without knowing the arrival to the waiting list in the upcoming forecast timeinterval. The recourse model

In Table 6.7, we provide the results of using the recourse model with N_{batch} is 1, 10, and 25. This table shows that using a limited version of the recourse model, with low N_{batch}, does not increase the resulting costs. An explanation for this is that the maximum number of patients that can arrive on the waiting list is relatively small if $\lambda_O = 1$. Therefore, the difference in the influence of possible arrivals is relatively slight. Furthermore, using more iterations than necessary increases the resulting computational costs.

Table 6.9 shows the results averaged over 1000 simulation runs using the recourse model with ten iterations. The costs are similar to those in Table 6.8 above. The addition of the recourse model does not highly influence the performance.

Table 6.7: Average costs and average computational time for applying the recourse model for 1, 10, and 25 iterations for 'full knowledge' and the two DLA-CFA policies, over 250 simulation runs.

	1 iteration	10 iterations	25 iterations
	Average costs		
Full knowledge	43.2 [34.3, 52.6]	42.5 [34.2, 52.7]	43.4 [34.3, 53.0]
All - Tuned	57.9 [47.4, 69.7]	58.1 [47.3, 69.7]	58.0 [47.1, 69.8]
All - Newsvendor	55.0 [44.0, 66.8]	54.7 [44.0, 66.6]	54.1 [43.7, 65.9]
	Average computational time (seconds)		
Full knowledge	1.7 ± 0.1	16.3 ± 0.8	41.1 ±2.0
All - Tuned	1.4 ± 0.1	13.3 ± 0.5	33.4 ± 1.2
All - Newsvendor	2.4 ± 0.1	33.7 ± 1.2	84.1 ± 3.1

Table 6.8: Performance of DLA-CFA policy for both a tuned as a newsvendor version, and 'full knowledge', 'no forecast' and 'cheapest direct cost' policies based on 3000 simulations of 52 weeks with knowledge of upcoming arrivals. (1) average cost, (2) median cost, (3) average number of weeks waiting per urgency level, (4) average maximum number of weeks waiting per urgency level (5) average number of weeks capacity exceeded per year per discipline, (6) average capacity exceeded per discipline, (7) average maximum peak overcapacity per discipline. Each includes a 95% confidence interval.

	Cost assessment		*Access time assessment*	
	(1) mean	(2) median	(3) mean	(4) maximum
Full knowledge	43.1 [40.6, 45.7]	12.45 [10.8, 14.7]	0.33 [0.32, 0.35] 0.24 [0.22, 0.25] 0.16 [0.15, 0.17]	3.5 [3.4, 3.7] 2.5 [2.4, 2.6] 1.6 [1.5, 1.7]
No forecast	66.3 [63.0, 69.7]	25.4 [21.1, 28.9]	0.11 [0.10, 0.12] 0.08 [0.07, 0.09] 0.05 [0.04, 0.06]	1.4 [1.3, 1.5] 0.9 [0.8, 1.0] 0.5 [0.4, 0.6]
Cheapest direct cost	77.1 [73.1, 81.2]	26.0 [20.0, 30.0]	0 0 0	0 0 0
All - Tuned	58.8 [55.7, 62.0]	20.8 [18.5, 24.6]	0.23 [0.22, 0.24] 0.16 [0.15, 0.17] 0.10 [0.09, 0.11]	2.4 [2.3, 2.5] 1.7 [1.6, 1.8] 1.1 [1.0, 1.2]
All - Newsvendor	54.9 [51.4, 57.6]	18.4 [15.6, 21.5]	0.41 [0.39, 0.43] 0.29 [0.27, 0.31] 0.20 [0.19, 0.22]	4.1 [4.0, 4.3] 3.0 [2.8, 3.1] 2.1 [1.9, 2.2]
All - Tuned Update	58.7 [56.2, 61.3]	20.4 [16.8, 24.0]	0.18 [0.12, 0.26] 0.11 [0.07, 0.17] 0.08 [0.04, 0.13]	2.6 [2.5, 2.7] 1.8 [1.7, 1.9] 1.1 [1.0, 1.2]

	Demand assessment		
	(5) # weeks	(6) average	(7) maximum
Full knowledge	2.4 [2.3, 2.6]	2.2 [2.1, 2.2]	1.7 [1.6, 1.8]
No forecast	7.5 [7.2, 7.8]	3.8 [3.7, 3.9]	4.5 [4.3, 4.6]
Cheapest direct cost	7.5 [7.2, 7.8]	5.2 [5.1, 5.2]	5.6 [5.4, 5.8]
All - Tuned	6.0 [5.7, 6.3]	3.3 [3.2, 3.4]	3.7 [3.5, 3.8]
All - Newsvendor	2.9 [2.8, 3.1]	2.2 [2.1, 2.3]	1.9 [1.8, 2.0]
All - Tuned Update	5.9 [5.7, 6.0]	3.2 [3.1, 3.2]	3.6 [3.5, 2.7]

Table 6.9: Performance of DLA-CFA policy for both tuned as newsvendor version, and 'full knowledge', 'no forecast' and 'cheapest direct cost' policies based on 3000 simulations of 52 weeks with knowledge of upcoming arrivals. (1) average cost, (2) median cost, (3) average number of weeks waiting per urgency level, (4) average maximum number of weeks waiting per urgency level (5) average number of weeks capacity exceeded per year per discipline, (6) average capacity exceeded per discipline, (7) average maximum peak overcapacity per discipline. Each includes a 95% confidence interval.

	Cost assessment		*Access time assessment*	
	(1)	(2)	(3)	(4)
Full knowledge	43.3 [40.8, 46.0]	12.5 [10.8, 14.7]	0.33 [0.31, 0.35] 0.24 [0.22, 0.25] 0.16 [0.15, 0.17]	3.6 [3.4, 3.7] 2.5 [2.4, 2.7] 1.6 [1.5, 1.7]
No forecast	66.3 [63.1, 69.6]	25.6[21.1, 28.9]	0.11 [0.10, 0.12] 0.08 [0.07, 0.09] 0.05 [0.04, 0.06]	1.4 [1.3, 1.5] 0.9 [0.8, 1.0] 0.5 [0.4, 0.6]
Cheapest direct cost	77.1 [73.1, 81.2]	26.0 [20.0, 30.0]	0 0 0	0 0 0
All - Tuned	58.9 [55.8, 62.0]	20.8 [18.3, 24.6]	0.23 [0.22, 0.24] 0.16 [0.15, 0.17] 0.10 [0.09, 0.11]	2.4 [2.3, 2.6] 1.7 [1.6, 1.8] 1.1 [1.0, 1.2]
All - Newsvendor	54.5 [51.4, 57.6]	18.2 [15.2, 21.3]	0.41 [0.39, 0.43] 0.29 [0.27, 0.31] 0.20 [0.18, 0.22]	4.2 [4.0, 4.4] 3.0 [2.8, 3.1] 2.1 [1.9, 2.2]
All - Tuned Update	58.7 [56.2, 61.3]	0.18 [0.12, 0.26] 0.11 [0.07, 0.17] 0.08 [0.04, 0.13]	2.6 [2.5, 2.7] 1.8 [1.7, 1.9] 1.1 [1.0, 1.2]	

		Demand assessment	
	(5)	(6)	(7)
Full knowledge	2.7 [2.5, 2.8]	2.2 [2.1, 2.2]	1.7 [1.6, 1.8]
No forecast	7.5 [7.2, 7.8]	3.8 [3.7, 3.9]	4.5 [4.3, 4.6]
Cheapest direct cost	7.5 [7.2, 7.8]	5.2 [5.1, 5.2]	5.6 [5.4, 5.8]
All - Tuned	6.0 [5.7, 6.3]	3.3 [3.2, 3.4]	3.7 [3.5, 3.8]
All - Newsvendor	3.0 [2.9, 3.1]	2.2 [2.1, 2.3]	1.9 [1.8, 2.0]
All - Tuned Update	5.9 [5.7, 6.0]	3.2 [3.1, 3.2]	3.6 [3.5, 2.7]

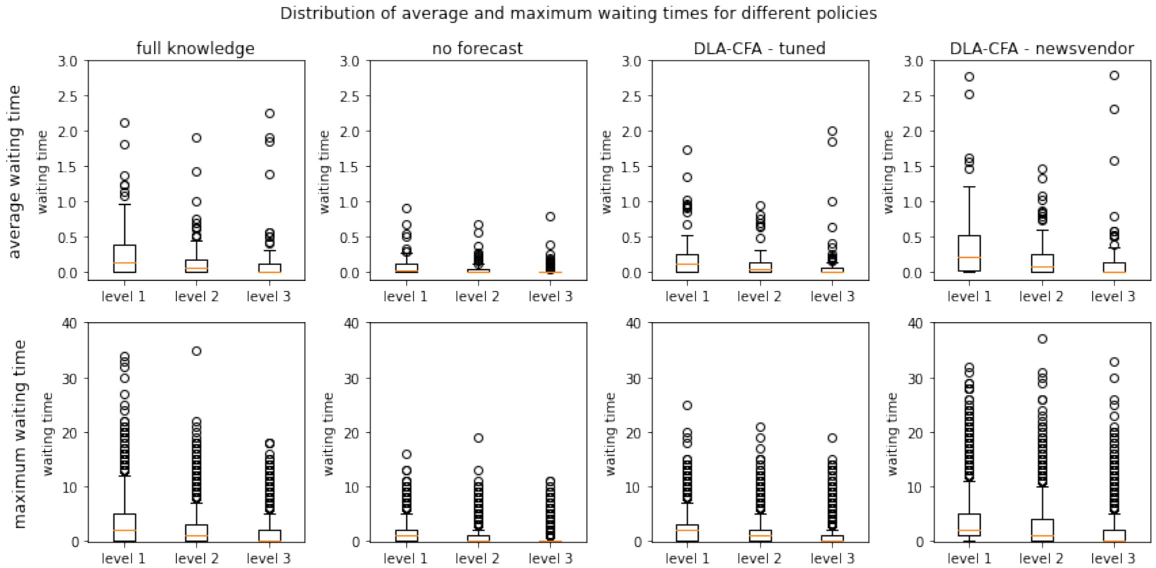


Figure 6.16: Boxplots of average access times and maximum access times per urgency level for various policies based upon 3000 simulation runs.

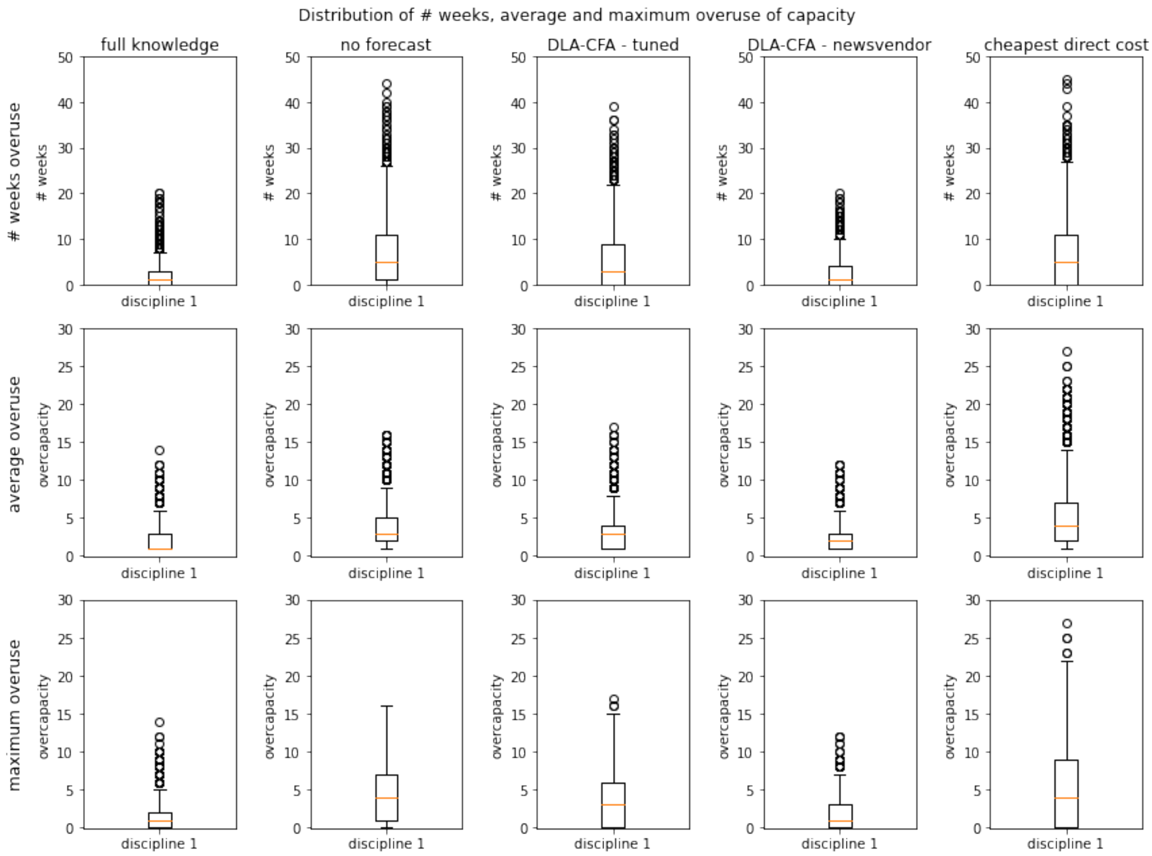


Figure 6.17: Boxplots of the number of weeks that capacity is exceeded, average overuse of capacity, and maximum overage of capacity for various policies based upon 3000 simulation runs.

6.4 A policy for the rehabilitation clinic

In this section, we develop a policy that fits the rehabilition clinic *Merem*. Before implementing the policy , we conduct a data analysis to configure the system settings. This is described in Section 6.4.1. We determine the values of the parameters in Section 6.4.2, after which we conclude whether or not the policy works in Section 6.4.3.

6.4.1 Data analysis

We use the same data used in Section 2 to determine the system settings. We determine the parameters for the (Poisson) arrival distribution, the probabilities for discharge after a decision moment, and the different treatment plans.

From the data analysis in Section 2.3.1, we can determine the average arrivals for both inpatients and outpatients. We set $\lambda_I = 3.0$, and $\lambda_O = 4.0$.

We can determine the value for T^{start} and T^{MDO} based upon the Figure 2.7, and expect knowledge. In reality, all patients are discussed every two weeks. Therefore, setting $T^{\text{MDM}} = 2$ for both groups is logical. Furthermore, based on the data, we determine $T_I^{\text{start}} = 4$ and $P_I^{\text{cont}} = \{0.75, 0.60, 0.33, 0.40, 0\}$, and $T_O^{\text{start}} = 8$ and $P_O^{\text{cont}} = \{0.75, 0.70, 0.70, 0.40, 0\}$.

Treatmentplan of inpatient and outpatients

In reality, the treatment plan can differ weekly per patient. Furthermore, as we only have data available on given treatments and not on asked treatments, it is challenging to specify fixed treatment plans. This is something that requires further research.

Instead, we perform an analysis of the most occurring sets of treatments per patient per week. Tables 6.10 and 6.11 are the eight most occurring combinations shown for inpatients and outpatients, respectively. We note that most combinations are very alike. For simplicity, we therefore choose the following:

$$\begin{aligned}
\mathcal{P} &= \{\{1,1,1\},\{1,1,2\},\{2,2,0\}\} \\
\mathcal{Q} &= \{\{3,4,4\},\{3,4,1\}\} \\
p_{\mathcal{P}}^{\text{plan}} &= \{0.6, 0.2, 0.2\} \\
p_{\mathcal{Q}}^{\text{plan}} &= \{0.5, 0.5\}
\end{aligned}$$

Capacity limits

We can set the capacity limits in two ways: look into the demand used on average weekly in the provided dataset or define the capacity by given clinic contracts. We choose the same capacities as in the discrete event simulation, see Table 2.2.

$$\mathbf{r}_1 = 132, 165.6, 98$$

Table 6.10: Top seven most occurring combinations of treatment per discipline in the dataset, for inpatients

occupational therapy	physiotherapy	speech therapy	occurrence
3	4	4	23
2	3	1	22
3	4	1	19
3	4	3	17
4	4	3	16
3	4	2	16
4	5	1	15

Table 6.11: Top seven most occurring combinations of treatment per discipline in the dataset, for outpatients

occupational therapy	physiotherapy	speech therapy	occurrence
1	1	1	81
1	1	2	34
1	2	1	31
1	2	2	28
2	1	2	27
2	2	1	26
2	2	2	25

However, we note that the set capacity for speech capacity is relatively low, whereas, from expert knowledge, we know that this group of therapists often work more for the neurological/surgical department than officially registrated.

On the other hand, the capacity for occupational therapy and physiotherapy is relatively high. Due to holidays, pregnancy, et cetera, the capacity is lower in reality. Due to these imbalances, the model does not precisely react to all three disciplines, but only to speech therapy; see as an example Figure 6.21. Therefore, a second and realistic capacity array should be specified to evaluate for a more optimal capacity.

$$\mathbf{r}_2 = 120, 145, 110$$

6.4.2 Parameter estimation

An overview of the settings used to resemble the rehabilitation clinic is given in Table 6.12. We are unsure about the cost settings for the rehabilitation clinic. We will analyze this further in Section 6.4.4. We first tune the value for θ^{new}, $\theta^{\text{dec},O}$ and $\theta^{\text{dec},I}$ to see how the developed policy reacts to the multi-disciplinary settings.

Table 6.12: Settings for a system resembling the rehabilitation clinic.

$C = 3$	$\kappa_c = \{1.5, 1.5, 1.5\}$
$U = 3$	$\kappa_u = \{0.55, 1.05, 2.05\}$
$\mathbf{r}_1 = \{132, 165.6, 98\}$	$\mathbf{r}_2 = \{120, 145, 110\}$
$\mathcal{P} = \{\{1,1,1\},\{1,1,2\},\{2,2,0\}\}$	$\mathcal{Q} = \{\{3,4,4\},\{2,3,1\}\}$
$p_{\mathcal{P}}^{\text{plan}} = \{0.6, 0.2, 0.2\}$	$p_{\mathcal{Q}}^{\text{plan}} = \{0.5, 0.5\}$
$T_O^{\text{MDM}} = 2$	$T_I^{\text{MDM}} = 2$
$T_O^{\text{start}} = 8$	$T_I^{\text{start}} = 4$
$\lambda_O = 4.0$	$\lambda_I = 3.0$
$p_O^{\text{cont}} = \{0.66, 0.4, 0.5, 0.5, 0\}$	$p_I^{\text{plan}} = \{0.75, 0.6, 0.4, 0.25, 0\}$
	$n_I^{\max} = 25$

Parameter estimation: stochastic approximation

In Figures 6.18 6.19 and 6.20, we show the tuning results for both capacity setting $\mathbf{r}_1$ and $\mathbf{r}_2$. For capacity $\mathbf{r}_1$, discipline 3, which describes speech therapy, is the primary discipline affecting the outcome. The other disciplines exhibit no response to the tuning adjustments. Additionally, we observe that the parameters associated with the decision moment fall below zero. Note again that the tuning value should be such that the best decisions are made concerning the linear program that defines the DLA-CFA policy. The demand forecast exhibits a decline when the value is zero or below. Consequently, this leads to an increase in patient admissions.

Conversely, if the demand forecast were higher, the associated parametrized costs would also increase, resulting in decreased patient admissions. As the parametrized capacity costs continue to rise, patient admissions will be delayed for an extended period. Thus, in this context, a low or even negative forecast ultimately yields a reduction in overall costs, as it correlates with shorter access times for patients.

For capacity $\mathbf{r}_2$, all three disciplines are tuned corresponding and do not reach under zero. We conclude to the following value $\theta^{\text{new}} = \{7, 7, 5.3\}$, $\theta^{\text{dec},O} = 0.35$ and $\theta^{\text{dec},I} = 0.25$.

Parameter estimation: newsvendor model

Firstly, we analyzed the influence of α similarly to the previous section, see Table 6.6. Again, we note that the influence of different α is not significant. We choose to use $\alpha = 0.1$ here.

We note that the computation time of using the newsvendor model has increased compared to the simple single-disciplinary system. The calculation per decision takes about 2 seconds, compared with 0.2 seconds for the smaller system analyzed in the section earlier. This is due to the increase in the number of patients in the system. Therefore, the number of decision moments increases, resulting in a more complex joint probability distribution.

Furthermore, given the constraints on the number of inpatients that can be admitted, the newsvendor model must incorporate these limitations into its analysis. To achieve this, it estimates the anticipated number of inpatients and subsequently adjusts the forecast in response. This process is elaborated upon in Algorithm 4.

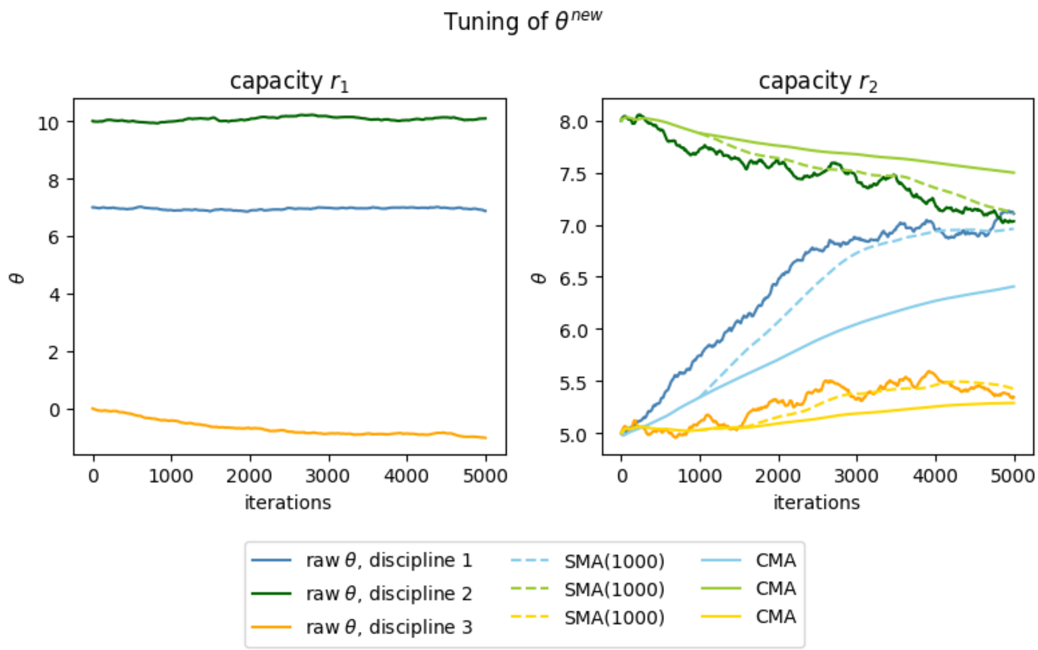


Figure 6.18: Tuning of θ^{new} for both system setting with $\mathbf{r}_1$ and $\mathbf{r}_2$. SMA(1000) indicates a simple moving average of the last 1000 iterations. CMU indicates the cumulative moving average.

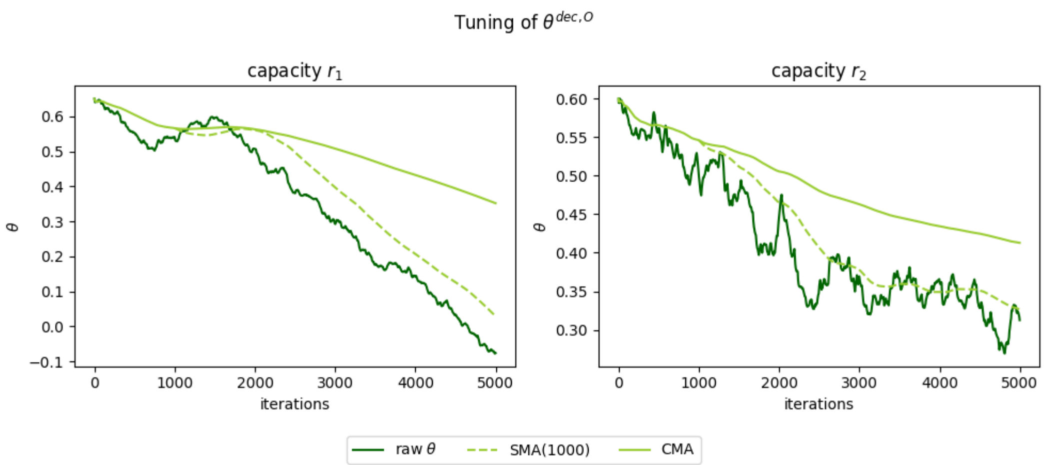


Figure 6.19: Tuning of $\theta^{dec,O}$ for both system setting with $\mathbf{r}_1$ and $\mathbf{r}_2$. SMA(1000) indicates a simple moving average of the last 1000 iterations. CMU indicates the cumulative moving average.

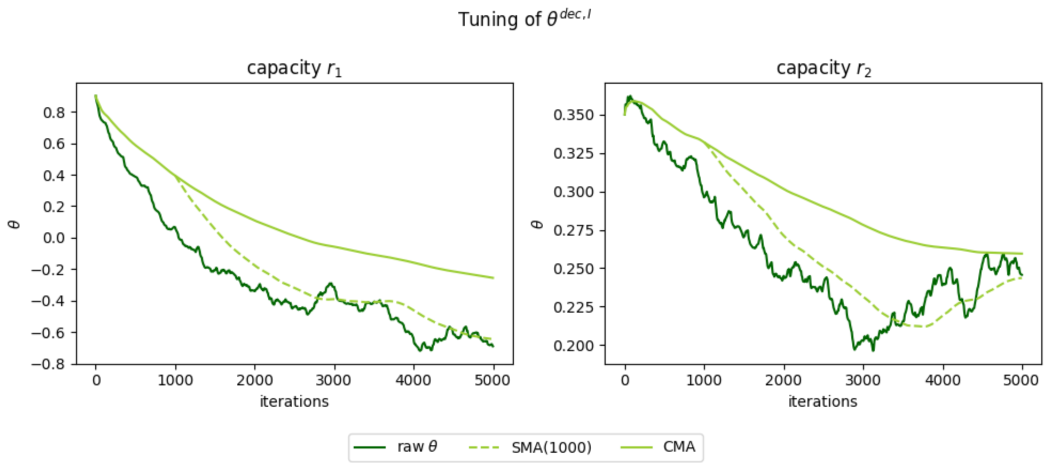


Figure 6.20: Tuning of $\theta^{dec,I}$ for both system setting with $\mathbf{r}_1$ and $\mathbf{r}_2$. SMA(1000) indicates a simple moving average of the last 1000 iterations. CMU indicates the cumulative moving average.

6.4.3 Comparison with other policies

We perform a simulation run for both capacity settings. Figure 6.21 shows the demand throughout the weeks and the waiting costs for various policies for capacity $\mathbf{r}_1$. The analysis reveals that the capacities of Discipline 1 and Discipline 2 have never been reached. At the same time, Discipline 3 serves as the critical discipline, with its capacity being exceeded for many weeks. Additionally, it is observed that the policy based on the 'cheapest direct cost' demonstrates the best performance during this simulation run.

Conversely, the 'full knowledge' policy does not yield better results than the DLA-CFA tuned policy. Table 6.13 gives the average costs for the different policies based on 1000 simulation runs. All policies perform similarly, apart from the newsvendor model. We think that the performance of the newsvendor model is worse, as it puts more priority on forecasting demand. As a result, it often admits fewer patients, leading to longer access times. However, since the demand for speech therapy never falls below capacity, this means that admissions remain low for an extended period with higher costs.

	Average cost
Full knowledge	443.4 [431.1, 456.1]
No forecast	430.7 [420.0, 441.5]
Cheapest direct cost	432.1 [421.8, 442.3]
DLA-CFA - Tuned	431.0 [420.3, 441.7]
DLA-CFA - Newsvendor	479.1 [457.7, 511.8]

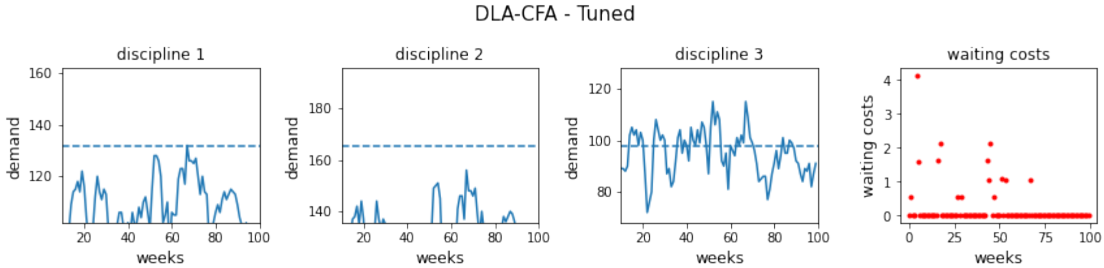
Table 6.13: Average costs for the developed and myopic policies based upon 1000 simulation runs, each of 50 weeks, using capacity settings $\mathbf{r}_1$.

Figure 6.22 shows a simulation run with capacity settings $\mathbf{r}_2$. The capacity of all three disciplines is exceeded, but not for all weeks. The developed policy decreases the peaks. For example, around week 40, a large peak for all three disciplines can be seen at the 'cheapest direct cost' policy. At both DLA-CFA methods, using tuning and newsvendor, this peak is lowered by later admission of outpatients.

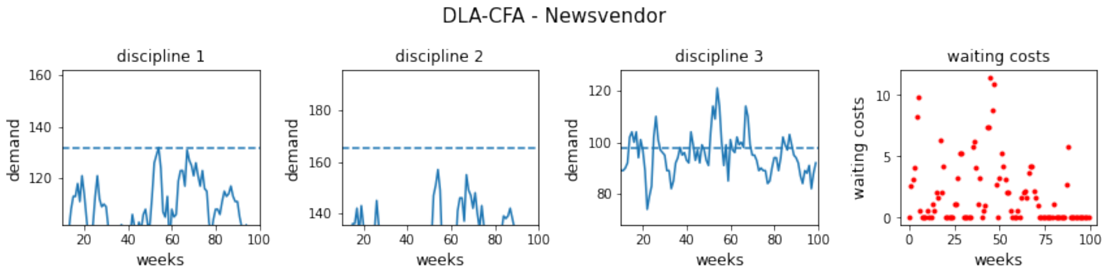
We continue analyzing with the modified capacity $\mathbf{r}_2$. In Table 6.14, we present the results of the assessment factors for both the developed DLA-CFA and myopic policies. We base this assessment on 3000 simulation runs. For each run, we simulated 52 weeks to define the assessment factors on a yearly basis. We note similar results in terms of costs and assessment factors to those in the single-discipline system. The developed DLA-CFA method outperforms the 'no forecast' and 'cheapest direct cost' myopic policies. Again, in terms of access times, the tuned version of the method performs better than the newsvendor version, and in terms of the number of weeks that the capacity is exceeded, the newsvendor model performs better.

Furthermore, in the tuned version of the developed policy, we note that the parameters related to the decision moment have little influence. If we set these values to 0, the resulting costs and results for the assessment factors are almost similar to the initial results. We explain these results by noting that the influence of the decision moments on the demand forecast is only on the last two weeks of the five-week forecast. Therefore, the impact of the new arrivals is more significant. We analyze the system with $T^{\text{MDM}} = 0$ in Section 6.4.5 to identify these possible reasons.

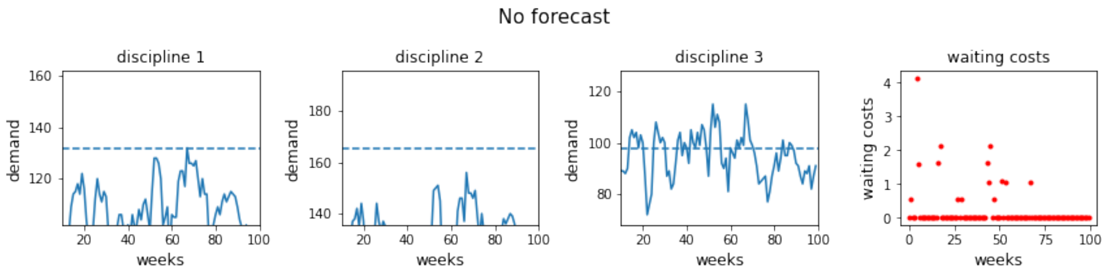
Moreover, we also analyzed the system with a recourse model to accommodate the lack of knowledge of the upcoming arrivals to the waiting list. However, like the single-discipline system, the recourse model yields the same results as when the arrivals are known. The relatively small difference between different arrival sample paths is a reason for this. Although the difference is more prominent than in the single-multidiscipline system with $\lambda_O = 1$ versus $\lambda_O = 4.0$ in this system, this is not significant. We will further analyze this in Section 6.4.6. Another reason that could explain the relatively good performance of the recourse model is the relatively small difference between urgency costs for the different urgent levels, such that the difference between a future arrival of a high-urgency or low-urgency patient in terms of costs is relatively small. We further investigate this in Section 6.4.4.



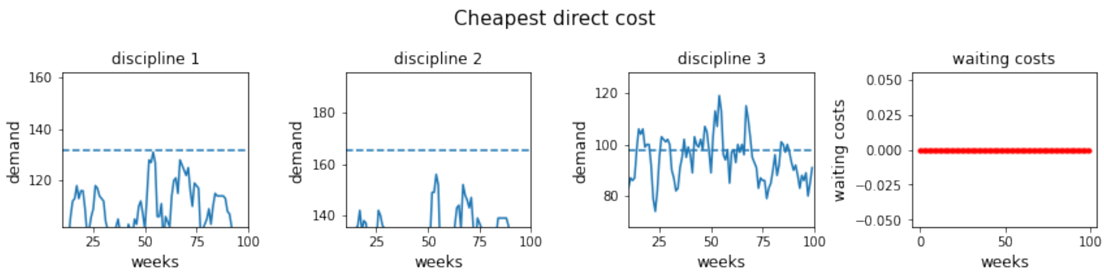
(a) 376.8



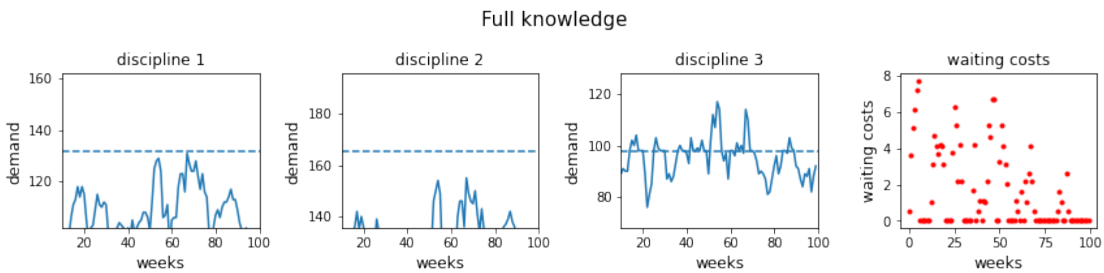
(b) 479.2



(c) 376.8

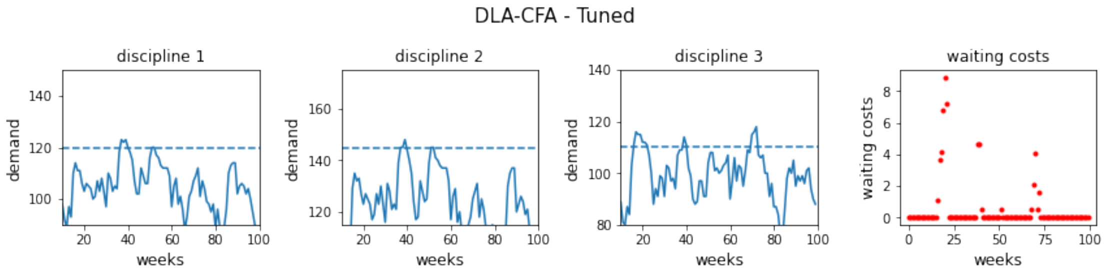


(d) 345.0

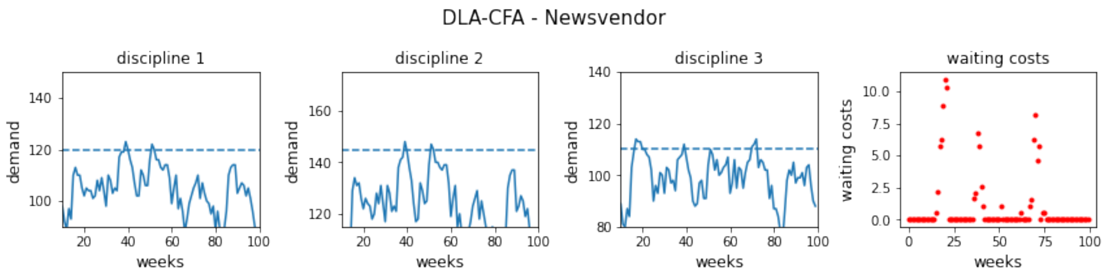


(e) 376.9

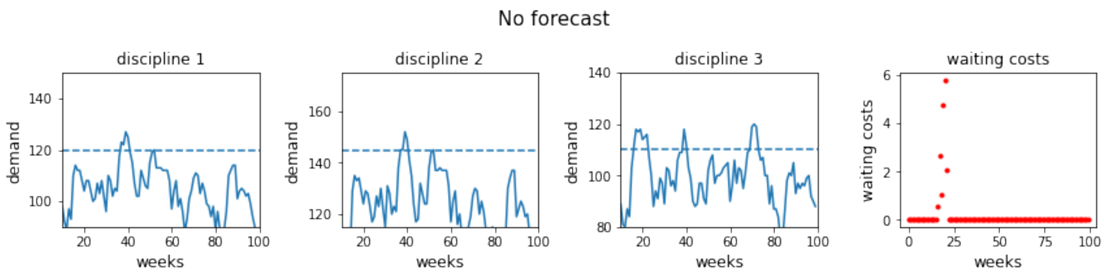
Figure 6.21: A simulation run of 200 weeks with different policies with capacity $\mathbf{r}_1$, the blue line is demand throughout the week, the dotted line is capacity, and the red dots are waiting costs made in a given week. Costs for this simulation run are given below the plot for each corresponding policy.



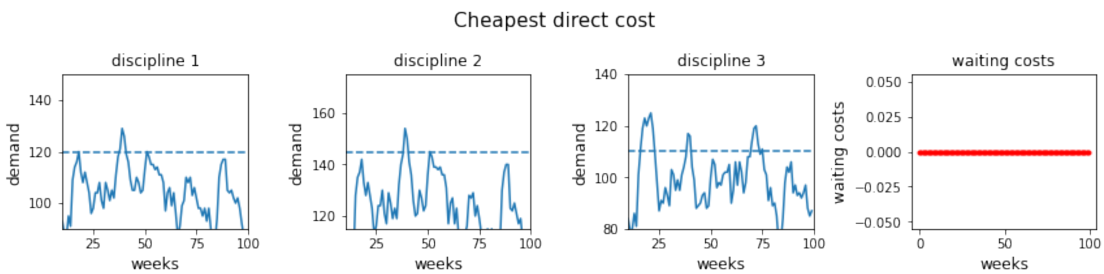
(a) 136.6



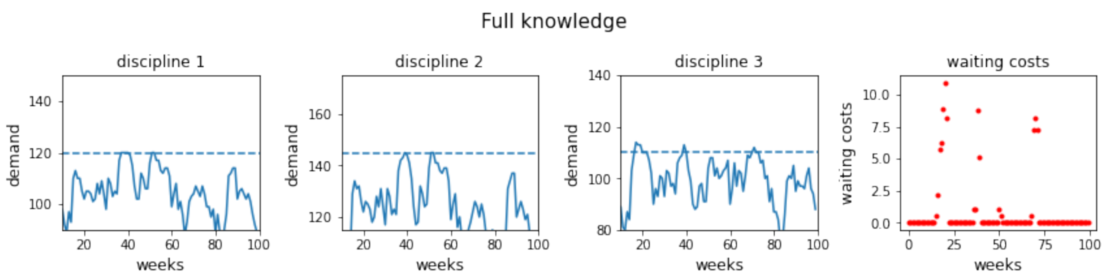
(b) 138.2



(c) 174.4



(d) 210.0



(e) 106.1

Figure 6.22: A simulation run of 200 weeks with different policies with capacity $\mathbf{r}_2$, the blue line is demand throughout the week, the dotted line is capacity, and the red dots are waiting costs made in a given week. Costs for this simulation run are given below the plot for each corresponding policy.

Table 6.14: Performance of DLA-CFA policy for both a tuned as a newsvendor version, and 'full knowledge', 'no forecast' and 'cheapest direct cost' policies based on 3000 simulations of 52 weeks with knowledge of upcoming arrivals. (1) average cost, (2) median cost, (3) average number of weeks waiting per urgency level, (4) average maximum number of weeks waiting per urgency level (5) average number of weeks capacity exceeded per year per discipline, (6) average capacity exceeded per discipline, (7) average maximum peak overcapacity per discipline. Each includes a 95% confidence interval.

	Cost assessment		*Access time assessment*	
	(1) mean	(2) median	(3) mean	(4) maximum
Full knowledge	143.8 [135.3, 152.4]	86.5 [76.1, 92.0]	0.42 [0.33, 0.52] 0.30 [0.21, 0.41] 0.17 [0.12, 0.24]	5.2 [5.0, 5.4] 2.7 [2.5, 2.8] 3.7 [3.6, 3.9]
No forecast	215.9 [205.2, 226.7]	148.3 [136.9, 158.0]	0.10 [0.07, 0.13] 0.09 [0.05, 0.12] 0.04 [0.02, 0.06]	1.9 [1.8, 2.0] 0.8 [0.7, 0.9] 1.2, [1.1, 1.3]
Cheapest direct cost	242.2 [229.8, 254.7]	153.0 [144.0, 169.5]	0 0 0	0 0 0
DLA-CFA - Tuned	185.5 [176.1, 195.2]	122.1 [114.5, 133.4]	0.29 [0.23, 0.36] 0.22 [0.16, 0.28] 0.13 [0.09, 0.18]	3.8 [3.7, 3.9] 1.9 [1.8, 2.0] 2.8 [2.7, 2.9]
DLA-CFA - Newsvendor	189.8 [175.3, 204.4]	132.8 [118.1, 144.4]	0.43 [0.37, 0.50] 0.30 [0.24, 0.38] 0.18 [0.14, 0.23]	4.9 [4.3, 5.5] 2.5 [2.0, 3.0] 3.4 [2.9, 4.0]

	Demand assessment		
	(5) # weeks	(6) mean	(7) maximum
Full knowledge	1.8 [1.6, 1.9] 0.6 [0.5, 0.6] 4.7 [4.5, 4.9]	2.2 [2.1, 2.3] 3.0 [2.9, 3.2] 4.5 [4.4, 4.6]	1.6 [1.5, 1.7] 1.1 [1.0, 1.2] 5.7 [5.5, 5.9]
No forecast	7.7 [7.4, 8.0] 5.8 [5.5, 5.9] 9.3 [9.0, 9.2]	5.3 [5.2, 5.4] 5.2 [5.2, 5.3] 6.5 [6.4, 6.5]	7.3 [7.0, 7.5] 6.7 [6.5, 6.9] 10.4 [10.1, 10.7]
Cheapest direct cost	7.5 [7.2, 7.9] 5.8 [5.6, 6.1] 9.0 [8.7, 9.3]	6.7 [6.6, 6.8] 6.8 [6.7, 6.9] 7.9 [7.8, 8.0]	8.6 [8.3, 8.9] 4.7 [4.5, 4.8] 12.1 [11.8, 11.4]
DLA-CFA - Tuned	5.6 [5.3, 5.8] 3.8 [3.6, 4.0] 7.8 [7.5, 8.1]	3.8 [3.7, 3.9] 3.9 [3.8, 4.0] 5.5 [5.4, 5.6]	5.2 [5.0, 5.4] 4.7 [4.5, 4.8] 8.7 [8.5, 9.0]
DLA-CFA - Newsvendor	4.4 [4.1, 4.8] 3.1 [2.9, 3.3] 6.9 [6.5, 7.3]	3.4 [3.3, 3.5] 3.4 [3.3, 3.5] 5.0 [4.9, 5.1]	4.3 [4.1, 4.6] 3.8 [3.5, 4.0] 8.1 [7.8, 8.5]

6.4.4 Influence of waiting and capacity costs

The sections above show that the developed policy applies to systems on the scale of a rehabilitation clinic. This section evaluates the impact of different waiting and capacity costs on the resulting assessment factors to determine the optimal cost settings that best fit the rehabilitation clinic's wishes.

We define three different cost settings. Figure 6.23 shows the tuning process for cost settings κ_{c_2}. Note that this tuning obtains different results than κ_{c_1} tuning, demonstrating that the tuning reacts to the system settings, as in Figures 6.18, 6.19 and 6.20. This shows the need for tuning for every system parameter adaptation.

$$\kappa_{c_1} = \{1.5, 1.5, 1.5\}, \kappa_{c_2} = \{0.5, 0.5, 0.5\}, \kappa_{c_3} = \{2, 2, 3\}, \kappa_{c_4} = \{10, 5, 15\}$$

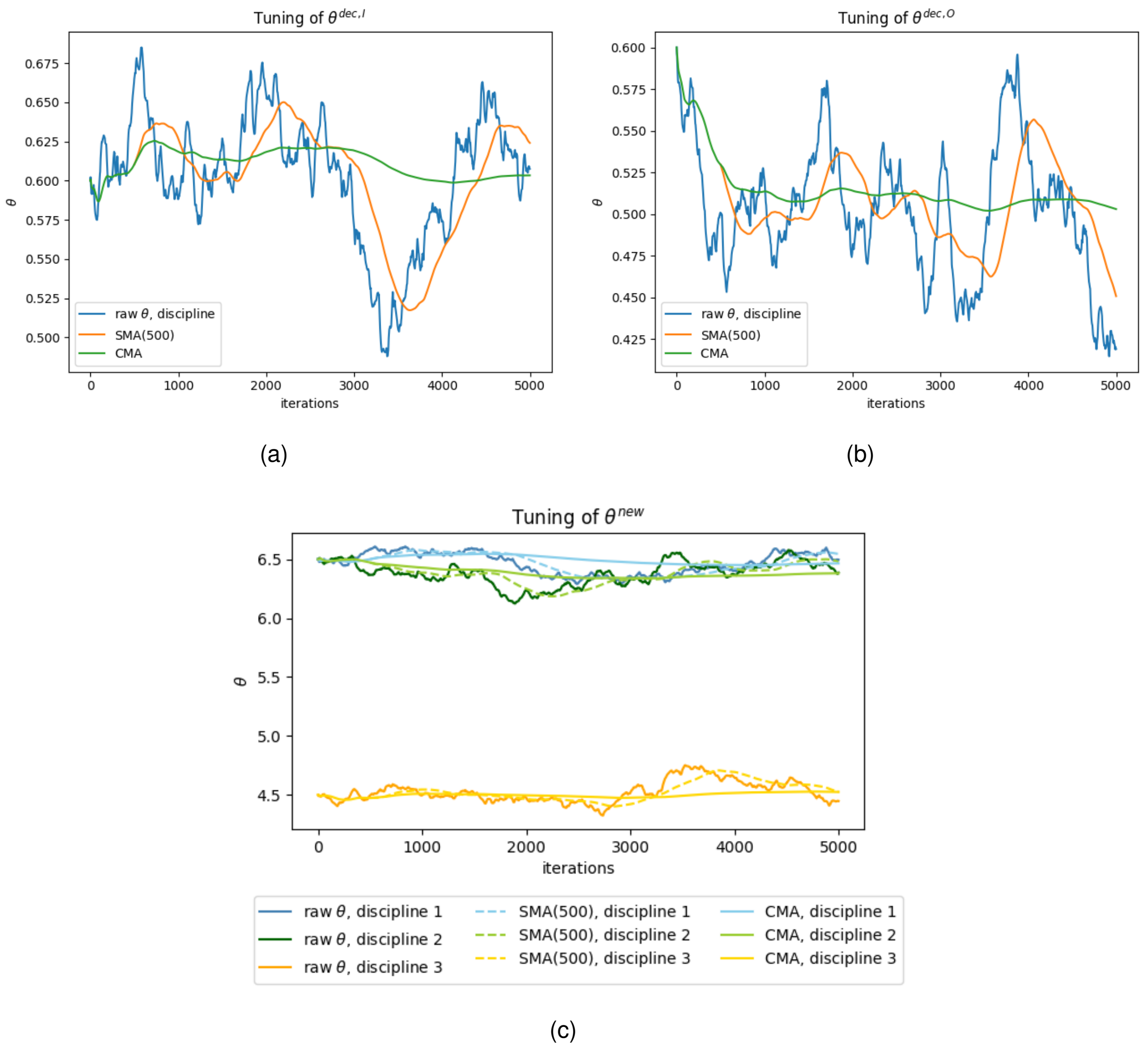
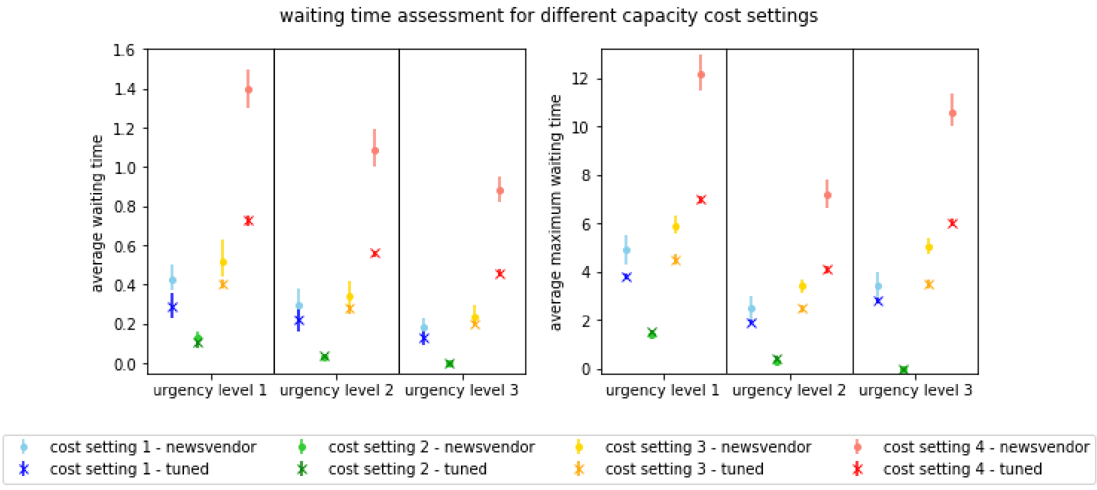
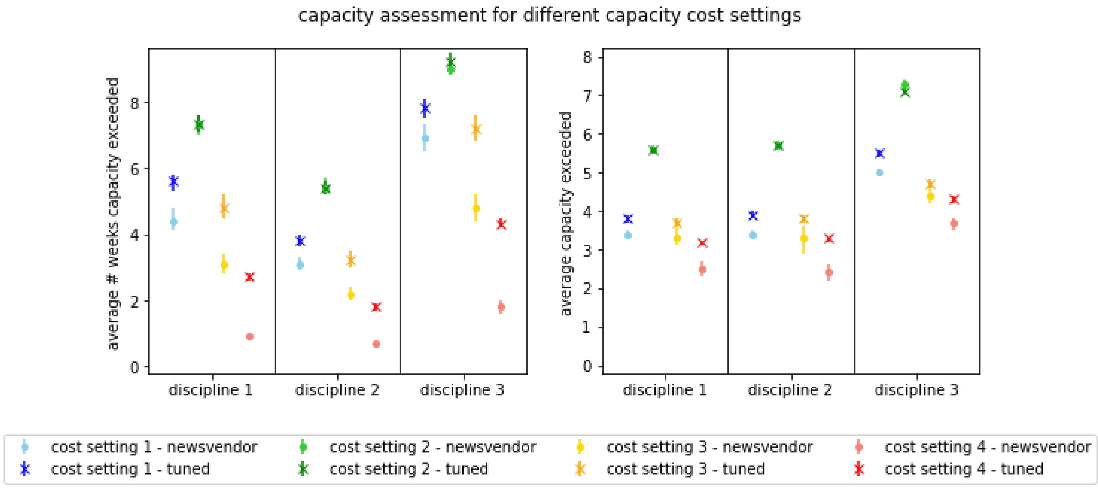


Figure 6.23: Tuning of θ^{new}, $\theta^{dec,I}$ and $\theta^{dec,O}$ for a system with settings as in Table 6.12 with with cost settings κ_{c_2}.

Figure 6.24 shows the results of the assessment factors for the various capacity cost settings for both the newsvendor version and the tuned version of the developed policy. Logically, lower capacity costs lead to shorter access times and more frequent overcapacity. And the highest capacity costs settings κ_{c_4} result in the highest access times with the lowest capacity exceeded, on average. The average access time is about nine days, closer to rehabilitation clinics' current average access time. But notably, the relative increase in access time is larger than the relative decrease in capacity exceeded in comparison with κ_{c_3}.



(a)



(b)

Figure 6.24: Assessment factor related to access times (a) and capacity (b) for costs settings $\kappa_{c_1}, \kappa_{c_2}, \kappa_{c_3}, \kappa_{c_4}$.

We also note that the difference between the assessment factors' results for the tuned version and the newsvendor version of the model becomes larger if capacity costs increase relatively. Moreover, we note that for these cases, the relative performance between the developed DLA-CFA policy increases, which is visualized in Figure 6.25. This means that the developed policy performs relatively better if overuse of capacity gets more priority relative to access time.

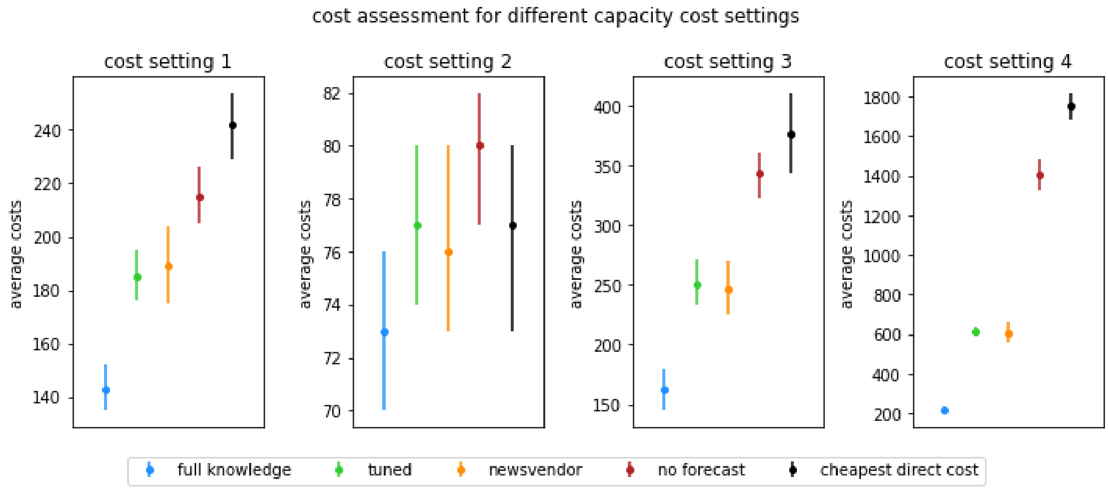
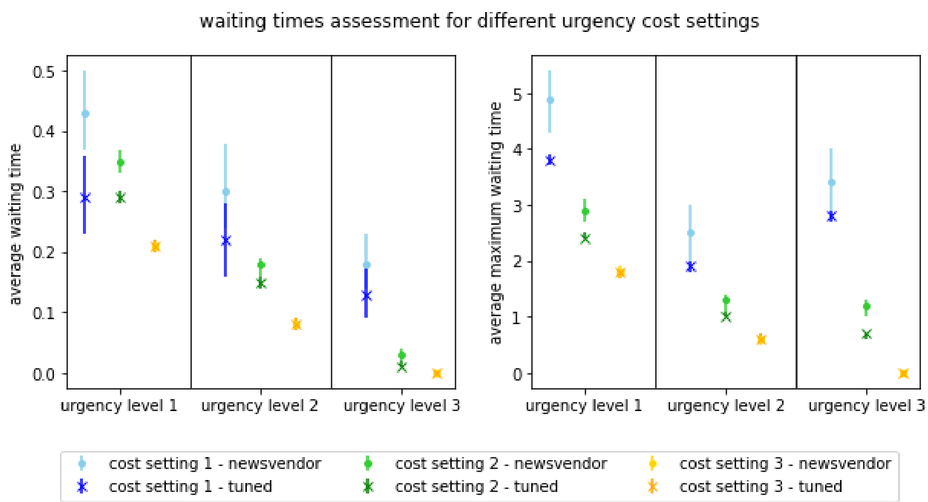


Figure 6.25: Average costs for the developed and myopic policies using different cost settings.

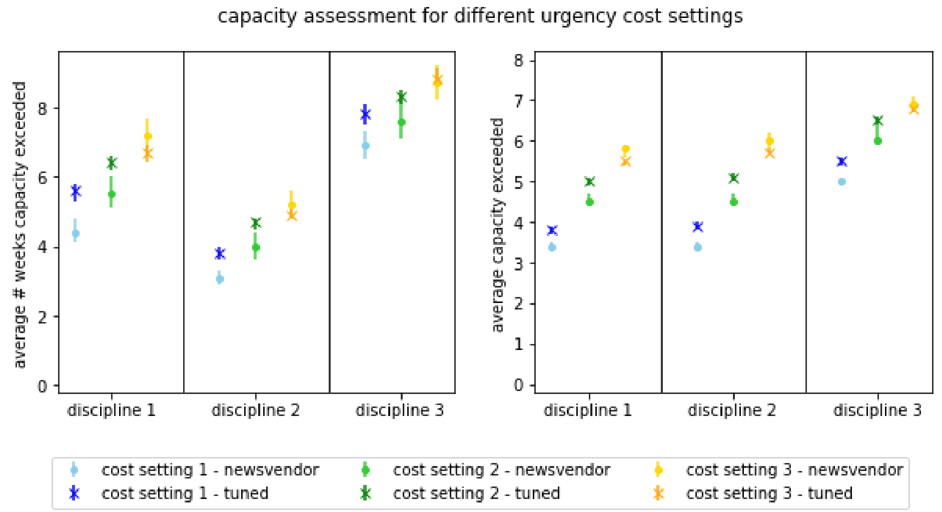
Similarly, we define three different urgency settings. Figure 6.26 visualizes the three settings' access time results and capacity assessment factors, where we use κ_{c1} as capacity cost settings.

$$\kappa_{u_1} = \{0.55, 1.05, 1.55\}, \kappa_{u_2} = \{0.15, 0.55, 4.55\}, \ \kappa_{u_3} = \{1.05, 2.05, 10.05\}$$

We note that with higher urgency costs, the access time decreases. We also note that the higher the urgency costs, the closer the results of the tuned and newsvendor version of the model. This is also what we expected. Furthermore, we evaluated whether variations in urgency costs lead to differences in the performance of the recourse model. Our observations suggest that this is not a significant difference.



(a)



(b)

Figure 6.26: Assessment factor related to access times (a) and capacity (b) for costs settings $\kappa_{u1}, \kappa_{u2}, \kappa_{u3}$.

6.4.5 Influence of the timing of the decision moment

In the above results, we noted that the influence of the forecast demand for the result of decision moments is minor. The results with only applying θ^{new} are almost similar to the results shown in Table 6.14 with using $\theta^{\text{dec},I}$ and $\theta^{\text{dec},O}$. We are interested in whether the influence of the decision moments changes if the time between the decision moment and the resulting extra demand becomes shorter. Otherwise, we could conclude that for all system settings, we can disregard the demand forecast for decision moments.

To test this, we use a similar system as in Table 6.12, but now with $T^{\text{MDM}} = 0$ for both inpatients and outpatients. We also apply a cost setting $\kappa_c = \{10, 10, 10\}$, as from the previous section, we note that the relative performance of the policy increases if the capacity costs are higher. We think that the decision moment is also more influential.

Table 6.15 shows the results of this system for a ‘full knowledge’ policy and a DLA-CFA policy using all three parameters and using only θ^{new}. The first DLA-CFA policy, which includes all parameters, performs much better cost-wise than without the parameters regarding the decision moments. This shows that one cannot always disregard the influence of the decision moments.

6.4.6 Influence on the recourse model

We implemented a recourse model to eliminate the assumption of knowledge about future arrivals on the waiting list. We found that employing the recourse model for the system resembling the rehabilitation clinic yielded results nearly identical to those obtained with prior arrival knowledge, similar to what we observed in the single-discipline system. A reason for this result is the relatively low number of possible arrivals in the system analyzed. To see if this can be explained as such, we perform a test for a larger system, where $\lambda_O = 15$. To eval-

Table 6.15: Performance of tuned DLA-CFA policy with and without decision moments, and the 'full knowledge' policy based on 3000 simulations of 52 weeks with knowledge of upcoming arrivals. (1) average cost, (2) median cost, (3) average number of weeks waiting per urgency level, (4) average maximum number of weeks waiting per urgency level (5) average number of weeks capacity exceeded per year per discipline, (6) average capacity exceeded per discipline, (7) average maximum peak overcapacity per discipline. Each includes a 95% confidence interval.

	Cost assessment		*Access time assessment*	
	(1) mean	(2) median	(3) mean	(4) maximum
Full knowledge	93.4 [80.9, 106.7]	35.4 [27.5, 39.8]	0.26 [0.21, 0.32] 0.17 [0.13, 0.23] 0.13 [0.09, 0.17]	4.2 [3.9, 4.5] 2.1 [1.9, 2.3] 3.0 [2.8, 3.3]
DLA-CFA - Tuned	352.3 [318.0, 388.2]	199.4 [178.9, 241.0]	0.20 [0.16, 0.25] 0.13 [0.09, 0.17] 0.11 [0.08, 0.14]	2.9 [2.8, 3.2] 1.5 [1.3, 1.6] 2.2 [2.0, 2.4]
DLA-CFA - Tuned with only θ^{new}	702.5 [630.2, 780.1]	370.0 [300.0, 430.0]	0.01 [0.0, 0.01] 0.01 [0.0, 0.01] 0.0 [0.0, 0.01]	0.3 [0.2, 0.3] 0.1 [0.0, 0.1] 0.2 [0.1, 0.2]

	Demand assessment		
	(5) # weeks	(6) mean	(7) maximum
Full knowledge	0.02 [0.01, 0.04] 0.02 [0.01, 0.03] 0.5 [0.4, 0.6]	3.0 [2.4, 3.9] 3.0 [2.3, 3.8] 3.4 [3.1, 3.7]	0.08 [0.03, 0.2] 0.07 [0.03, 0.1] 1.0 [0.9, 1.2]
DLA-CFA - Tuned	1.5 [1.3, 1.7] 1.2 [1.1, 1.3] 4.6 [4.3, 5.0]	3.1 [2.9, 3.1] 3.0 [2.8, 3.1] 4.7 [4.6, 4.8]	2.1 [1.9, 2.3] 1.9 [1.7, 2.1] 6.1 [5.7, 6.5]
DLA-CFA - Tuned with only θ^{new}	3.2 [2.9, 3.6] 2.7 [2.4, 3.0] 6.1 [5.6, 6.5]	5.0 [4.8, 5.2] 5.0 [4.8, 5.2] 6.7 [6.5, 6.9]	4.3 [3.9, 4.6] 4.1 [3.7, 4.4] 8.5 [8.0, 9.1]

uate this, we will assume complete awareness of the other exogenous information that arrives analogous to the previously employed 'full knowledge' policy. In response to the growing number of outpatients, we have adjusted our capacity which is set to $\mathbf{r} = \{180, 200, 160\}$.

We conduct an analysis of the recourse model, as outlined in Algorithm 1, with 1, 10, 25, 50, and 100 iterations at every decision epoch. To assess performance effectively, we perform 1000 simulation runs. The resulting costs are compared with the costs that would have been incurred had the policy possessed complete foresight regarding the forthcoming arrivals to the waiting list. Table 6.16 presents the average differences in costs across all simulation runs, as well as the average differences for those simulation runs where this difference is non-zero.

In our analysis of the recourse model applied to this system, we observed that its performance is worse than that of future arrivals on the waiting list, which is known in advance. This result aligns with our expectations, as having prior knowledge enhances the use of the DLA-CLA LP as given in (5.17). Additionally, we noted a slight improvement in performance with an increased number of iterations used within the recourse model between 25 and 50 iterations. However, it remains relatively modest, indicating that while more iterations can be beneficial, the degree of such benefits may be limited. When examining the performance of the 'full knowledge' policy with the knowledge of arrivals to the waiting list using the 'cheapest direct cost' policy, we observe an average cost of 1005.8, with a confidence interval ranging from 927.2 to 1087.6. It is evident that the difference attributable to the recourse model is relatively negligible when compared to the difference with the 'cheapest direct cost'.

Table 6.16: Compared with the knowledge of the arrivals to the waiting list, average difference costs and average difference costs with excluding same results applying the recourse model for 1, 10, 25, 5,0 and 100 iterations with the 'full knowledge' policy. Including a 95% confidence interval.

	Average difference	Average difference excluding same results
1 iteration	4.0 [3.3, 4.8]	5.1 [4.2, 6.1]
10 iterations	3.9 [3.2, 4.6]	5.2 [4.3, 6.1]
25 iterations	3.9 [3.2, 4.6]	5.1 [4.2, 6.0]
50 iterations	3.0 [2.4, 3.7]	4.0 [3.1, 4.8]
100 iterations	3.3 [2.7, 3.9]	4.5 [3.6, 5.4]

6.4.7 A policy description

The main goal of this project is to develop a policy for outpatient admissions at the multidisciplinary rehabilitation clinic. Given the extensive number of states involved in the developed MDP, it is impractical to provide a detailed description of the recommended actions for each individual state. In this section, we will present an illustrative example of a specific state, outlining the advised action and comparing it to the recommendations made by the myopic 'no forecast' policy. Furthermore, we discuss with the planner to assess their alignment with the recommended outcomes.

In states when the number of admitted patients is low, the recommended admission policy is to admit all outpatients on the waiting list. For instance, in the extreme scenario where no patients are currently admitted, and considering that the arrival rate of inpatients suggests a very low probability of the inpatient ward reaching total capacity in the forthcoming week, this advice is provided. From a logical perspective, the planner concurs with this assertion.

We introduce a more interesting case in which the best admission choice is not obvious. The current demand of this state is given in Figure 6.27. Nine patients are on the waiting list, as shown in Figure 6.28 using fictitious patients' names. In the MDP, we denote this waiting list as follows

$$H_t = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 0 \\ 3 & 1 & 1 \end{bmatrix}.$$

The current demand is insufficient to determine which patients are best to admit. Therefore, we look into the future. Figure 6.29 shows the' no forecast' myopic policy vision. The demand decreases for the next week as patients leave the clinic, and the policy does not consider new admissions. This myopic policy advises admitting all patients, as it does not foresee the possible increases in demand for the coming weeks.

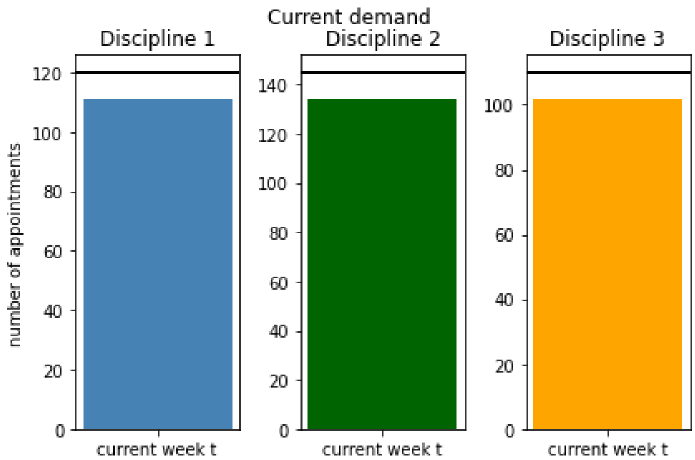


Figure 6.27: Visualization of the demand in the current week of the given state for three disciplines, being (1) occupational therapy, (2) physiotherapy, and (3) speech therapy. The black line indicates the capacity for each discipline.

	Treatment plan		
	Occupational therapy	Physio-therapy	Speech therapy
Urgency level 3			
Ms. Hesen	1	1	1
Mr. van der Poel	1	1	2
Ms. de Jong	1	1	1
Ms. Dobbelsteen	1	1	1
Mr. van Vleuten	2	2	0
Urgency level 2			
Mr. Van Aert	1	1	1
Mr. Vollering	1	1	1
Urgency level 1			
Ms. de Kroon	1	1	1
Mr. Hilhorst	2	2	0

Figure 6.28: Waiting list for the given state, described by random patient names.

The demand forecast of the DLA-CFA policy, using the tuned version of the developed method, is given in Figure 6.30. This forecasts a higher demand for the next week and future weeks within the forecast time interval. Whereas the myopic policy advises admitting all patients from the waiting list, the DLA-CFA policy only advises admitting two out of seven patients, as is visualized by Figure 6.31. The mathematical notation of advised action is described as follows:

$$\mathbf{a}^*(S_t) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 3 & 0 & 1 \end{bmatrix}.$$

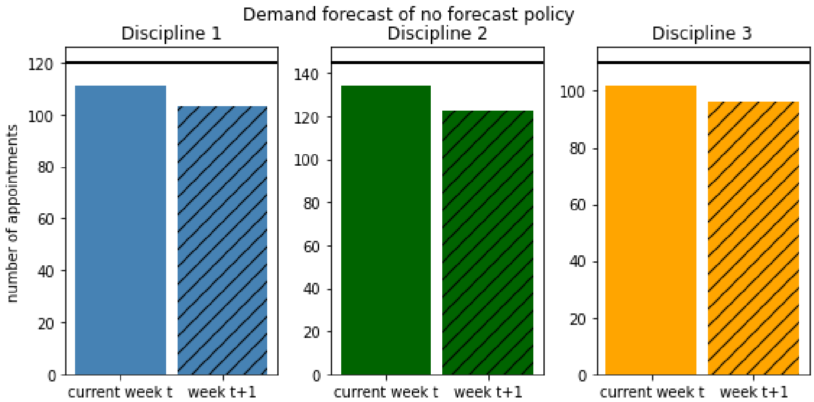


Figure 6.29: Demand forecast of the myopic 'no forecast' policy, which looks one week in advance and does not consider new arrivals of inpatients. The black line indicates the capacity set for each discipline.

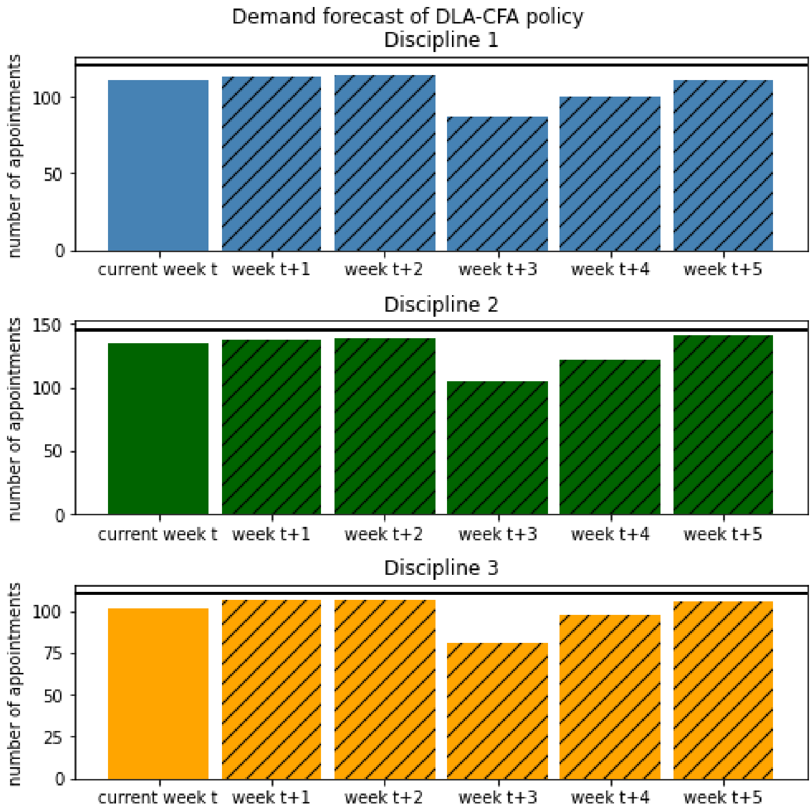


Figure 6.30: Demand forecast of the DLA-CFA policy, newsvendor version. The black line indicates the capacity set for each discipline, and the shaded demand is a forecast.

We explained the main idea of the developed policy to the main planner of the neurological/surgical department of the rehabilitation clinic *Merem*. We presented him with the above and two other examples and asked what his choice of admission would be and whether this coincides with the advice provided by the developed policy.

Firstly, he observed that the developed policy articulates the intuitive judgment he uses in making decisions regarding the admission of outpatients. He usually takes into account the possible arrivals of two to three inpatients but does not consider the results of decision

	Treatment plan		
	Occupational therapy	Physio-therapy	Speech therapy
Urgency level 3			
Ms. Hesen	1	1	1
Ms. de Jong	1	1	1
Ms. Dobbelsteen	1	1	1
Mr. van Vleuten	2	2	0
Urgency level 2			
Mr. Van Aert	1	1	1
Urgency level 1			

Figure 6.31: Advised patients to admit for the given state, described by random patient names.

moments as he does not have a clear overview of this, and the decision moments are not always performed as assumed in the model. He describes he does not have an overview of the current and future demand clearly such as in Figure 6.30. He needs to derive this from the actual patient planning, such as in Figures 2.2 and 2.3, describing the planning by a schedule for each patient individually.

Regarding the example given above, he says that he initially might include admitting Mr. van der Poel. He admits he would choose that option since he lacks an overview of current demand, as shown in Figure 6.30, and instead relies on his intuition. Additionally, he acknowledges that speech therapists often facilitate additional appointments by reducing administrative time, on which he bases this decision. Nonetheless, should a forecast be developed in accordance with the DLA-CFA policy, allowing for a more thorough understanding of demand, and should the speech therapists implement stricter criteria, we would align with the suggested admissions strategy. He presents a similar rationale for the additional examples we provided to him.

Given the explanation, he highlights that the implemented policy has enhanced his understanding of his impact on the planning process. He notes that a significant portion of his daily responsibilities is managing appointment schedules to ensure alignment. Consequently, this has resulted in outpatient admission being a secondary task. This project has underscored the critical importance of timing in these admissions and its broader implications.

Nevertheless, he observes that, in practice, admission may face more significant challenges due to the introduction of various constraints applicable to specific patients. For instance, some patients may require treatment to commence within a designated timeframe or may only be eligible to particular practitioners who may already be fully booked for the upcoming weeks. He emphasizes that the developed method does not consider these implications, a viewpoint with which we concur.

He concludes that implementing the developed policy might be difficult due to these implications and considering the current scheduling application, which has existed for nearly twenty years. However, considering the implementation of a new planning application in the near future, he may request the inclusion of an overview of demand such as in Figure 6.27 or, the other way around, available appointment slots. Moreover, a forecast of new arrivals, as outlined in the developed policy as in Figure 6.30, could also be integrated.

Chapter 7

Discussion

In this thesis, we exploited a DLA-CFA policy to decide on outpatient admissions in a rehabilitation clinic. In this chapter, we present a comprehensive overview of our findings and their implications for practice and future research. This chapter consists of three sections. First, we discuss the results obtained in Section 7.1, describing the impact of tuning, the newsvendor model, and the recourse model. Second, we discuss implementing the method in the rehabilitation clinic and what must be done to implement the model there in Section 7.2. Lastly, we recommend future research based on the discussion points in Section 7.3.

7.1 Discussion of the results

In this research, we model the arrival and length of stay of inpatients and outpatients by a Markov Decision Process (MDP). For each decision epoch, the planner needs to decide which outpatients to admit from the waiting list. To solve this MDP, we used a combination of a Cost Function Approximation and a Direct Lookahead Policy (DLA-CFA). Section 5.1 explains the four policy classes as defined by Powell. We conclude that this combined policy class would be the most fitting. DLA-CFA policies are intuitive but have a downside: they involve finding parameter values, which cannot always be performed easily. This is highlighted in Powell's statement, 'Tuning is an art rather than a science.' [32].

This search evolves in forecasting the demand resulting from the exogenous arriving information, consisting of the demand from new arriving inpatients and the extra demand resulting from the extension of the length of the treatment after a decision moment. Due to the policy design, we do not necessarily need the closest-to-reality estimation but rather a forecast that leads to the best choice for the model's objective function. This can be observed during the tuning process of the multi-disciplinary system across two distinct capacity settings, where despite utilizing the same parameters for the exogenous information, the tuning outcomes varied significantly.

We want to highlight that, based on the single-discipline systems and the multi-disciplinary system tested, the impact of the parameter regarding new inpatient arrivals is more significant than that of the decision moments. This is primarily due to the fact that the decision moments only affect the estimated demand after the initial dismissal date has elapsed. This is a fundamental design choice, which makes the total influence of an incorrect parameter estimation smaller. By adapting the system setting, this period can be increased or decreased. We analyzed a system in which the decision moments directly influenced the week after, and we concluded that the decision moment has a more significant influence in this system. However, the developed policy's overall performance has worsened in comparison with the 'full knowledge' policy, which has correct knowledge of the evolution of the demand for the forecast time interval.

To evaluate the value of the parameters for the exogenous information, we used both stochastic approximation and a newsvendor model. Stochastic approximation is tuning the values, which can be done separately for each model parameter. We show that an Adam algorithm works. Due to the high stochasticity of the model, not necessarily one value is perfect, but rather, the tuning will end up fluctuating within a range. It can be difficult to evaluate if the value converged enough. The resulting parameter value, therefore, asks for a certain amount of guessing. The tuning can take several days, depending on the batch size and number of iterations. However, decisions can be made quickly once the parameter value is determined.

The newsvendor model uses knowledge of the transition function to forecast the extra demand. The main advantage is that this makes the chosen value state-dependent. Therefore, it can better react to different scenarios, for example, if the bed capacity is reached and no new arrivals can occur. The newsvendor model, on average, outperforms the tuned version for single-discipline systems, but not significantly. We have made a contribution to the literature by introducing the newsvendor model for parameter determination within a DLA-CFA policy, or otherwise described in a multi-period "predict-then-optimize" framework. We see great potential in this application, as it directly determines the forecast without the need for timely tuning. Additionally, the newsvendor model can respond to various distribution functions for different scenarios, as we illustrate in the next paragraph, and can directly address changes in the system setting. Therefore, we express a favor towards the newsvendor model in comparison with the tuning method.

The tuned version of the model has lower access times, and the newsvendor version values overuse of capacity better. A possible reason for this is limiting bed capacity. The newsvendor model can better assess the possible number of arrivals in the forecast due to state dependency. The tuned version has a consistently lower demand forecast for new arrivals to accommodate the limited number of inpatients that can be admitted. However, this can forecast a low demand, increasing the number of admissions and resulting in higher overuse of capacity. We analyzed a tuning parameter that considers whether the clinic is full. However, we obtained similar results as the initially set tuning parameters. To conduct a more thorough investigation, it is advisable to evaluate the adjustment of the inpatient arrival parameter corresponding to the current number of admitted patients. However, an increase in tuning parameters may not be desirable, as indicated by Powell [32].

We showed that the developed method can be used for larger, multi-disciplinary systems resembling the size of a rehabilitation clinic, given that the system's capacity settings are in balance with the asked demand. However, we note that the computational time of utilizing the newsvendor model increases as the distribution function of the exogenous arrivals becomes more complicated. The tuning process is analogous to the single-disciplinary system, requiring approximately two days to determine the parameter values.

By modifying the capacity and waiting cost parameters, we observed that the model balances these two factors differently. It is up to management to determine the ideal balance. We analyzed the results for a couple of different settings. We saw that with high waiting costs, the access time decreased considerably. Similarly, with high capacity costs, the overuse of capacity became less. However, the increase in access time was also considerable. This shows that striving for a fully optimal assessment for access time or overuse of capacity comes at a price. Moreover, we noted that the difference between the tuned version and the newsvendor version of the model becomes more prominent if the capacity costs increase and the access time increases. This is because the newsvendor's forecast will be better tuned to the state, considering the probability of how many patients can arrive. The higher costs result in a higher forecast demand in the newsvendor model and, therefore, higher access times.

The developed DLA-CFA policy is built upon the assumption that we possess knowledge of future arrivals. Acknowledging that this assumption does not reflect reality, we have introduced a recourse model to address the discrepancies. We observed that, upon evaluating the impact of this recourse model in the context of single-disciplinary systems and the system analogous to a rehabilitation clinic, performance remains consistent compared to a model that relies on complete and accurate knowledge of future arrivals. This was even the case if the recourse model only consisted of one iteration. A reason for this is the relatively small range of arriving patients. To test this theory, we analyzed a system with an arrival rate of 15 outpatients per week and a capacity set accordingly. In this system, we observe a notable distinction in performance when utilizing complete information regarding impending outpatient arrivals, in contrast to the recourse model. However, the differences are relatively minor compared to the disparity between the obtained results and those derived from a myopic policy.

7.2 Implementation in the rehabilitation clinic

A direct implementation of the developed method will be challenging due to management and ICT issues, but the method can still provide insights into various aspects. A lack of consistent planning makes a direct comparison between the developed policy and reality difficult. However, we note that the system gives a certain amount of freedom to solve the planning with a pen-and-paper solution. The inconsistency in planning arises from numerous mutations, for example, resulting from changes in treatment plans for patients and changes in capacity because of holidays. These make the scheduling challenging. In this project, we assume fixed treatment plans, now defined by data, but which can be assembled in conversation with doctors and practitioners. The current system is challenging to navigate not only because of mutations but also because of the lack of visibility regarding upcoming capacity. Furthermore, late communication about inpatient admissions and how mutations and admissions are processed often arrives through email, phone, or paper and is overwritten in the data systems, causing unstable data collection.

Assuming the above complications are solved, we could apply the developed model. Firstly, the clinic must decide on the treatment plans, as the developed model does not consider mutations on a weekly basis, as it does in reality. Then, one must apply the tuning algorithm to find reasonable values for the tuning parameters. One should define a joint distribution function by describing (an estimation) the arriving exogenous information. We applied both in our analysis. Unless significant changes in capacity or the distribution of arriving patients change, this only needs to be done once. Then, the method advises on the best admission strategy every week. Even if the advice is not followed, the system updates according to the chosen actions and then sets a new advice for the next week.

Moreover, the model does not only provide advice on outpatient admission. Given that one follows the admission policy, it can also show the expected access time and the predicted overuse of capacity. Additionally, it gives the clinic insight into the balance between capacity and access times, which has been unclear until now. In our analysis of a system similar to a rehabilitation clinic, we observed that the current average access time shows overcapacity issues occurring for only a few weeks per year. The overuse is relatively low in this case, averaging two overuse appointments. To achieve this, we used relatively high capacity costs, tenfold in comparison with the waiting costs. Admissions occur when there is sufficient free capacity or when the access time becomes excessively high. We show that if we lower the capacity costs, the access time decreases, but the overuse of capacity increases. This increase is relatively smaller compared to the decrease in access times. It is the responsibility of management to decide which settings are most appropriate.

Moreover, the method can provide valuable insights if the clinic needs to adjust its capacity. If a clinic needs to reduce capacity in a discipline, the method can analyze how this decrease affects access times and capacity overuse. By doing this, management can better analyze the impact of specific decisions before implementing them.

The developed methodology incorporates a demand forecast to assess the optimal admission choices. We discussed the results of the model with the department's main planner. He is enthusiastic about the model. For the given examples, he agrees with the model's advice. While he acknowledges the importance of direct implementation, he also recognizes that it is not feasible within the framework of the current planning application. However, he notes that a visualization of the anticipated demand, or conversely, the availability of appointment slots, could significantly enhance the planning process. Furthermore, integrating a demand forecast as used in the method could serve as a valuable enhancement in the ongoing updates of the planning program.

7.3 Recommendations for future research

To use the DLA-CFA policy, we addressed a tuning algorithm and a newsvendor model to determine the parameters' values. Tuning is a timely process but can be done ahead of time. The newsvendor model becomes more computationally heavy the larger the system. Both models can be addressed for further research. Other different algorithms can be researched to tune the parameters. Ideally, they should perform similarly to the Adam algorithm but have less computational time. Additional research can also be done in the design of the DLA-CFA policy. In the current design, we have three parameters. It is interesting for future research to redesign the model so that more or different parameters are introduced. For example, defining a parameter regarding inpatient arrivals that is dependent on the number of inpatients currently admitted. This can give insight into the differences between the tuned and newsvendor models in the developed method.

For the given application of the newsvendor model, we did not note a significant difference in the results for different values of α. A more comprehensive study can be conducted on the newsvendor model defined in this research and the influence of α. Furthermore, already much research has been performed on various applications of the newsvendor model, for example, the review of Qin et al. [36]. Often, these use more complex stochastic search methods to solve the newsvendor model. In this research, we introduce the use of the newsvendor model as a forecast method in a DLA-CFA policy. Further research could be conducted on the integration of more complex newsvendor models within the DLA-CFA policy framework. A primary focus should be optimizing computational efficiency by establishing an entirely data-driven distribution. Additionally, it would be valuable to explore whether a more intricate newsvendor model can maintain performance comparable to the tuning version should a more complex system be developed.

Moreover, in this project, we analyzed two different system settings using tuning and the newsvendor version of the developed model. To see how the assessment factors depend on the changes in these settings, further research could be performed on the performance of the developed method for different system settings, such as the number of weeks between two succeeding MDM meetings or the set weeks of dismissal. An interesting example would be the settings where the recourse model breaks down and significantly decreases performance compared to a policy that knows the arrivals on the waiting list. Another part of this research could be making specific system settings week-dependent. To enhance the accuracy of our capacity planning, it would be beneficial to adopt a week-dependent capacity model that takes into account variables such as holidays and maternity leave. While the current newsvendor approach of the developed method is straightforward to implement, it will be essential to modify the algorithm to better align with the tuning model requirements.

Lastly, another area for future research could explore is the addition of an appointment scheduler. This was the initial description of the model, but it was left out due to complexity. This method would be two-fold. When all outpatient appointments are scheduled for a specific day and time, one first needs to plan the appointments for the selected patients from the waiting list. These appointments can be organized for a specific day and time. Secondly, the model should plan the appointment for inpatients. In this project, the developed method already makes sure that too much capacity is used. In this way, a linear program would probably solve this remaining puzzle. One solution would be to consider various assessment factors, such as evenly spread appointments for inpatients and little waiting between appointments for outpatients.

Chapter 8

Conclusion

This thesis aimed to improve the planning process at a rehabilitation clinic, specifically by addressing the admission policy for outpatients. To reach this goal, we developed a versatile model that balances the access time and the overuse of capacity to define which patients from the waiting list to admit. The model considers random arrivals of inpatients, which a planner cannot influence but use up a part of the capacity. The planner regulates the admission of outpatients. Ideally, this is done such that outpatients exactly fill up the remaining capacity. For the method, we designed a Markov Decision Process (MDP), which models the arrival and length of stay of both inpatients and outpatients. The model accounts for the multi-disciplinary environment exhibited in a rehabilitation clinic.

We use a combination of a Direct Look Ahead and a Cost Function Approximation (DLA-CFA) policy as a solution approach to solve the MDP approximately. The policy forecasts the arriving demand and extra demand resulting from patients staying longer in the clinic in order to decide which patients to admit, such that the access time is balanced with the overuse of capacity. The policy is constructed in two ways. One method involves tuning, where a stochastic approximation algorithm determines the values of the defined parameters. The other method uses a newsvendor model, eliminating the need for lengthy tuning processes.

The developed policy can be applied beyond making weekly decisions about outpatient admissions. It also facilitates the analysis of access times and the overuse of capacity. Overusing capacity results in appointment cancellations or less administration time for practitioners. Furthermore, one can identify the factors that influence this, such as the arrival rate and set capacity. If the right balance between access times and overcapacity is found, the actual planning of appointments should be eased. Next to the planning of a neurological and surgical rehabilitation clinic, the model is also applicable in any multi-disciplinary, multi-appointment setting, in which one should consider unknown exogenous arriving information in the current state to make the optimal decision.

We tested the performance of the developed method through a simulation study. We did this for single-discipline systems, and a system that resembles the rehabilitation clinic. We can conclude that the developed method performs better than the tested myopic policies, which do not use such forecasts. The performance is measured through the costs of the objective functions, balancing the access time and overuse of capacity. By adopting the cost setting, a manager can prioritize decreasing the overuse of capacity over decreased access time or vice versa. We noted that given the current access time of the rehabilitation clinic, the overuse of capacity can be decreased.

In conclusion, we developed a method that supports and optimizes patient admission at multi-appointment and multi-disciplinary rehabilitation clinics, such that the right balance between access times and overuse of capacity is applied. Furthermore, we introduced the use of a newsvendor model as a method to include a multi-period forecast in a DLA-CFA Policy.

Bibliography

[1] P. R. Harper, “A framework for operational modelling of hospital resources,” *Health Care Management Science*, vol. 5, no. 3, pp. 165–173, 2002, ISSN: 13869620. DOI: 10.1023/A:1019767900627/METRICS. [Online]. Available: https://link.springer.com/article/10.1023/A:1019767900627.

[2] J. Bai, A. Fügener, J. Gönsch, J. O. Brunner, and M. Blobner, “Managing admission and discharge processes in intensive care units,” *Health Care Management Science*, vol. 24, no. 4, pp. 666–685, Dec. 2021, ISSN: 15729389. DOI: 10.1007/S10729-021-09560-6/FIGURES/11. [Online]. Available: https://link-springer-com/article/10.1007/s10729-021-09560-6.

[3] A. Braaksma, N. Kortbeek, G. F. Post, and F. Nollet, “Integral multidisciplinary rehabilitation treatment planning,” *Operations Research for Health Care*, vol. 3, pp. 145–159, 2014. DOI: 10.1016/j.orhc.2014.02.001. [Online]. Available: http://dx.doi.org/10.1016/j.orhc.2014.02.001.

[4] Merem, *Merem: over ons*. [Online]. Available: https://www.merem.nl/over-ons.

[5] Nederlandse Zorgautoriteit, *Regeling Wachttijden en wachttijdbemiddeling medisch specialistische zorg*, Oct. 2017. [Online]. Available: https://puc.overheid.nl/nza/doc/PUC_2034_22/1/.

[6] Y. Li, Z. Zhou, N. Chen, L. He, and M. Zhou, “Seasonal variation in the occurrence of ischemic stroke: A meta-analysis,” *Environmental Geochemistry and Health*, vol. 41, no. 5, pp. 2113–2130, Oct. 2019, ISSN: 15732983. DOI: 10.1007/S10653-019-00265-Y/METRICS. [Online]. Available: https://link.springer.com/article/10.1007/s10653-019-00265-y.

[7] P. J. Hulshof, N. Kortbeek, R. J. Boucherie, E. W. Hans, and P. J. Bakker, *Taxonomic classification of planning decisions in health care: a structured review of the state of the art in OR/MS*, Dec. 2012. DOI: 10.1057/hs.2012.18.

[8] A. G. Leeftink, I. A. Bikker, I. M. Vliegen, and R. J. Boucherie, *Multi-disciplinary planning in health care: a review*, Apr. 2020. DOI: 10.1080/20476965.2018.1436909.

[9] P. Kolesab, “A markovian model for hospital admission,” Columbia University, Tech. Rep., Feb. 1970. [Online]. Available: https://www.jstor.org/stable/2628725.

[10] C.-F. Chien, F.-P. Tseng, and C.-H. Chen, “An evolutionary approach to rehabilitation patient scheduling: A case study q,” 2007. DOI: 10.1016/j.ejor.2007.01.062. [Online]. Available: www.elsevier.com/locate/ejor.

[11] S. N. Ogulata, M. Koyuncu, and E. Karakas, “Personnel and Patient Scheduling in the High Demanded Hospital Services: A Case Study in the Physiotherapy Service,” *Journal of Medical Systems*, vol. 32, no. 3, pp. 221–228, Jun. 2008, ISSN: 0148-5598. DOI: 10.1007/s10916-007-9126-4.

[12] J. Zhang, W. Xie, and S. C. Sarin, “Multiproduct Newsvendor Problem with Customer-Driven Demand Substitution: A Stochastic Integer Program Perspective,” *INFORMS Journal on Computing*, vol. 33, no. 3, pp. 1229–1244, Jul. 2021, ISSN: 1091-9856. DOI: 10.1287/ijoc.2020.0996.

[13] B. Cardoen, E. Demeulemeester, and J. Beliën, “Operating room planning and scheduling: A literature review,” *European Journal of Operational Research*, vol. 201, pp. 921–932, 2009. DOI: 10.1016/j.ejor.2009.04.011.

[14] S. Samiedaluie, B. Kucukyazici, V. Verter, and D. Zhang, “Managing patient admissions in a neurology ward,” *Operations Research*, vol. 65, no. 3, pp. 635–656, May 2017, ISSN: 15265463. DOI: 10.1287/opre.2016.1574.

[15] S. C. Kim, I. Horowitz, K. K. Young, and T. A. Buckley, “Analysis of capacity management of the intensive care unit in a hospital,” *European Journal of Operational Research*, vol. 115, no. 1, pp. 36–46, May 1999, ISSN: 0377-2217. DOI: 10.1016/S0377-2217(98)00135-0.

[16] L. Garg, S. McClean, B. Meenan, and P. Millard, “A non-homogeneous discrete time Markov model for admission scheduling and resource planning in a cost or capacity constrained healthcare system,” *Health Care Management Science*, vol. 13, no. 2, pp. 155–169, Jun. 2010, ISSN: 13869620. DOI: 10.1007/S10729-009-9120-0/TABLES/5. [Online]. Available: https://link-springer-com/article/10.1007/s10729-009-9120-0.

[17] W. Shen, L. Luo, L. Luo, L. Zhang, and T. Zhu, “A data-driven newsvendor model for elective-emergency admission control under uncertain inpatient bed capacity,” *Journal of Evidence-Based Medicine*, vol. 17, no. 1, pp. 78–85, Mar. 2024, ISSN: 1756-5383. DOI: 10.1111/jebm.12599.

[18] H. Zamani, F. Parvaresh, N. Izady, and R. Zanjirani Farahani, “Admission, discharge, and transfer control in patient flow logistics: Overview and future research,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 191, p. 103 722, Nov. 2024, ISSN: 13665545. DOI: 10.1016/j.tre.2024.103722.

[19] G. Zhu, D. Lizotte, and J. Hoey, “Scalable approximate policies for Markov decision process models of hospital elective admissions,” *Artificial Intelligence in Medicine*, vol. 61, no. 1, pp. 21–34, May 2014, ISSN: 09333657. DOI: 10.1016/j.artmed.2014.04.001.

[20] J. Marynissen and E. Demeulemeester, “Literature review on multi-appointment scheduling problems in hospitals,” *European Journal of Operational Research*, vol. 272, no. 2, pp. 407–419, Jan. 2019, ISSN: 03772217. DOI: 10.1016/j.ejor.2018.03.001.

[21] K. Schimmelpfeng, S. Helber, and S. Kasper, “Decision support for rehabilitation hospital scheduling,” *OR Spectrum*, vol. 34, no. 2, pp. 461–489, Apr. 2012, ISSN: 01716468. DOI: 10.1007/s00291-011-0273-0.

[22] A. Sauré, J. Patrick, S. Tyldesley, and M. L. Puterman, “Dynamic multi-appointment patient scheduling for radiation therapy,” *European Journal of Operational Research*, vol. 223, no. 2, pp. 573–584, Dec. 2012, ISSN: 0377-2217. DOI: 10.1016/J.EJOR.2012.06.046.

[23] I. M. Raschendorfer and H. W. Hamacher, “Hierarchical Edge Colorings and Rehabilitation Therapy Planning in Germany *,” Tech. Rep., 2014. [Online]. Available: https://d-nb.info/1054328269/34.

[24] S. Kling, S. Kraul, and J. O. Brunner, “Customized GRASP for rehabilitation therapy scheduling with appointment priorities and accounting for therapist satisfaction,” *OR Spectrum*, vol. 46, no. 3, pp. 821–872, Sep. 2024, ISSN: 0171-6468. DOI: 10.1007/s00291-023-00742-y.

[25] A. G. Leeftink, I. M. H. Vliegen, ·. E. W. Hans, and G. L. Nl, “Stochastic integer programming for multi-disciplinary outpatient clinic planning,” *Health Care Manag Sci*, DOI: 10.1007/s10729-017-9422-6. [Online]. Available: https://doi.org/10.1007/s10729-017-9422-6.

[26] Y. Hur, J. F. Bard, and D. J. Morrice, “Appointment scheduling at a multidisciplinary outpatient clinic using stochastic programming,” *Naval Research Logistics (NRL)*, vol. 68, no. 1, pp. 134–155, Feb. 2021, ISSN: 0894-069X. DOI: 10.1002/nav.21895.

[27] M. L. Puterman, *Markov decision processes: Discrete stochastic dynamic programming*. 2008. DOI: 10.1002/9780470316887. [Online]. Available: https://onlinelibrary-wiley-com/doi/book/10.1002/9780470316887.

[28] R. Bellman, *Dynamic Programming*. Princeton University Press, Aug. 2021, ISBN: 9781400835386. DOI: 10.2307/j.ctv1nxcw0f.

[29] W. B. Powell, “Approximate Dynamic Programming: Solving the Curses of Dimensionality: Second Edition,” *Approximate Dynamic Programming: Solving the Curses of Dimensionality: Second Edition*, pp. 1–638, Sep. 2011. DOI: 10.1002/9781118029176. [Online]. Available: https://onlinelibrary.wiley.com/doi/book/10.1002/9781118029176.

[30] W. B. Powell, *Reinforcement Learning and Stochastic Optimization*. 2022, ISBN: 9781119815037. DOI: 10.1002/9781119815068.

[31] W. B. Powell and S. Meisel, “Tutorial on Stochastic Optimization in Energy - Part II: An Energy Storage Illustration,” *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1468–1475, Mar. 2016, ISSN: 08858950. DOI: 10.1109/TPWRS.2015.2424980.

[32] W. B. Powell and S. Ghadimi, “The Parametric Cost Function Approximation: A new approach for multistage stochastic programming,” Jan. 2022. [Online]. Available: http://arxiv.org/abs/2201.00258.

[33] X. Chen, K. Li, S. Lin, and X. Ding, “Technician routing and scheduling with employees’ learning through implicit cross-training strategy,” *International Journal of Production Economics*, vol. 271, p. 109 208, May 2024, ISSN: 0925-5273. DOI: 10.1016/J.IJPE.2024.109208.

[34] J. C. Spall, *Introduction to Stochastic Search and Optimization*. Wiley, Mar. 2003, ISBN: 9780471330523. DOI: 10.1002/0471722138.

[35] K. J. Arrow, T. Harris, and J. Marschak, “Optimal Inventory Policy,” *Econometrica*, vol. 19, no. 3, p. 250, Jul. 1951, ISSN: 00129682. DOI: 10.2307/1906813.

[36] Y. Qin, R. Wang, A. J. Vakharia, Y. Chen, and M. M. Seref, “The newsvendor problem: Review and directions for future research,” *European Journal of Operational Research*, vol. 213, no. 2, pp. 361–374, Sep. 2011, ISSN: 0377-2217. DOI: 10.1016/J.EJOR.2010.11.024.

[37] U. Sadana, A. Chenreddy, E. Delage, A. Forel, E. Frejinger, and T. Vidal, “A survey of contextual optimization methods for decision-making under uncertainty,” *European Journal of Operational Research*, vol. 320, no. 2, pp. 271–289, Jan. 2025, ISSN: 03772217. DOI: 10.1016/j.ejor.2024.03.020.

[38] N. Halman, J. B. Orlin, and D. Simchi-Levi, “Approximating the Nonlinear Newsvendor and Single-Item Stochastic Lot-Sizing Problems When Data Is Given by an Oracle,” *Operations Research*, vol. 60, no. 2, pp. 429–446, Apr. 2012, ISSN: 0030-364X. DOI: 10.1287/opre.1110.1031.

[39] X. Chang, J. Wu, H. Sun, and X. Yan, “A Smart Predict-then-Optimize method for dynamic green bike relocation in the free-floating system,” *Transportation Research Part C: Emerging Technologies*, vol. 153, p. 104 220, Aug. 2023, ISSN: 0968090X. DOI: 10.1016/j.trc.2023.104220.

[40] J. Jiang, D. Sheng, X. Chen, Q. Tian, F. Li, and P. Yang, “Data-driven collaborative healthcare resource allocation in pandemics,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 192, p. 103 828, Dec. 2024, ISSN: 1366-5545. DOI: 10.1016/J.TRE.2024.103828.

[41] K. Amaruchkul, “Planning migrant labor for green sugarcane harvest: A stochastic logistics model with dynamic yield prediction,” *Computers & Industrial Engineering*, vol. 154, p. 107 016, Apr. 2021, ISSN: 0360-8352. DOI: 10.1016/J.CIE.2020.107016.

[42] R. Sutton and A. Barto, "Reinforcement Learning: An Introduction," *IEEE Transactions on Neural Networks*, vol. 9, no. 5, pp. 1054–1054, Sep. 1998, ISSN: 1045-9227. DOI: 10.1109/TNN.1998.712192.

[43] A. B. Philpott and V. L. De Matos, "Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion," *European Journal of Operational Research*, vol. 218, no. 2, pp. 470–483, Apr. 2012, ISSN: 0377-2217. DOI: 10.1016/J.EJOR.2011.10.056.

[44] S. Ghadimi and W. B. Powell, "Stochastic search for a parametric cost function approximation: Energy storage with rolling forecasts," *European Journal of Operational Research*, vol. 312, no. 2, pp. 641–652, Jan. 2024, ISSN: 03772217. DOI: 10.1016/j.ejor.2023.08.003.

[45] V. Krätschmer, "First order asymptotics of the sample average approximation method to solve risk averse stochastic programs," *Mathematical Programming*, vol. 208, no. 1-2, pp. 209–242, Nov. 2024, ISSN: 0025-5610. DOI: 10.1007/s10107-023-02036-1.

[46] X. Peng, L. Li, and F. Y. Wang, "Accelerating Minibatch Stochastic Gradient Descent Using Typicality Sampling," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 11, pp. 4649–4659, Nov. 2020, ISSN: 21622388. DOI: 10.1109/TNNLS.2019.2957003.

[47] G. Hinton, N. Srivastava, and K. Swersky, "Neural Networks for Machine Learning Lecture 6a Overview of mini--batch gradient descent," Tech. Rep. [Online]. Available: https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf.

[48] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," Dec. 2014. [Online]. Available: http://arxiv.org/abs/1412.6980.

[49] S. P. Adam, S. A. N. Alexandropoulos, P. M. Pardalos, and M. N. Vrahatis, "No free lunch theorem: A review," *Springer Optimization and Its Applications*, vol. 145, pp. 57–82, 2019, ISSN: 19316836. DOI: 10.1007/978-3-030-12767. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-12767-1_5.