Reducing waiting times at out-of-hours general practitioner departments

A data-driven simulation modelling and optimization study

Hanneke Post
November 2021
MSc. Engineering & Policy Analysis

REDUCING WAITING TIMES AT OUT-OF-HOURS GENERAL PRACTITIONER DEPARTMENTS

A thesis submitted to the Delft University of Technology in partial fulfillment of the requirements for the degree of

Master of Science in Engineering & Policy Analysis

by
Hanneke Post
Student number: 4374665
to be defended publicly on November 17, 2021 at 2:00 PM

Chair & First supervisor:	Prof. dr. Alexander Verbraeck
Second supervisor:	Dr. Yilin Huang
Advisor	Dr. Saba Hinrichs-Krapels
External supervisor	Dr. Thijs de Bruijn

An electronic version of this thesis is available at
http://repository.tudelft.nl/

TUDelft

EXECUTIVE SUMMARY

The pressure on healthcare systems is increasing all over the world. With an ageing world population, the costs for healthcare and the shortages in medical staff are continually increasing. In the Netherlands, out-of-hours general practitioner departments (*"huisartsenposten"*) suffer from this increasing pressure: the departments are crowded and the staff capacity is too low to adequately handle the amount of patients. This often leads to long waiting times for patients in need of potentially urgent medical care and to high pressure work environments for staff. There is no insight into when and why it is crowded, how high waiting times emerge from this, and how changes can be made in the departments internally and beyond to reduce this problem.

In this thesis, research is presented that addresses the practical and scientific lack of knowledge of the factors that influence the out-of-hours general practitioner departments and that identifies how waiting times can be reduced. This is done by answering the following main research question:

How can waiting times at the telephone triage of out-of-hours general practitioner departments in the Netherlands be reduced?

A data-driven simulation modelling study is performed to give an informed answer to this question, which makes use of real-life data of the telephone system of two out-of-hours departments. The research can be split up into two parts: the system data analysis and the implementation and use of the simulation model which led to practical recommendations.

PART I: SYSTEM DATA ANALYSIS

An extensive system data analysis was performed to research factors that have an effect on demand for healthcare and on the time it takes to help a patient on the phone: the service times. When these factors are known, scenarios of the demand and the service times can be identified to simulate reality as closely as possible in the simulation model implemented in part II of this research. The researched system is the telephone triage system of the out-of-hours department, which helps patients through the phone when their regular general practitioner is not on duty, for example during the evenings or the weekends. They can be seen by a doctor after the call, but that part of the system is not researched in this thesis. Demand therefore indicates the amount of calls to the department, and service times indicate the time that a call takes.

From analyzing the telephone data of the out-of-hours departments, different conclusions on the factors that impact demand and service times were drawn:

- Demand and service times are different between seasons
- During the week, the demand for healthcare is lower than during the weekends
- During Fridays, the demand for healthcare is higher than during the rest of the week
- On Saturday it is more busy than on Sunday

- The more urgent the medical problem is, the longer the service time is
- Demand can in some cases be different if it is very cold or very hot

By identifying these factors, the real-life data was split up into different groups in which demand and service times are the same: the scenarios. With the identification of these scenarios, insights into what factors can lead to situations with high waiting times at out-of-hours departments were identified.

PART II: SIMULATION MODEL

Next, a discrete event simulation model was implemented that can simulate any day in the out-of-hours department. This simulation model uses the identified demand and service times scenarios from the system data analysis to identify the arrival rates of patients at the department, the necessary service times to help them and the urgency of their medical problems. With scenarios that come from real life data, reality is simulated as closely as possible.

First of all, it was validated that the model accurately calculates waiting times of patients on the phone at out-of-hours departments. It can therefore be used as a means for departments to foresee how busy it is going to be at the department on a specific day, by choosing the corresponding scenario and using that in the model. Also, when validating the model it was found that the regular staff capacity at out-of-hours departments is never enough to help all patients within the officially required time to do so. Therefore, experiments with the simulation model were performed to quantitatively test what system changes other than increased staff capacity can lead to a promising reduction of waiting times at out-of-hours departments.

PART II: PRACTICAL RECOMMENDATIONS

These promising system changes when it comes to waiting time reduction can be split up into quick wins that can be easily and quickly implemented by out-of-hours departments and whose results can be measured directly, and long term interventions that focus more on behavior change, require more cooperation between stakeholders and whose results have to be measured over a longer time frame. These results were discussed with three out-of-hours departments in the Netherlands resulting in practical recommendations to achieve waiting time reductions.

The quick win interventions are the following:

- **Shift patients from peak demand to less busy periods of the day.** This can be implemented by adding extra information to the audio fragments that are played to patients in the queue about peak hours and when to call back. A shift of 1 patient per hour in the peak hours of the day to the less busy hours reduces waiting times by at least 10-20%, which increases to 50% when shifting 4 patients per hour on the busiest day of the week: Saturday. The simulation model needs to be consulted to find the optimum amount of patients to shift, as shifting too many patients only moves the waiting time peak to another part of the day.
- **Let work shifts overlap by starting them every one or two hours, based on how crowded it is.** With this flexibility, sudden demand increases that were not foreseen can be handled with staff capacity at hand. Optimizing the schedule hourly reduces peak waiting time by at least 10%, which can increase to percentages around 50% on Saturdays if a shift takes around 4 hours.

- **Retrieve personal patient information automatically in the queue in stead of by a staff member**. This reduces the time a call takes, which greatly impacts the waiting times: a 10% call time reduction reduces the waiting times by more than 50%. The call times have been increasing over the past years, so it is key to try to reduce them again or at least avoid an extra increase.

The long term recommendations have a solid basis in literature, where they have a demand reducing effect in emergency healthcare departments. Demand reduction has quantitatively proven to reduce waiting times in out-of-hours departments already by 50% when a demand reduction of 10% is achieved. Also, the consulted out-of-hours departments recognize the potential effect of these recommendations. Therefore, the implementing these recommendations has great potential to reduce waiting times over a longer time frame.

- **Increase accessibility and understanding of the primary healthcare system.** By working together with regular general practitioners, *("huisartsen")*, to coordinate accessible opening hours and by informing patients on making appointments and when to use what service, less of the burden lays on out-of-hours departments.
- **Implement a small financial (dis)incentive for out-of-hours care.** People are found to choose the regular general practitioner over out-of-hours care if the former is free and the latter is not.
- **Implement separate lines for home care and nursing homes.** This reduces the amount and length of these types of calls in weekends.
- **Monitor patients who regularly contact out-of-hours departments.** Cooperate with the regular general practitioners of these patients and plan proactive check-ins to reduce their use of out-of-hours care.
- **Implement working from home for triagists in the near future.** This could enable short and spontaneous shifts that reduce waiting time peaks on busier days like Friday and in the afternoons of the weekend days.

PREFACE

Rotterdam, November 2021

The thesis that lies in front of you concludes my time as a master student Engineering and Policy Analysis at TU Delft. A thesis that combines my interest in the healthcare system with my love for data science, programming and simulation modelling which grew over my years as a student. By combining these two, I wanted to perform research that addresses the pressure on healthcare systems, especially relevant in the current pandemic.

When I first started thinking about my thesis proposal, a curfew was still in place and going to university was only possible with a reservation. Over the course of my research process, my hopes to have a graduation at the faculty increased and as I am writing this today, it is (relatively) certain that my graduation can take place in person. It is also certain that I have had great supervision during my research process, and I want to thank all my supervisors for their help, guidance and time. First of all, I want to thank Saba for the many meetings we have had, starting with my proposal and ending just a few days before I finished this thesis. You greatly helped me with conceptualizing and executing my research: I often went into a meeting feeling a bit chaotic and with a lot of questions, and coming out of them my thoughts and next steps were organized again. This greatly helped me to in the end identify the implications of my research. Also, I want to thank Alexander and Yilin for your technical feedback on the modelling and analysis choices I made. You often pointed me in directions that I had not thought of before and made me think critically about why I make certain choices in my research. I also want to give a special thanks to the Esculine team and in particular Thijs for giving me the opportunity to work with real-life data from healthcare departments, to speak to them about my results, for challenging me to build a model where real insights can be derived from and of course for our weekly meetings in which we discussed my progress, but also discussed your work and often had a laugh.

Lastly, I want to thank my friends and family for being there for me throughout my whole study process. My parents and sister for (almost) always being supportive of every choice I made and for their interest in my thesis process. Willem, I want to thank you for your confidence in me and your ability to help me see things in perspective if I felt stuck. Of course, I also want to thank my roommates with whom I spent a lot of days inside the house the past 20 months, for our daily coffee and lunch breaks and for all the fun and distracting activities we did at home.

Hanneke Post

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

The pressure on healthcare systems is increasing. In 2050, 22% of the world population will be older than 60 years old, compared to 12% in 2015 (World Health Organization, 2018). This demographic shift leads to changing employment, retirement and housing, to a decreasing working share of the population and to increasing costs for healthcare, all while there already is a shortage in medical personnel (Levit, Smith, Benz, & Ferrell, 2010; Tinker, 2002). Because of this existing shortage, it is important that healthcare departments can use their resources efficiently while avoiding queues of people in need of medical attention, dissatisfied patients and personnel and high costs for idle employees. To achieve this, they need to be able to foresee when and why it will be crowded at healthcare departments and how internal and external factors lead to long waiting times and a stressful work environment in different situations.

In the Netherlands, inhabitants have multiple options to choose from when they are in need of primary medical care. Each citizen has a general practitioner or GP, in Dutch *huisarts*, who plays a central role in the Dutch healthcare system. An appointment can be made to see a GP during office hours on weekdays. A GP can, if necessary, refer a patient to a hospital to see a medical specialist, or refer them to other types of care. Such a referral is required to receive specialist care. If someone needs after-hours care, they can contact out-of-hours general practitioner departments, in Dutch *"huisartsenpost"*, which is a cooperative of GP's in which each GP has to work a certain amount of hours each year (Smits, Keizer, Huibers, & Giesen, 2014). In case of emergencies, someone will be referred to the emergency department, or the phone number 112 can be called (Tikkanen, Osborn, Mossialos, Djordjevic, & Wharton, 2020).

The out-of-hours departments in the Netherlands experience the problem that it is often very crowded: the demand for care is high and departments are unable to foresee this demand. Also, it takes time to come to help a patient and this combination leads to waiting times for patients in need of medical attention. In these departments *telephone triage* is used: the process that starts when a patient in need of medical attention calls the department, after which a triage nurse or doctor picks up and determines the urgency of the call. Advise on the phone might suffice, if not the patient comes to the department to see a doctor, a doctor visits the patient or an ambulance can be dispatched. Often, a patient has to wait in a queue on the phone longer than the designated norms before being able to speak to a medical professional. This practical problem is supported in literature, where multiple reasons for the increasingly high demand for out-of-hours care all over Europe are mentioned, like ageing, population growth, patient behavior and problems with the availability and amount of healthcare personnel (L. Huibers, Philips, et al., 2014).

It is therefore necessary to gain insight into the factors that influence the demand for care and the waiting times at the telephone triage of out-of-hours departments. This understanding might ensure better help and small queues for patients in need of care and lead to a healthier work environment for the healthcare workforce of the department.

A data-driven simulation modelling approach is used in this research, with data from two out-of-hours departments in the Netherlands, leading to a model where different demand and service time scenarios can be simulated to predict a day in the system as it is now, but also leading to a model which identifies system changes that lead to lower waiting-times, based on which accurate interventions by the out-of-hours department and beyond can be identified. It leads to a model-based advise on reducing waiting times within the telephone triage system in different scenarios for out-of-hours healthcare departments. The main research question is therefore as follows:

How can waiting times at the telephone triage of out-of-hours general practitioner departments in the Netherlands be reduced?

The remainder of this thesis entails a literature review on this subject, the research design with subquestions, the conceptualization of the model, the analysis of the data towards different demand and service time scenarios, the identification of the distributions used in the model, the model implementation, and the results, discussion and conclusion.

2 LITERATURE REVIEW

In this chapter, existing literature on out-of-hours care is explored to see what is currently at play within the field and to identify links between research to see whether the practical problem that out-of-hours departments experience is supported by the existing scientific literature and in what way a knowledge gap exists. Then, simulation research within the healthcare system field and beyond is explored to identify main methods and concepts used for simulation modelling, especially looking at methods that might correctly represent a telephone triage system. Lastly, important modelling processes and concepts are discussed based on existing research, which serve as a framework on which the modelling process in the following chapters is based.

2.1 OUT-OF-HOURS CARE

To find the reviewed literature for each section, keywords and search terms were defined and used to search literature from the scientific databases Scopus and PubMed. For Scopus, the search was within the abstract, keywords and title of the literature. For PubMed, all fields were searched. The inclusion criterion for a paper was that it should be about the system of telephone triage in out-of-hours general practitioner care: about reasons and factors influencing the use of out-of-hours care, about the availability of workforce, about urgency determination or about the way the system works, is organized and might be improved. Papers that were medical or specifically about a disease, that were about emergency departments or other hospital departments or that were about self-referrals were not included. The main papers and subjects within this theme are discussed to see if the practical problem formulated in Chapter 1 is seen in academic literature, or what other types of research are conducted on out-of-hours care. The table with all used search terms is visible in Table 2.1. Note that for Scopus, ”{}” is used for an exact search, while for PubMed, ”” is used. In the search term, the PubMed version is visible. It can be seen that only very few results were found when looking for simulation studies in the out-of-hours field, and none were relevant. The reviewing process towards the included papers for the out-of-hours part of the review is visualized in a flowchart in Figure 2.1.

Table 2.1: Search terms and corresponding results per database

Search term	Found papers PubMed	Found papers Scopus	Relevant papers
((”out of hours” OR ”out-of-hours”) AND (”care” OR ”GP” OR ”general practice*” OR ”primary care”) AND Netherlands)	187	173	37
((”out of hours” OR ”out-of-hours”) AND (”care” OR ”GP” OR ”general practice*” OR ”primary care”) AND ”triage”)	188	229	37
((”out of hours care” OR ”out-of-hours care” OR ”out-of-hours GP” OR ”out-of-hours primary care” OR ”out-of-hours general practice*”) AND ”simulation”)	7	9	0

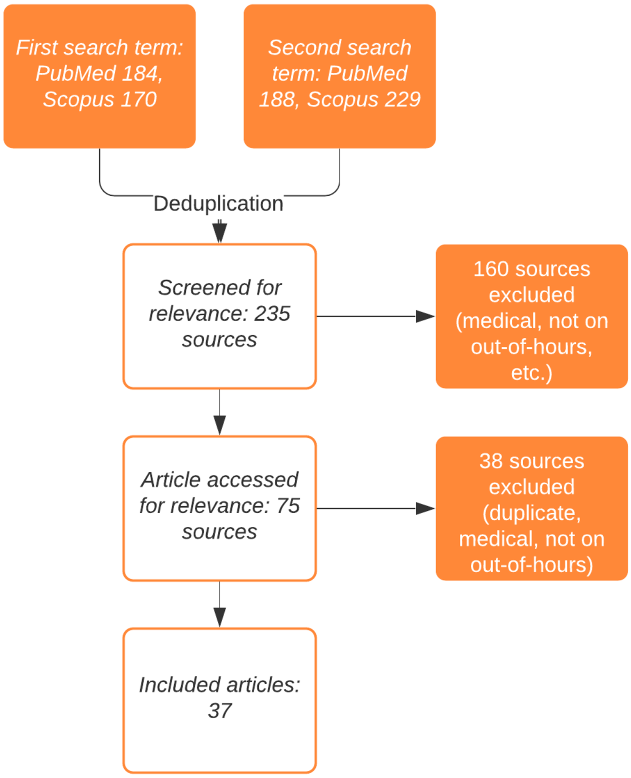

Figure 2.1: Reviewing process flowchart based on PRISMA

Based on the found literature when using these search terms, several categories that most of the papers fall in can be identified. These categories are:

- Busyness at out-of-hours departments
- Urgency determination of patients
- Demographic indicators for use of out-of-hours care
- System improvements and policies
- Comparative studies between countries
- Organizational models of out-of-hours care

2.1.1 Busyness at out-of-hours departments

Many of the research on out-of-hours care in the Netherlands emphasizes the existence of high demand and busyness in healthcare and being able to better manage this demand is a common used reason why research is performed. Keizer, Senn, Christensen, and Huibers (2021) mentions that the problem of high demand is especially visible in out-of-hours care and emergency departments. Because of the experienced busyness due to ageing, the growth of the population, the behavior of patients and problems with availability and amount of workforce in out-of-hours care all over Europe, a European Research Network was even set up to share knowledge between out-of-hours care departments, called EurOOHnet (L. Huibers, Philips, et al., 2014). These papers support the experienced practical problem of high demand, busyness and workforce problems that out-of-hours departments in the Netherlands have, that was discussed in Chapter 1. On the other hand, not that much can be found on the length of conversations at the telephone triage of out-of-hours departments. This length might be a reason for experienced busyness and queues, but up and till now it is only found that variation between the healthcare employee who takes the call exists, and between mental health or non-mental health calls (Mohammed, Clements, Edwards, & Lester, 2012).

2.1.2 Urgency determination

Many other papers look at ways in which the urgency of patients is determined in the telephone triage process in the department. It is often found that variation exists between allocated urgency of similar patients, which is most often attributed to differences between the people calling, like their age and gender (Zwaanswijk, Nielen, Hek, & Verheij, 2015). On the other hand, in Exalto et al. (2021), it is assessed whether a difference exists between the allocated urgency, symptoms and call times of men and women with TIA-like symptoms, and it is concluded that the gender of a person is no direct indication for a triagist to take specific action. This is supported by Van Der Meer et al. (2019), where similar urgency allocations were found for men and women who called with chest symptoms. Another study where differences are analyzed is of Erkelens, Zwart, et al. (2020), where the time of calling of people with TIA symptoms and with mimic symptoms are analyzed, but no differences are found. It is interesting to see that when looking at the whole patient population, differences for similar problems are found, but when looking at more severe issues such as TIA's, these differences are not present.

Not only differences in urgency and call characteristics between patient groups are analyzed, also the accuracy of the allocated urgency that high risk patients receive is often researched, for example by Erkelens, Rutten, et al. (2020); Wouters, Rutten, et al. (2020) to see if the telephone triage process accurately allocates urgency to patients with a TIA, acute coronary syndrome or other life-threatening events. It is found that the Netherlands Triage System underestimates the urgency of these patients in 27% of the cases. If triagists interfere and overrule the automatic urgency, it reduces to 14% with poor efficiency of the triage process. This reduction is supported by Wouters, Zwart, et al. (2020), which states that triagists sometimes feel a misalignment between the computer decision support systems choice of urgency level for a patient with acute cardiac symptoms and their own choice of urgency.

Lastly, it is often found that out-of-hours services are not appropriately used. For example, in Denmark, it was found that one in four of the calls to an out-of-hours department are not urgent and could have been avoided (Nørøxe, Huibers, Moth, & Vedsted, 2017). In Germany, it was found that for the most calls, the reason for encounter (RFE) was in fact not urgent enough to call the out-of-hours department, and people could have waited until the opening hours of their regular general practitioners (Leutgeb, Engeser, Berger, Szecsenyi, & Laux, 2017). Reasons for this are mentioned in the next section.

2.1.3 Demographic indicators for use of out-of-hours care

Reasons why people call the out-of-hours departments and differences in calling behavior between groups within the population are often analyzed in literature. As the demand for care is so high, it is often mentioned that it is key to know why certain people call at all, why they might call often and how it can be reduced to only necessary care. One of the most researched factors that influences calling behavior is found to be socioeconomic status of a patient or whether or not someone is a non-western migrant. In Keizer et al. (2021), it is found that non-western migrants have more contacts with emergency care than native born people or western migrants. These non-western migrants also tend to call their general practitioner often. These differences can be partially explained by employment status, anxiety levels, attitude towards healthcare and the ability of people to access their own day-time general practitioner. They feel fewer barriers to contact the out-of-hours services and are more anxious about symptoms, have problems planning a consult with their own general practitioner, which was already found in Keizer et al. (2017), and are more often unemployed. The papers mention the need for an increase in awareness and better access to daytime general practitioners, to help people make the right choice within the healthcare system. Also, it is recommended that people are educated about the purpose of the out of hours department and the fact that it is easier to access daytime care services. Next to these indicators, there are also behavioral aspects that might disturb the process at out-of-hours care departments. It is found that often, non-western migrants expect action in the form of a visitation or medication from healthcare services and can be dissatisfied with only advice trough the phone. This appears to be a general tendency of people calling the department as well, as people tend to be unhappy when they receive advice on the phone and are not seen by a doctor (Van Uden, Ament, Hobma, Zwietering, & Crebolder, 2005) and can sometimes even become aggressive (P. Giesen, Mokkink, Hensing, van den Bosch, & Grol, 2008). To address this, it is important that expectations of people calling the department are managed and that the emotions of the caller are identified.

As mentioned above, the socioeconomic status (SES) of patients and the sociodemographic characteristics of a neighborhood are found to be an important indicator for (poor) use of out-of-hours care (Jansen, Zwaanswijk, Hek, & De Bakker, 2015): the probability of using out-of-hours care increases for each lower income group Jansen, Hek, Schellevis, Kunst, and Verheij (2020). It is also a good indicator for underlying health problems as symptoms differ between socioeconomic groups and for how good a person is at expressing the need for help (Jansen, Hek, Schellevis, Kunst, & Verheij, 2021; Jansen et al., 2018). Neighbourhoods were more women live, where the general income level is lower and where there are many immigrants tend to have a higher demand for out of hours care. Jansen et al. (2020) recommends further research into how the telephone triage process can be improved for vulnerable groups with a lower SES. Similar to the paragraph above, where education on the purpose of out-of-hours care is recommended, in Jansen et al. (2018) improvement of the use of healthcare services by these groups with a lower SES is recommended by improving their health literacy.

Next to socioeconomic status or migrant status, an intellectual disability is found to be an important indicator for how often out-of-hours departments are called: they tend to call more often and their calls are rated with lower urgency levels than calls of people without an intellectual disability (Heutmekers et al., 2017). Lastly, it is found that older people who contact the out-of-hours services tend to be in urgent need of medical care (Smith & Carragher, 2021), tend to have a high call rate that increases with age (Haraldseide, Sortland, Hunskaar, & Morken, 2020) and when they also have a consultation at the department itself they are more likely to be admitted to a hospital later on or to use home care: not a direct indicator for an increased use of out-of-hours care, but it is an indicator for the use of further care. Training of triagists to recognize geriatric problems and identification of these patients before and after their visit to the department could help these patients (Bloemhoff et al., 2020).

2.1.4 System improvements and policies

In the past few subsections, different categories of research within the field of out-of-hours care were discussed. Many of these papers end with some kind of recommendation to improve or reduce the use of out-of-hours care, to better educate people on what its purpose is, to improve overall health literacy of more fragile groups and to make day-time general practitioner care more easily accessible and available (Jansen et al., 2018; Keizer et al., 2017; Keizer, Maassen, Smits, Wensing, & Giesen, 2016; Keizer et al., 2021, 2015; Smits et al., 2015) Also, identification and support of people with higher risks of health problems or who tend to contact the out-of-hours services more often, respectively elderly people (Bloemhoff et al., 2020) and people with intellectual disabilities (Heutmekers et al., 2017), was recommended.

Next to these recommendations, some papers go deeper into what changes or incentives can be used to improve the functioning of the out-of-hours care system as a whole. One of the proposed changes is the use of online advice by patients prior to contacting the out-of-hours department, to see whether contact with the department is necessary. Often, some online advice is enough, or a visit to the daytime general practitioner for example the next day suffices. In M. J. Giesen et al. (2017), it is found that giving this type of online advice has a high potential to reduce unnecessary use of out of hours services. Other incentives like giving an overview of the medical costs or a next day appointment with the general practitioner also had influence on patients decisions for urgent care. Also co-payment for patients, where now out-of-hours care is free, is an incentive to reduce the use (Keizer et al., 2016). A step further for the online advice strategy is used in Verzantvoort, Teunis, Verheij, and Van Der Velden (2018), where the app "Should I see a doctor?" is tested (*'moet ik naar de dokter?' in Dutch*) to see if it can help guide people to appropriately contact the out-of-hours department. It is found that it could be valuable for patients and also for the department, as 65% of people intend to follow the apps advice, and 81% of the urgency levels given by the app correspond to the urgency level that a triagist would assign to the patient.

Other proposed changes are more in the direction of the triage system, to make it more efficient or to make sure urgency is more adequately determined: research which base is not specifically the high demand for care like in most of the aforementioned papers, but which base is the need for a high quality triage system. One of the recommendations is to alter the display of the Netherlands Triage System to have fewer options for diagnosis, so that questions asked by triagists are less ambiguous and have less possible answers and will provide more clear answers (Erkelens et al., 2021). This is especially interesting, as in Section 2.1.3 it was found that people with a lower SES have more trouble expressing the need for help (Jansen et al., 2021), and this proposed change might make that easier due to less ambiguity. Another proposed way to adequately determine urgency is proposed in L. Huibers et al. (2012), where they state that it might be more important for a triagist to recognize patterns to identify health problems in a call, than to ask all the questions the triage system prescribes during a triage. This is in line with what was found in Erkelens, Rutten, et al. (2020), earlier mentioned in Section 2.1.2, where overruling by the triagist of the automatically allocated urgency by the system ensured a reduction of the underestimation of urgency from 27% to 14% in patients with coronary problems or other life-threatening events. This gives the image that leaving room for human interactions and decisions in relation to the triage system ensures better quality of the triage system.

2.1.5 Comparative studies

Some of the research on out-of-hours care compares the systems between countries, which are shortly explained in Section 2.1.6. Another interesting comparison that is made is to see how behavior of patients from different countries can vary. For example, L. Huibers, Moth, et al. (2014) finds that the Danish have a larger amount of contact with their out-of-hours services than the Dutch do: in some categories even almost double. In L. Huibers et al. (2018) it is also found that there are some differences between the behavior of Dutch, Danish and Swiss people when it comes to seeking out of hours help, especially for parents with young children. It is however unclear if this is due to personal preferences, cultural differences or differences within the healthcare system of the countries. Furthermore, it is found that the scope of diagnoses in out-of-hours care is very similar across many European countries L. A. Huibers et al. (2011). This is in line with what L. Huibers, Moth, et al. (2014) recommends: to find out what factors might be connected to contacting the out-of-hours services.

2.1.6 Organizational models in out-of-hours care worldwide

There is a large variation in organizational models used by countries for out-of-hours care. The most often used model is the general practitioner cooperative, or the GP cooperative where many general practitioners work together and take turns to serve their patients during night time and out of hours. This is also the model the Netherlands uses (Steeman et al., 2021). Many other models are seen internationally: there are smaller GP cooperatives called rota groups, or even out-of-hours centers that are run by only one general practitioner. It is also possible that the emergency department of a hospital takes care of all out-of-hours services. There are also countries where there are primary care centres and minor injury centres, where you can visit without appointment.

2.1.7 Conclusion

A main conclusion of the past sections is that the practical problem of a high demand for care and the waiting times that are experienced by out-of-hours departments in the Netherlands is also a phenomenon that is often addressed in literature. It is - next to the need for a high quality triage process and accurate allocation of urgency - the base for a lot of research into demographic reasons and other indicators for contacting out-of-hours care. Many recommendations for improvement are made in literature that might lead to reduction and improvement of the use of out-of-hours departments. It can however also be seen that no simulation studies have been performed within the out-of-hours literature to test there recommendations, and therefore no prior research has been performed on the performance of the telephone triage of the out-of-hours department system with the use of real data, to be able to implement system changes to reduce waiting times and to see what has an impact on the hourly and daily demand for care and its service times. This also means that some of the recommendations that have been proposed in the reviewed literature, like for example finding out what the factors are that contribute to contacting the out-of-hours-department, are carried out in this research.

2.2 SIMULATION MODELLING IN HEALTHCARE SYSTEMS

Next to getting an overview of the research within the out-of-hours field, simulation studies on (a part of) healthcare systems are consulted, to get an overview of the main methods used within healthcare systems simulation research. This gives a grounded reason for the chosen simulation modelling method Discrete Event Simulation, as further discussed in Chapter 3.

First, research in the field of simulation modelling in healthcare systems was explored, as visualized in Figure 2.2, to identify the main simulation techniques and its applications. Research on healthcare systems that involved discrete event simulation was then reviewed. As no literature was available on simulation of telephone triage and out-of-hours care, some literature on simulation studies of call centers was reviewed to identify the main issues and methods in that field.

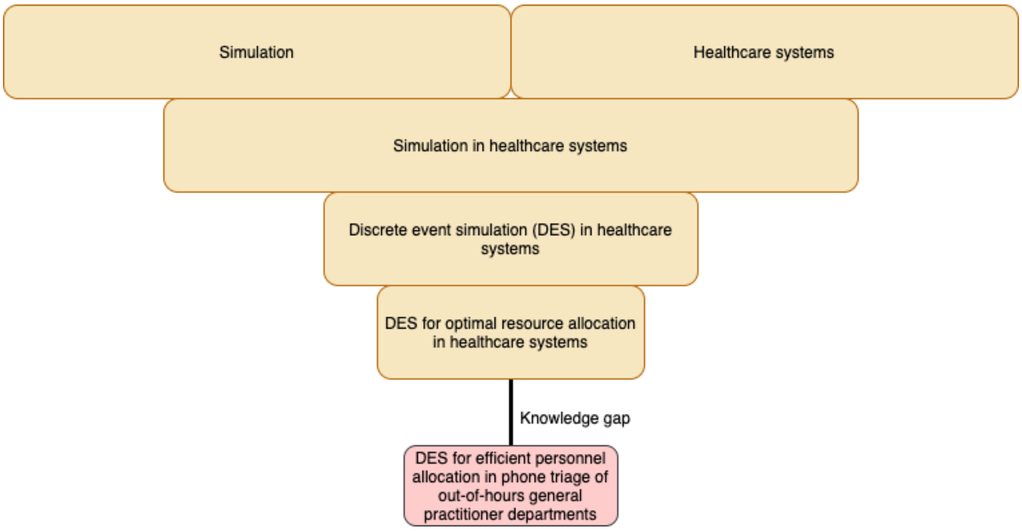

Figure 2.2: Reviewing process towards knowledge gap

In Table 2.2, all used search terms and the corresponding amount of results per database are displayed.

Table 2.2: Search terms and corresponding results per database

Search term	Found papers PubMed	Found papers Scopus
”simulation” AND ”healthcare systems”	177	1175
”general practitioner” AND ”simulation” AND ”staff”	5	27
”emergency department” AND ”simulation” AND ”staff”	182	306
”discrete event simulation” AND ”healthcare”	17	578
”staff allocation” AND ”healthcare”	24	19
”call center” AND ”simulation”	1	328

2.2.1 Simulation Methods in Healthcare Systems

This subsection dives into simulation techniques used in healthcare systems, to define the key methodology and to identify which technique is used for what application within the healthcare system domain. This enables a review of the methods that are often used for the optimization of resources and which might be used to correctly represent a telephone triage system.

According to Mielczarek (2016), simulation modelling is recognized as one of the most common approaches to healthcare management problems, with the main modelling methods being Monte Carlo Simulation, Agent-Based Modelling, System Dynamics and Discrete Event Simulation.

Monte Carlo Simulation is often used to determine optimal capacities like for beds (Akkerman & Knip, 2004), or adequate vaccine supply (Dhamodharan & Proano, 2012), because of its ability to use historical data and to run thousands of potential scenarios (Mielczarek, 2016). Agent Based Modelling models individual behavior of an 'agent', and is in the healthcare domain often used to model disease spread or demand for healthcare (Knight, Williams, & Reynolds, 2012). System Dynamics is often used when people can be divided into large groups and the state of the individual patient is not followed. It aims to understand the the structure of a system and the relation it has with its behavior (Demir, Gunal, & Southern, 2017). Lastly, multiple healthcare simulation surveys state that Discrete Event Simulation is the most often used technique in healthcare management. It is a stochastic method, often used when research is looking at arrival times and process times (Mielczarek, 2016), which are variables that are at play in the telephone triage system at an out-of-hours department. It is used for improving patient flows, for the management of beds and for resource scheduling like for example personnel planning. In general it is the most versatile methodology for modelling healthcare systems (Jun, Jacobson, & Swisher, 1999; Zhang, Grandits, Härenstam, Hauge, & Meijer, 2018), and complete reviews have been written on its wide application and problem solving abilities within the healthcare domain (Günal & Pidd, 2010). It can therefore be concluded that Discrete Event Simulation is a main simulation technique used in literature to optimize resources like beds, rooms and staff.

2.2.2 Discrete Event Simulation in healthcare systems

According to Jun et al. (1999), healthcare simulation research that uses Discrete Event Simulation often has an objective that relates to scheduling and patient flow or to sizing and planning of beds, rooms and staff (workforce planning), or it involves a discussion on future research areas. The focus is on reviewing literature on optimizing sizing and planning of staff and reducing waiting times in healthcare systems.

Research on staff scheduling has been performed in various ways, with great variety in their definitions of optimization. Ghanes et al. (2015) simulates the 'Length of Stay' within the emergency department and the time it takes before a patient sees the first doctor in emergency rooms, under different percentages of increased staffing budget. They find that budget increase for personnel has a decreasing effect on the Length of Stay at the emergency room, where in Zeltyn et al. (2011), budgeting is not mentioned, but it is found that staff gives better performance when workload is allocated over time in stead of calculated at the arrival of a patient. These two papers focus on performance indicators that are affected by changes in staff policies, but do not conclude how resources should be allocated in various system situations. This observation matches the review of Salmon, Rachuba, Briscoe, and Pitt (2018), which states that future research should have a focus on increase in demand, crisis scenarios and how systems can recover from a fall in performance.

A paper that does address different strategies for the allocation of staff, but does not address variability of demand and the cause of it, is Mukherjee (1991). In this paper, the waiting times of patients were reduced because of good scheduling, while staying in control over the costs for resources. It can therefore be observed that optimal scheduling leads to beneficial outcomes for healthcare departments and that some interventions like increased budgeting positively affect performance, but that there is a need for research that takes into account the variability of demand and what causes it, to confidently understand the emergence of waiting times.

2.2.3 Telephone triage simulation

In Section 2.1, it is found that when it comes to telephone triage and out-of-hours care, no previous simulation modelling studies have been conducted. Systems that technically work similarly to telephone triage systems are call centers. Research has been performed for call centers to improve the performance of call centers, customer's satisfaction and to reduce costs (Kim, Lee, & Choi, 2005), but these types of costumer service call centers are hard to compare to telephone triage system within a healthcare system where real health problems have to be dealt with and lives could be at stake. The performance of a telephone triage system should therefore be measured differently. The research on out-of-hours departments and telephone triage has been discussed extensively in Section 2.1, and generally aims at the safety of telephone triage and the quality of the calls (Montalto, Dunt, Day, & Kelaher, 2010), the mismatch of evaluation criteria for welfare between patients and healthcare personnel (Brasseur et al., 2019) or demographic factors that impact the demand and use of out-of-hours departments (Jansen et al., 2015; Keizer et al., 2021), and does not quantitatively analyze the performance of the system for different external and internal variables.

2.3 CONCEPTS WITHIN DISCRETE EVENT SIMULATION

Within Discrete Event Simulation (DES) research, different concepts, theories and methods are used. The remainder of this literature review explores these concepts in literature, to come to a grounded modelling framework which can be used to answer the main research question and the subquestions. The concepts defined below are used for conceptualization of the model in Chapter 4 and for implementation of the model in Chapter 6.

2.3.1 Demand variability and arrival process

A first main characteristic of DES is that it is often used when research incorporates arrival rates and service times (Mielczarek, 2016). Within simulation research in healthcare systems, the arrival, waiting and handling times often encompass the arrival of patients at a healthcare department, the amount of time it takes before they receive medical attention and the time the treatment or conversation with a healthcare professional lasts.

The arrival of patients to a healthcare system can be seen as the demand side of the system to be modelled in a discrete event simulation model. This arrival often follows a Poisson process, where the inter arrival times follow an exponential distribution and arrivals are independent of each other (Jahn, Theurl, Siebert, & Pfeiffer, 2010), like for example in e Oliveira, de Vasconcelos, Almeida, and Pinto (2020) and in Pinto, De Campos, Perpetuo, and Ribeiro (2015) where it is used in the context of patients arriving to the hospital and the allocation of hospital beds, or in Keshtkar, Salimifard, and Faghih (2015) where it is used with the aim to allocate resources effectively in an emergency department.

A first method, out of three, seen in literature to retrieve the inter arrival times and arrival rates, is directly drawing from the operational data. It can be used as input into a model directly and no statistical distribution is used within the simulation model to model the inter arrival times. An example of this approach is seen in (Zhu, Hen, & Teow, 2012), where ICU bed capacity is estimated using discrete event simulation. To get the inter arrival times, the actual arrival times from the operational data are used as input for the discrete event simulation model, as well as the length of stay at the ICU. They note that a problem with this is that if data is not complete, for example if a patient did not enter the ICU in the case of this paper, no service times can be drawn from the data. Then, a distribution still has to be fit on the existing data of other patients.

As stated previously, in Salmon et al. (2018) it was mentioned that future research on staff scheduling should incorporate increase in demand - so the arrival rate - and the impact of scenarios on staffing. This statement is supported when looking at recent literature on simulation optimization research for hospital beds (Keshtkar et al., 2015) and similar research for improvement of patient experience in emergency departments (Abo-Hamad & Arisha, 2013), where inter arrival times of patients from various data sources are used but remain static throughout the simulations and do not take into account possible variability of demand for care. More and more research stresses the importance of using some sort of prediction method for predicting arrival rates for different processes. Methods like statistical testing, distribution fitting and regression as used in (McCarthy et al., 2008) and mentioned in (Gul & Celik, 2020) are the second often proposed way of retrieving arrival rates. The last method goes a bit beyond that and proposes artificial neural networks and other forms of machine learning for the retrieval of arrival rates, like Xu, Wong, and Chin (2013) where they use an artificial neural network for daily patient arrivals at an emergency department, in Höpken, Eberle, Fuchs, and Lexhagen (2021) where it is used to predict tourist arrivals and in Hill and Böse (2017) where it is used to predict truck arrivals and waiting times. It should be noted that while these papers do use a predictive or machine learning method to predict arrivals in a system, after prediction the arrival rates they do not model or simulate that system using discrete event simulation. In fact, literature reviews have been written on predictive methods for arrivals at, for example, a hospital emergency department (Gul & Celik, 2020), but many of the papers reviewed do not use their predictions in a (DES) simulation model, and the ones who do often use static inter arrival times with no variability.

Based on this knowledge, it can be concluded that inter arrival times are retrieved in various ways in literature: direct drawings from operational data Zhu et al. (2012), using historical data for statistical testing and distribution fitting (Pinto et al., 2015), or by developing predictive models that assess which factors contribute to arrival rates and predict those rates (Höpken et al., 2021; McCarthy et al., 2008; Xu et al., 2013). The last of these two methods have the attribute that they can account for variability in demand based on known temporal, demographic or climatic factors, which is often used already in papers where some sort of arrival process in a healthcare system is predicted, like in McCarthy et al. (2008) that was mentioned before, but also in Hamrock, Paige, Parks, Scheulen, and Levin (2013) and in Marcilio, Hajat, and Gouveia (2013) where seasonal, daily and temperature factors are taken into account. It should however be noted that none of these papers use this variability of demand as input into a discrete event simulation model.

2.3.2 Queuing and server process

Another concept within discrete event simulation often used when healthcare departments are simulated is the process of queuing for service after arrival, and the server process. The server process can be seen as the supply side of the system. This server process in healthcare systems can consist of one or more processes at a healthcare department, such as triage, registration, screening, observation, or a procedure (Choon, Dali, Beng, & Magdalene, 2016). These processes can follow different distributions, like Gamma, Beta, Weibull, Exponential or Lognormal distributions (Keshtkar et al., 2015), or empirical distributions derived from historical data are used. In healthcare applications often the exponential distribution is not applicable, since treatments tend to take some time (Gupta, 2013). These service times are often derived by the simulation by taking an average, by taking certain intervals or by type of patient for which service times differ from each other (McCarthy et al., 2008) or they are simulated by the model and thus not used as input, like in Hossain, Debusk, Hasan, Jaradat, and Khasawneh (2017), where the service time of patients in a blood test lab is simulated.

Types of queuing models
Different types of queuing models are used within discrete event simulation models. A queuing model handles the before-mentioned demand and supply side of a system and is used in many different areas, such as telephone and communications and logistics (Lin, Wu, Chen, & Chen, 2019) and has as its most basic form the M/M/1 model - written in Kendall notation - where arrivals are Poisson distributed (first M), service times are exponentially distributed (second M) and there is one server (Tiwari, Gupta, & Joshi, 2016), so in a healthcare application that would mean one doctor, nurse, bed, operation room, etc. There are many versions of the queuing model, including models where more servers than one are present, M/M/s, or even more complex models to determine bed allocation like the M/Ph/c/N model where the N is added to note the maximum capacity of the system, c are the amount of beds and the service times follow a phase-type distribution (Gorunescu, McClean, & Millard, 2002). In other models, the arrival rates and service times are not stationary and vary over time, like for example in Green (2006), where it is proposed to construct multiple M/M/s models for the different times of the day where arrival rates or service times are different. Another approach is the M(t)/M(t)/c model, used in Chen, Govindan, Yang, Choi, and Jiang (2013) for truck arrivals, where arrival rates and service times vary over time. This approach is interesting for this research, as many papers state that the demand for healthcare varies over temporal variables (Hamrock et al., 2013; Marcilio et al., 2013; McCarthy et al., 2008). Another characteristic of a queuing model is its service protocol when it comes to the queue. This can follow the 'First in First Out' principle (FIFO), but could for example also follow the 'Last in First Out' principle (LIFO), where some type of priority queuing is implemented (Gupta, 2013).

2.4 CONCLUSION: SCIENTIFIC CONTRIBUTIONS

After the literature review, several scientific contributions of the research performed in this thesis can be identified.

- A simulation study is conducted, using discrete event simulation, where the telephone triage system in an out-of-hours department is represented.
- The demand for care at out-of-hours departments is analyzed, important indicators that impact demand and scenarios in which demand varies are identified, using data of two departments in the Netherlands. This is a new addition to the demographic indicators that are identified in research, as mentioned in Section 2.1.3.
- The indicators that effect the service times within a telephone triage system are identified. This adds knowledge to the found variation in call lengths between personnel and between mental-health and non mental-health calls as addressed in (Mohammed et al., 2012).
- The different demand and service time scenarios based on the real data are incorporated and combined in a simulation model, to represent reality as well as possible within the simulated system.
- System changes that reduce waiting times are identified based on which accurate interventions recommendations can be made.

3 RESEARCH DESIGN

3.1 RESEARCH APPROACH

The main research question is answered using a modelling approach, as there is a lack of understanding how different system inputs and internal variables impact the out-of-hours system and its performance. Using a modelling approach, it can be explored how system factors impact system outcomes that relate to waiting times and performance on norms, which reflect patient and employee satisfaction. The modelling approach is data-driven and uses the real-life data of 2 out-of-hours general practitioner departments in the Netherlands. The data of the out-of-hours departments was analyzed to identify demand and service time scenarios, and the patterns within the two data-sets are compared to each other to see whether out-of-hours departments have similar patterns and trends in their data. With the implemented model, the system as it currently can be simulated with different demand and service times scenarios that are derived from the data, in Chapter 5. Experiments were run using the model and the real-life data in order to find the system changes that lead to waiting time reductions. This led to the identification of system interventions that could potentially lead to those system changes.

Specifically, the modelling approach entailed a discrete event simulation modelling approach. This simulation approach is often used in healthcare systems studies when the focus is on optimizing allocation of for example beds, patients and personnel (Jun et al., 1999). It is a stochastic method that is often used when a study looks at arrival-times (Mielczarek, 2016), which was used in this study. Its stochasticity can however also be seen as a limitation and something to be aware of as the research is conducted, as results from the model are a little different each time the model is run (Caro & Möller, 2016; Caro, Ward, Deniz, O'Brien, & Ehreth, 2007). This calls for a large number of runs, which takes a long time specifically when complex simulations are run. This can constrain uncertainty and sensitivity analysis of the model (Caro & Möller, 2016; Caro, Möller, et al., 2007).

3.2 SUB QUESTIONS

The main research question was answered by answering 4 subquestions. The research followed the steps of a modelling approach, with the following steps: problem definition (executed in Chapters 1 and 2), conceptualization (Chapter 4 and 5), formalization, implementation, analysis, setup and model use (Chapters 6 and 7). (EFSA PPR Panel, 2014; PBL Netherlands Environmental Assessment Agency, 2013). The identification of main concepts of the model and the relations between them, the conceptualization, was conducted in the first two subquestions below. Based on the conceptualization, the model was formalized and implemented. The built model was then verified and validated after which it was extensively used. The results of the model use provided an answer to subquestion 3 and 4.

For each subquestion the required data and its sources, the used research methods to gather the data and the tools used to analyse the data with are mentioned and discussed.

1. **What concepts and performance indicators are needed to accurately model telephone triage at an out-of-hours department?**
2. **What variables effect the demand and service time behavior of the out-of-hours telephone triage system?**
3. **What system changes can reduce waiting times in the out-of-hours telephone triage system compared to the current situation?**
4. **What possible interventions can lead to the system changes (identified in subquestion 3) of the out-of-hours telephone triage system in which waiting times are reduced?**

3.3 RESEARCH METHODS

In this section, for each subquestion the required data and its sources, the used research methods to gather the data and the tools to analyse the data with are mentioned and discussed. Each subquestion is related to one or more steps within the modelling approach, as explained in Section 3.2. This provides a comprehensive flow of the research design. A visualization of the flow of the research is found in Figure 3.1: the research started with an analysis of the system itself and its outcomes by data analysis and conceptualization in subquestions 1 and 2 to be able to implement a model that correctly displays the waiting times, followed by the identification of system changes that lead to a reduction of waiting times by means of experimenting, which in its turn led to the search for and identification of interventions that might result in these system changes and thus in waiting time reduction.

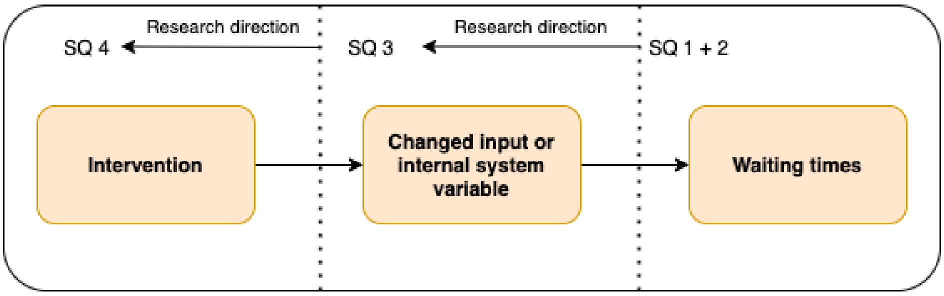

Figure 3.1: Research flow

1. **What concepts and performance indicators are needed to accurately model telephone triage at an out-of-hours department?**
This subquestion covered the conceptualization phase of the modelling approach, together with subquestion 2. Data from real-world out-of-hours general practitioner departments is necessary to identify the main concepts and indicators of the system. This data was provided by two out-of-hours general practitioner departments with different geographical locations in the Netherlands. The research method was desk research: next to the use of the provided data, literature on previous studies related to the field was consulted to validate and possibly extend the choice of concepts within the model, and to identify possible ways of incorporating data into the model used in literature. Also, people from out-of-hours departments were consulted to verify the use of the correct concepts and indicators in the model. The tools used were

the online literature databases Scopus, PubMed and Google Scholar and the resource manager Mendeley to keep track of the consulted literature.

2. **What variables effect the demand and service time behavior of the out-of-hours telephone triage system?**
   The conceptualization of subquestion 1 was needed to adequately analyze the data from the out-of-hours department and to be able to draw valid conclusions on its behavior. In this research question it was tested whether the behavior of the system, especially for demand (patient arrivals) and service time (call duration) is influenced by temporal or weather factors, to see whether scenarios can be identified in which the behavior of the system is significantly different from other scenarios. The research method to identify these scenarios was data analysis on the data of two out-of-hours departments using the tool Python and within Python the data analysis packages *Pandas*, *NumPy*, *SciPy* and *statsmodels* and the visualization packages *matplotlib* and *seaborn*. It was checked whether the found patterns are similar for the two data-sets, to be able to say something about the to what extent the results that the model generates can be used by other out-of-hours departments with the same type of system.

3. **What system changes can reduce waiting times in the out-of-hours telephone triage system compared to the current situation?**
   This subquestion covered the formalization and implementation phase of the discrete event simulation model of the telephone triage system, performed in Chapter 6. This was done by implementing the model as described in the conceptualization phase, in Chapters 4 and 5. The model was implemented with the tool Python using the simulation package *salabim*. When the model was implemented, the model validation and verification were executed to define to what extent it represents the actual system (Cook & Skinner, 2005). This was done by conducting sensitivity analyses on the gathered model results and by comparing it to the conceptual model and by using the identified demand and service time scenarios from subquestions 1 and 2. The results are also compared to real-world data and to experiences of experts and people working in the field, found in Section 5.6. Scopus, PubMed and Google Scholar and non-academic engines are used. With a validated and verified model, it could be used to make valid predictions of current situation and of what would happen with certain system changes. Different experiments were run to identify system changes that lead to a reduction of waiting times at out-of-hours departments.

4. **What possible interventions can lead to the system changes (identified in subquestion 3) of the out-of-hours telephone triage system in which waiting times are reduced?**
   This subquestion builds on the findings of subquestion 3. In subquestion 3, system changes that lead to a reduction of waiting times in the out-of-hours system were identified. To accomplish these system changes, interventions needed to be identified that could push the system as it is now to a changed system that leads to a reduction of waiting times. In this subquestion, those interventions were identified: some of them within the decision power of the out-of-hours department but also some of them out of its decision power. These interventions could lead to changes in internal system behavior such that the waiting times are reduced, leading to increased norm performance, a better work environment for triagists and better service for patients. The interventions were identified based on consultations with people working at the

out-of-hours department and on literature, using PubMed, Scopus and as a first exploration means Google Scholar will be used.

3.4 LIMITATIONS OF RESEARCH METHODS

It is important to address some of the limitations of the main research methods used. The limitations of discrete event simulation can be read in Section 3.1. For the desk research method, a limitation is that it is hard to get a complete and unbiased image of the literature. To address this, it is important to identify the goal of the literature research and be thorough in finding main papers and authors within the field. This is accounted for by using multiple search engines and by combining and scanning all found papers for relevance, as visible in Figure 2.1 in Chapter 2. For empirical data gathering, a pitfall is personal bias when creating survey or interview questions. A way to address this is to have others look over your questions and test them. For data analysis it is important that the source of the data is verified and validated, to avoid faulty conclusions being drawn from the analysis. For exploratory modelling, the stochastic discrete event simulation model will be used, so many runs are needed to avoid conclusions based on a few runs with coincidental results.

3.5 CONCLUSION

In this chapter, the modelling approach that is used in this research was introduced as well as the reason why it is most suitable for this type of research. Next, the four subquestions that lead to the answer to the main research question were identified and for each of them it was explained what data and methods were used to answer them. Lastly, the limitations of the used research methods were discussed.

4 SYSTEM CONCEPTUALIZATION

In this Chapter, subquestion 1 is answered. This chapter combined with Chapter 5 forms the conceptualization phase of the modelling approach towards a discrete event simulation model on telephone triage implemented in Chapter 6.

4.1 SUBQUESTION 1: CONCEPTS FOR MODELLING TELEPHONE TRIAGE

In this section, the concepts and research methods necessary for modelling a telephone triage system are explained. For some of the concepts used, a reference will be made to the literature review in Chapter 2, where some of the main concepts of discrete event modelling especially in healthcare systems have been discussed.

4.1.1 Data from out-of-hours departments

The data that is used in this research is from two out-of-hours departments in the Netherlands. Their systems work similarly, and all analyses performed on the data of the first department in Chapter 5 are also performed on the data of the second department in Appendix A.

4.1.2 Telephone triage processes

The telephone triage process is logged in the available data. Based on this data, together with a visit to the first out-of-hours department all aspects of the telephone triage process were identified. This information leads to the following description of the system process:

At a telephone triage department, calls come in from people in the surrounding region of the department. These calls come in at a certain **rate** per hour. There is medical personnel present taking the calls. These people are called **triagists**. The amount of triagists varies for different shifts and days, where shifts take 8 hours. The time it takes for them to handle a call, is called the **service time**. When someone calls the department, they are placed in a **queue** if none of the triagists is available. If a caller is at the front of the queue and an available triagist accepts the call, the caller is removed out of the queue and a conversation concerning the medical problem is started. A call can be a call to the emergency line, or a call to the normal line. A call to the emergency line immediately goes to the front of the queue: they are seen as very urgent calls. Normal calls have to get in line in the back of the queue. When the conversation starts, the triagist asks questions based on the information the caller provides and walks through the questions of the Netherlands Triage System. When this process is finished, the caller gets a certain urgency level allocated, between

0 and 5, from most urgent to least urgent. Based on this level and the type of problem, further action is taken with the patient, outside the scope of the telephone triage system. Sometimes, no urgency is allocated to a patient in (often) shorter calls. This means that no (complete) triage was performed on the patient, and these conversations tend to last shorter than full triage conversations. There are therefore two subprocesses in the system: a call where no full triage took place, or a full triage call. When a call is finished, the triagist has 40 seconds to finish all administrative tasks that come with a call. This is a guideline, a triagist can of course only start answering a phone call when they have finished this task. It can therefore happen that a triagist takes longer or shorter than 40 seconds for this last task in the process. A visualization of the telephone triage processes (arrival and service process) is visible in Figure 4.1.

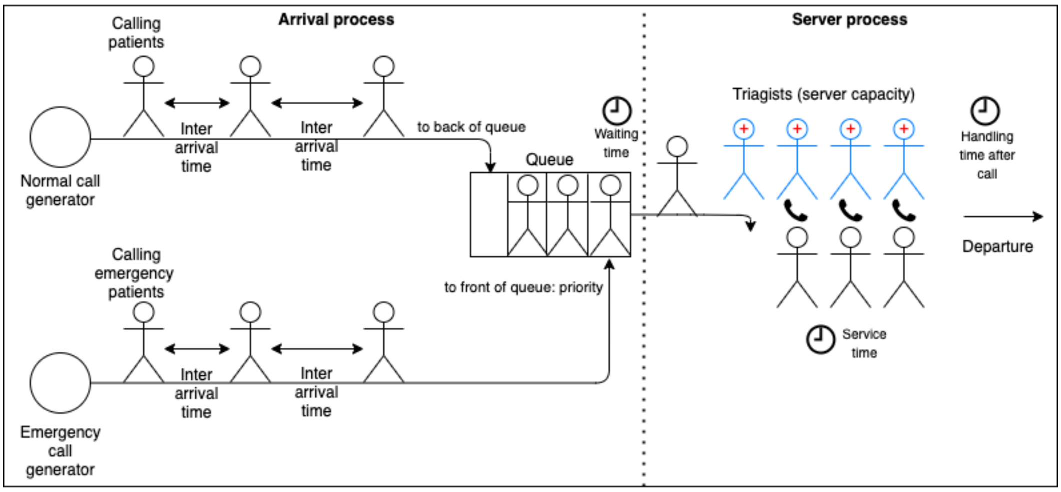

Figure 4.1: Processes in telephone triage of out-of-hours department

The data that was made available by the departments, is data that logs a phone call and stores information on the call. Also, data was provided that gives the urgency of a calling patient. This data was merged and its shape is displayed in Figure 4.2. The columns are in Dutch as the data comes from a Dutch department. From left to right, the columns stand for Urgency, the exact time the call was answered by a triagist, the registered time that the call was answered by a triagist in the second system (often a few seconds later), the time that the caller started the call, the line the caller called to (normal or emergency), the waiting time and the service time.

	Urgentie	answer_time	Oproep_aangenomen	start_call	wachtrij	wachttijd	gesprekstijd
0	U3	2018-08-22 01:24:34.303	2018-08-22 01:24:46	2018-08-22 01:23:19.303	Patienten	0 days 00:00:36.000028800	0 days 00:05:47.000025600
1	NaN	2018-08-22 01:28:57.087	NaT	2018-08-22 01:27:54.087	Patienten	0 days 00:00:23.000025599	0 days 00:02:36.000038400
2	U4	2018-08-22 01:36:47.477	2018-08-22 01:36:57	2018-08-22 01:35:24.477	Patienten	0 days 00:00:43.999977600	0 days 00:07:40.999987200

Figure 4.2: Available data combined with patient urgency

To be able to accurately model this system, certain information needs to be available as input variable or as internal variable. In this list, these necessary variables are mentioned, along with the necessary methods to be able to retrieve this information.

- **Inter arrival times/arrival rates** This information needs to be gathered from the historical arrival data. As mentioned in the section 2.3.1, in literature there are three ways in which arrival rates are deducted from data: direct drawing from the data, statistical testing and predictive methods (Gul & Celik, 2020; Pinto et al., 2015). To be able to model varying arrival rates, it is necessary that analyses are performed that determine if there are factors that impact this arrival rate. This can be done by statistical testing and by predictive methods. The former is chosen, because statistical testing involves analysing the data visually and step-by-step, in stead of using more of a 'black-box' model like an Artifical Neural Network (Hill & Böse, 2017; Höpken et al., 2021). In Chapter 5, these analyses are performed on the data, based on hypotheses from literature and from within the field. By doing this, different scenarios in which the system performs differently, for example between days, hours and weather conditions, can be identified and simulated in the simulation model by drawing from the empirical distribution of the data as mentioned in Chapter 5.

- **Service times** For service times, the same goes as for arrival rates, and in 5 statistical testing on service time data can be found. Important to note is that the service times are different for a call where there is no full triage performed and where there is full triage performed as these are two different subprocesses.

- **Delay time before a call** The out-of-hours system is a system in which people operate. When a triagist is available and a patient is in the queue, the triagist needs to accept the call. This often does not happen immediately, and therefore a delay before acceptance of a call should be modelled in the system.

- **Handling time after call** This time is set at 60 seconds. In the system, it is officially 40 seconds, but during a visit at the department it was identified that many of the triagists never meet this deadline and need more time. Modelling it at 60 seconds will create a more accurate system performance.

- **Composition of subprocesses over demand** For an arriving patient, it should be known what the probability is that it will not receive full triage and will thus be assigned to subprocess 1 and what the probability is that the patient will receive full triage and will thus be assigned to subprocess 2. This probability is derived from the data in Section 5.5.10.

- **Composition of urgency levels over demand** Similarly to the subprocesses, information needs to be available on the chance that this patient has a certain urgency level. This urgency level impacts the service time of the patient, as analyzed in the next Chapter 5. The urgency composition over arriving patients is found in Section 5.5.8.

- **Triagist schedule and capacity** There is a standard triagist schedule for the out-of-hours department. The standard schedule is found in Table 6.3 and slightly altered when validating the waiting times of the simulation model in Section 6.4.2 to accurately display the waiting times, because in the schedule no breaks or other calls and problems at the department are taken into account which are present in reality and in the data. This altered schedule is tested and used for verification and validation of the implemented simulation model in Chapter 6. In Chapter **Interventions chapter**, experiments are run in which the model deviates from the standard schedule to efficiently allocate workforce of an out-of-hours department.

- **Performance norms for waiting times** Certain norms are set by the department for their performance on average waiting times for patients. These are be discussed in 4.2.2.

4.1.3 Boundaries to system

The system to modelled in a simulation model encompasses only the telephone triage system of an out-of-hours department. Whether a person will be visited by a general practitioner, comes to the hospital or only needs advice on the phone is out of the scope of this system. Capacity of hospital rooms and of general practitioners is therefore not considered, healthcare staff means triagists in this case.

4.1.4 Assumptions

Certain assumptions are made about the system, prior to analyzing the system data and implementing the model. These are the following:

- There are no differences in service times between triagists.
- All triagists need on average 60 seconds of handling time after a call.
- Conversations that last below 60 seconds in the data are not taken into account in the data analysis and thus later in the model
- Emergency conversations that last below 2 minutes in the data are not taken into account in the data analysis and thus later in the model
- Only inbound calls are taken into account. Inter-site calls between health professionals or to other healthcare departments are left out of the data.

4.2 SUBQUESTION 1: INDICATORS TO REPRESENT PERFORMANCE OF TELEPHONE TRIAGE SYSTEM

To measure the performance of the telephone triage system within the simulation model it is important to identify relevant indicators and implement the model in such a way that these metrics can be retrieved from it. The most important ones are the waiting times and the norms, which are set by the to measure the adequacy of the healthcare service. In Chapter 6, the mentioned visualizations of these indicators as output from the implemented model are visible.

4.2.1 Waiting times

Waiting time is the most important indicator of performance of the out-of-hours telephone triage system, on which the main research question is based. The waiting times and the amount of people in the line are measured in the model to be able to see the impact of scenario changes or of changes in the system. Also, if waiting times are really high or really low, it is important to be able to retrieve this from the model to see what causes this. It should be noted that in an out-of-hours system, the waiting time begins after an audio fragment is played to the calling patient. In the implemented simulation model in Chapter 6, waiting time measurement starts immediately at the arrival of a patient, so an arrival indicates that the patient already listened to the audio fragment and entered the queue after. Waiting time is therefore measured per patient in minutes and the waiting times of all patients can be visualized, which then displays the average waiting times at all hours of the day with a 95% confidence interval. This is useful when comparing the impact of system changes to normal system performance.

4.2.2 Performance requirements

Next to the waiting times itself, there are three performance requirements, or norms, set for waiting times by the department:

- In 98% of the cases, a call to the emergency line needs to be answered by a triagist within 30 seconds.
- In 75% of the cases, a call to the normal line needs to be answered by a triagist within 2 minutes.
- In 98% of the cases, a call to the normal line needs to be answered by a triagist within 10 minutes.

These are hourly measured in the model by calculating the percentage of patients in that hour were the norm was met. Over a whole simulation of a day, the hourly performance can then be visualized. Also, the performance on the norms can be visualized over other system variables such as triagist capacity or new variables that emerge when implementing system changes. This is useful to compare norm performance over different variable values.

4.2.3 Staff occupancy

The third important indicator is the occupancy of staff. When implementing system changes with the simulation model, it needs to be measured whether the staff that is scheduled is efficiently used. This means that occupancy should be high, but not always at 100% as that indicates that all triagists are handling a call and waiting time is emerging. The occupancy is measured between 0 and 1 in the simulation model - as a share of the triagists - every time it changes. It can be visualized over the whole day, thus indicating when occupancy is high and when low. This is useful when comparing the impact of system changes and different triagist capacities on triagist occupancy.

4.3 CONCLUSION

In this chapter, the telephone triage system and its processes were explained. Next, the different concepts and variables that are needed as input and internal variables in the model were identified, after which the indicators that are needed as output of the model to measure the performance of the system were identified.

By ending this Chapter, subquestion 1 is answered. In the next chapter, subquestion 2 is answered by extensively analyzing the system data. When the model is implemented in Chapter 6, subquestion 3 can be answered in that same chapter, after which the interventions can be identified - subquestion 4 - in Chapter 7.

5 RESULTS: SYSTEM DATA ANALYSIS

As mentioned in the literature review, there are three main ways in research used to predict or estimate arrival rates in healthcare systems and beyond, used as input into discrete event simulation models but also used as just as predictive models alone. To identify the factors that impact arrival rates and service times to be able to use correct scenarios in the simulation model implemented in Chapter 7, the second method mentioned in the literature review will be used: statistical testing. The data is tested on more factors than mostly done in literature: in for example Keshtkar et al. (2015) and in Gul and Celik (2020), a distribution is fitted on big parts of the data, like weeks, days or hourly intervals. For this research, more specific temporal, climatic and other variables are taken into account and the data grouped based on those variables. These groups of data are statistically compared and, if the difference is significant, split into groups (scenarios) of data for demand or service time. It could for example be the case that the demand for healthcare is higher on a very cold day than on a very hot day. An empirical distribution of the demand or service time data in that scenario is computed and can be used as input in the simulation model if predictions need to be made for that specific scenario (the reason why an empirical distribution is chosen can be read in Section 5.9). Analyzing and grouping the data this way combines the clarity and communicability of visualizing grouped parts of the data on a variable and basing a statistical test on that, with the complexity of what more predictive models such as artificial neural networks or regression models can do by integrating more factors than just weekly and hourly patterns.

This means that in this chapter, different hypotheses based on literature and expertise from people working as a triagist or general practitioner* and based on what can be visually hypothesized from plots and graphs, are statistically tested on the historical data from two out-of-hours healthcare department, one of which can be found in Appendix A. In Section 5.6, an interpretation section can be found where the findings of the analyses on the data are discussed and compared to expert opinions of people working at an out-of-hours department and to findings from literature, after which all identified scenarios and conclusions on the hypotheses can be found. Also, conclusions can be found on performing the same data analysis on a second out-of-hours department data-set. Based on the identified scenarios, an analysis is performed on the distributions of the demand and service time data to make a decision between theoretical and empirical distributions for sampling in the model.

*These people shed their light on the found results from the data analysis and gave ideas for hypotheses to test. They work as a triagist, floor manager, data analyst or general practitioner.

5.1 HYPOTHESES

In Table 5.1, hypotheses from literature and from triagists working at an out of hours care department are displayed. More hypotheses will follow from exploration and visualization of the data. In the rest of this chapter, these hypotheses are statistically tested to be able to group the data into significant groups for input into a discrete event simulation model.

Table 5.1: Hypotheses from triagists and from literature

Hypothesis	Source
Temporal variables have impact on arrival pattern	McCarthy et al., 2008, Hamrock et al., 2013, Marcilio et al., 2013
During holidays, the arrival pattern is different	McCarthy et al., 2008
In the 'darker' months, the arrival pattern is different	Triagists first department
When it freezes, it is more busy	Triagists first department
On days after a holiday, it is more busy	McCarthy et al., 2008
The higher someones urgency level, the higher the service time	McCarthy et al., 2008

5.2 AVAILABLE DATA

In Chapter 5, in Figure 4.2 the available data is shown. This data is directly used to analyze the factors that impact the service time and to test the hypotheses with. For the arrival rates, so the demand for healthcare, the data is converted to count data to be able to see the impact of certain factors on hourly arrival rates. An illustration of the count data is displayed in Figure 5.1.

	weekday	season	hour	year	day	holiday	after_holiday	count	temp
0	Friday	Spring	0	2019	2019-04-05	0	0	4	8.2
1	Friday	Spring	1	2019	2019-04-05	0	0	8	8.2
2	Friday	Spring	2	2019	2019-04-05	0	0	7	8.2
3	Friday	Spring	3	2019	2019-04-05	0	0	5	8.2

Figure 5.1: Processes in telephone triage of out-of-hours department

5.3 STATISTICAL TESTS

To test the above-mentioned hypotheses, statistical testing is used to compare different groups of data. When checking whether there is a statistical difference between groups, often the ANOVA test - analysis of variance - is used. This test assumes normally distributed data. In Section 5.9, it can be seen that neither the demand data, which has the form of count data as visible in Figure 5.1 nor the service time data are normally distributed. Because this assumption of the ANOVA test is not met, a similar non-parametric test must be used to compare groups of data for demand and service times in the out-of-hours department. A good option for this is the Kruskal-Wallis test, which does not assume an underlying distribution and is used to compare three or more groups. It compares the ranks of the data points between the groups, rather than comparing the data itself. The Kruskal-Wallis test can be interpreted for a test of differences between medians in groups, when observations are identically distributed between a group (Dransfield, 2021) In the rest of this Chapter, the Kruskal-Wallis test is used to see whether differences occur between groups, like for example between seasonal patterns for demand and service times. The posthoc-Dunn test is used to compare the groups pairwise, to see where exactly the differences between the groups occur. All statistical analyses are performed in Python, using statistical packages *scipy*, *statsmodels* and *scikit*. **p values uitleg**

5.4 SUBQUESTION 2: DEMAND FOR HEALTHCARE

The count data is used to see if differences between groups of data exist. If this difference exists and it is statistically significant, these data groups are separated and will be used as separate demand scenarios when later using it as input into the model.

5.4.1 Yearly differences

First of all, it is checked whether the arrival patterns of the past years are the same and can be aggregated, or are different from each other and some data has to be left out of the analysis. In Figures 5.2 and 5.3, the arrival patterns for the past 4 years with 95% confidence interval for the mean amount of patients that calls in that specific hour of the day are plotted for weekend days and weekdays (opening times from 17:00 (5:00 pm) till 8:00 am), for the normal line and for the emergency line respectively. It is important to note that for now, weekend days and weekdays are separated, but it will be statistically tested whether this is an accurate choice in subsection 5.4.3. In the plots, it is visible that there are some differences between the years, mostly for arrivals around noon. To statistically test whether these differences between the years are significant, the Kruskal-Wallis test is used to compare the distributions of the groups for each separate hour of the day. The hypotheses of this test for yearly differences can be found in Table 5.2.

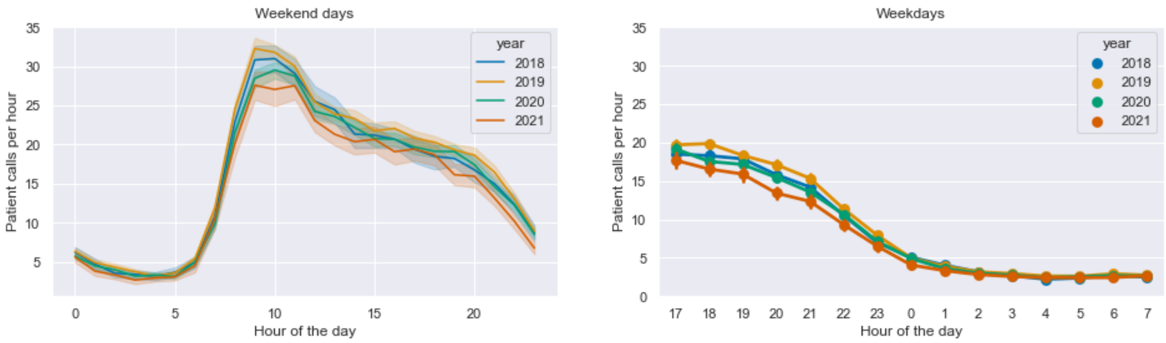

Figure 5.2: Comparison of hourly arrivals between years - normal line

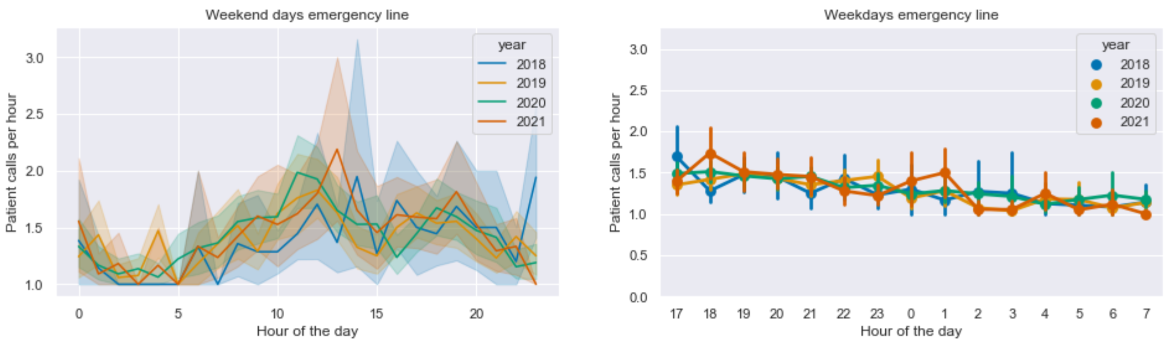

Figure 5.3: Comparison of hourly arrivals between years - emergency line

Important to note is that for the statistical test, the data of 2020 and 2021 have been aggregated, because the data of 2021 is not complete yet and covers only a few months.

Table 5.2: Hypotheses for Kruskal-Wallis test for comparing arrival patterns between years for normal and emergency line

Hypothesis
H0: No differences in distributions between years for hourly arrivals
H1: Differences in distributions between years for hourly arrivals

Therefore, three groups are compared: 2018, 2019, and 2020/2021 together. The results of the Kruskal-Wallis test can be found in Table 5.3. It displays the hours of the day in which differences exist between the years, so where the p-values are below 0.05, and the specific group(s) that are different from the rest.

Table 5.3: Results of Kruskal-Wallis test for comparing arrival patterns between years

Arrival line	Significantly different hours	Different group
Normal line - weekend	1, 3, 7, 8, 9, 10, 16, 18, 20, 21	2019
Normal line - weekday	1, 4, 17-23	2019
Emergency line - weekend	4, 23	2018/2019
Emergency line - weekday	-	-

It can be seen that 2019 is statistically different for some hours for the normal arrival pattern in the weekend and on weekdays, which is a notion that the plots support. For the emergency line, only during the weekend two hours are slightly different, respectively for 2019 in hour 4 and for 2018 in hour 23. For the sake of the completeness of the data, and because it can be expected that some natural variation exists between years, all years are taken into account for further data analysis.

5.4.2 Seasonal differences

In Figures 5.4 and 5.5, similar to the yearly patterns, the arrival patterns are displayed for the four different seasons of a year. The 95% confidence intervals are displayed in the plots as well. In the plots, it is visible that some differences exist for some hours of the day between seasons. Again, it is tested whether differences between the seasons are significant, using the Kruskal-Wallis test. The hypotheses can be found in Table 5.4

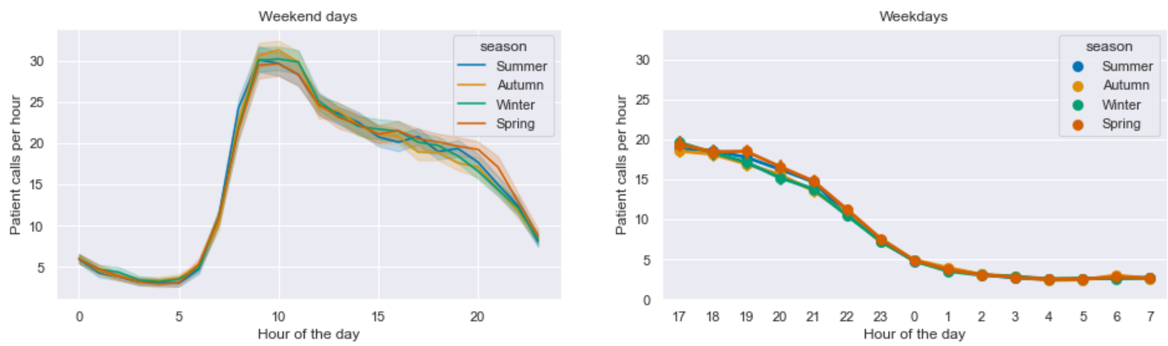

Figure 5.4: Comparison of hourly arrivals between seasons - normal line

The results of the test can be found in Table 5.5. It can be seen that for the normal arrival line, Spring and Winter tend to sometimes have a different pattern than the other seasons. For winter, this only happens for one hour (hour 20 during weekdays), which is why it is chosen to still aggregate winter with summer and autumn for both weekend and weekdays. Spring is different in all other displayed hours, and is therefore taken as a separate arrival

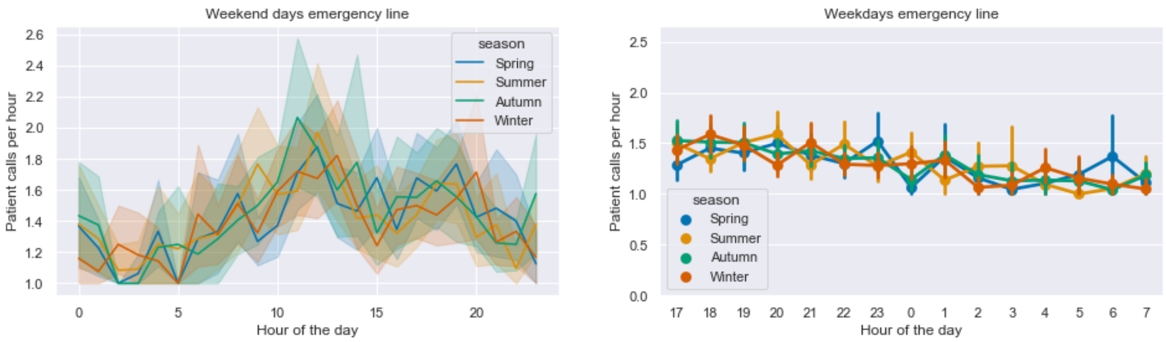

Figure 5.5: Comparison of hourly arrivals between seasons - emergency line

Table 5.4: Hypotheses for Kruskal-Wallis test for comparing arrival patterns between seasons for normal and emergency line

Hypothesis
H0: No differences in distributions between seasons for hourly arrivals
H1: Differences in distributions between seasons for hourly arrivals

pattern. For the emergency line, no differences are seen and therefore all seasons are aggregated.

The found differences are statistically significant, but appear small in the plots. It is chosen to still take these small differences into account as different groups for input into the queueing model, as a difference of a few callers in an hour might lead to big increases in waiting times and the performance of the KPI.

Table 5.5: Results of Kruskal-Wallis test for comparing arrival patterns between seasons

Arrival line	Significantly different hours	Different group
Normal line - weekend	8, 19, 20, 21	Spring
Normal line - weekdays	19, 20, 21	Spring/Winter
Emergency line - weekend	-	-
Emergency line - weekdays	-	-

5.4.3 Weekend days versus weekdays

It can be seen that the patterns of the weekends are by definition different from weekday patterns. The opening hours are different, and it can be seen in Figure 5.6 that for the weekend days, the arrivals increase a lot more rapidly in the mornings than they do for weekdays. This could be due to the regular general practitioners opening at 8 during weekdays, for which people might wait, while in weekends no such options exists and people have to wait till Monday to see their regular general practitioner. Such a pattern can also be seen for emergency arrivals in Figure 5.7, where in the mornings, starting from 5, the arrivals increase more rapidly in the weekends than during weekdays.

Because of these visual differences and because their different opening hours, weekend days and weekdays are seen as different groups for input into the queuing model. In Section 5.6, an expert and literature interpretation is given on this conclusion.

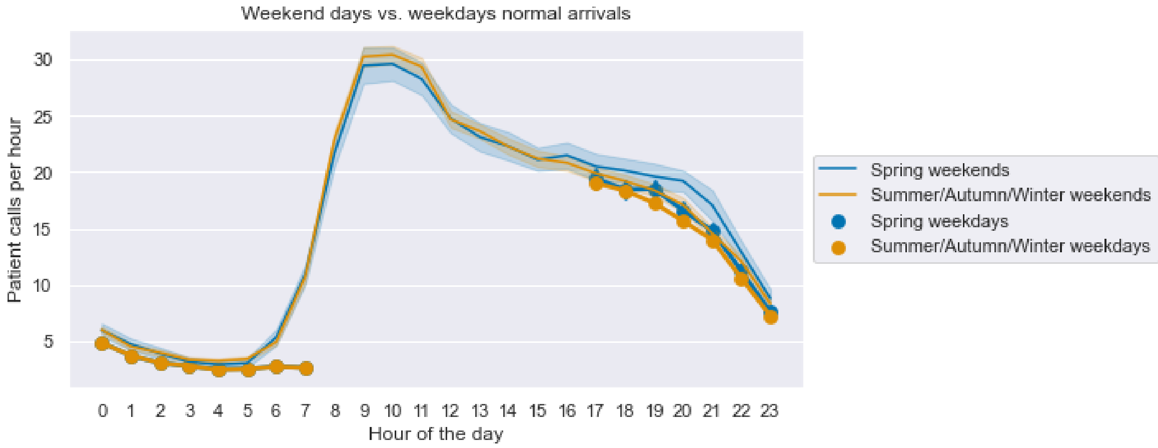

Figure 5.6: Comparison of hourly arrivals between weekdays and weekend days - normal line

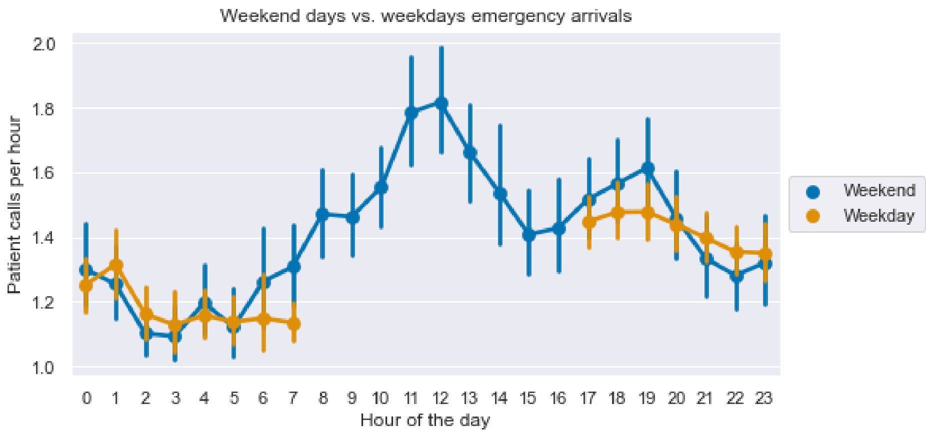

Figure 5.7: Comparison of hourly arrivals between weekdays and weekend days - emergency line

5.4.4 Weekend days, holidays and days after holidays

The daily arrival patterns for weekend days, holidays and days after holidays are displayed in Figures 5.8 and 5.9. The figures both consist of three plots, respectively displaying the difference between arrivals on a Saturday and on a Sunday, the difference between a weekend day and a holiday and the difference between a day after a holiday and a weekend day. Again, it is tested whether statistical differences exist between the plotted arrival patterns. The hypotheses for the Kruskal-Wallis test to hourly compare the amount of arrivals are displayed in Table 5.6.

Table 5.6: Hypotheses for Kruskal-Wallis test for comparing arrival patterns between weekend days, holidays and days after holidays for normal and emergency line

Hypothesis
H0: No differences in distributions between weekend days, between weekend days and holidays and between weekend days and days after holidays for hourly arrivals H1: Differences in distributions between weekend days, between weekend days and holidays and between weekend days and days after holidays for hourly arrivals

The results of the test can be found in Table 5.7. For the normal arrival line, the test is conducted for the data group that consists of the seasons Summer, Autumn and Winter together and for Spring separately as is also visual in the plots and is based on the results

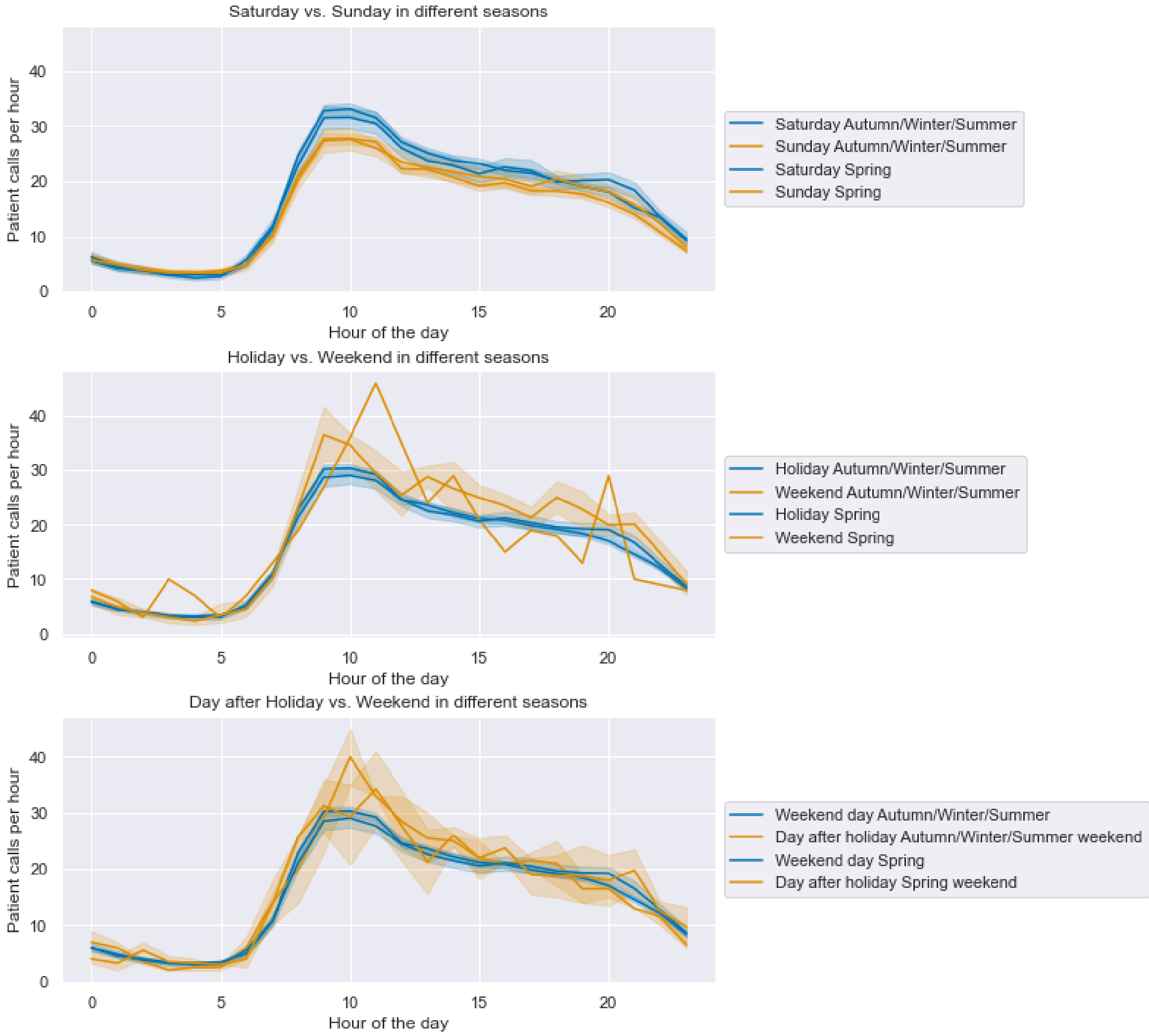

Figure 5.8: Comparison of hourly arrivals between weekend days and (day after) holidays - normal line

of previous testing, which concluded that Spring is different from the three other seasons in the normal arrival line. For Summer, Autumn and Winter it can be concluded that differences exist between Saturdays and Sundays, as is clearly visible in the first plot of Figure 5.8. The differences between weekend days and holidays and days after holidays are not significantly different from a Saturday or a Sunday, and they can thus be aggregated. For Spring, again, Saturday and Sundays are different. Different from the other seasons, in Spring a holiday and a day after a holiday can only be aggregated with Saturdays, and not also with Sundays. In Section 5.6, an expert and literature interpretation is given on this conclusion.

For the emergency line, all seasons are aggregated, so only one group has to be tested for differences between weekend days, holidays and days after holidays. It is found that only in three hours of the day statistically significant differences occur. Days after holiday are only slightly different from holidays on hour 12, and Saturdays and Sundays slightly differ on hour 6 and hour 15. As these are minor differences and to be able to aggregate more data, it is assumed that these hours are the same over the groups, and thus weekends days, holidays and days after holidays are aggregated for the emergency line.

Figure 5.9: Comparison of hourly arrivals between weekend days and (day after) holidays - emergency line

Table 5.7: Results of Kruskal-Wallis test for comparing arrival patterns between weekend days, holidays and days after holidays

Arrival line	Significantly different hours	Different group
Normal line - weekend - Summer/Autumn/Winter	5, 7-23	Saturday/Sunday
Normal line - weekend - Spring	4, 7-13	Saturday/Sunday/Holiday
Emergency line - weekend	6, 12, 15	Day after holiday/Saturday/Sunday

5.4.5 Differences between weekdays

In Figures 5.10 and 5.11, the arrival patterns during weekdays and days after holidays that are on a weekday are visualized for the normal arrival line for the two seasonal groups (Summer, Winter, Autumn together and Spring separate), and for the emergency line. It is again statistically tested whether there are differences between the groups, in this case the weekdays. From the plots, one might hypothesize that Fridays tend to be different from the other days. The hypotheses for the Kruskal-Wallis test can be found in Table 5.8.

The results of the test are visible in Table 5.9. It is visible that for the normal line in both season groups only Fridays are significantly different from the other days. This is also visible in the plots in Figure 5.10, where Friday clearly has a higher trend. This might be due to the weekend coming up in which general practitioners are not opened for another two days and people do not want to walk around with their symptoms all

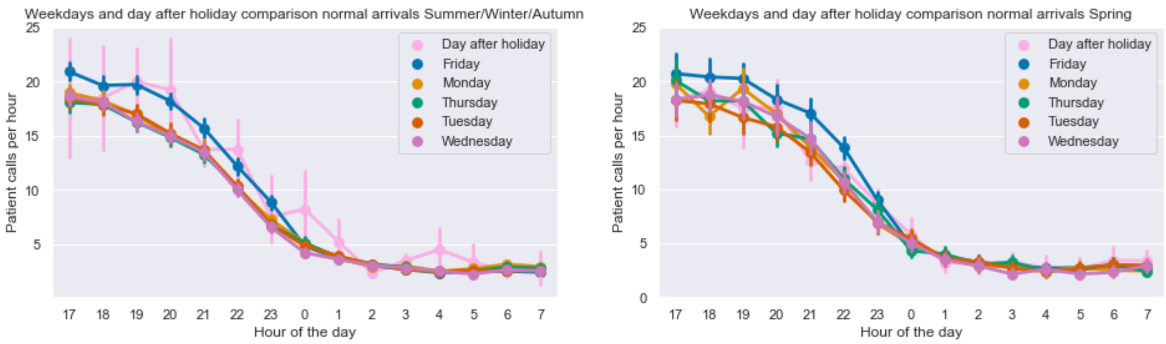

Figure 5.10: Comparison of hourly arrivals weekdays - normal line two season groups

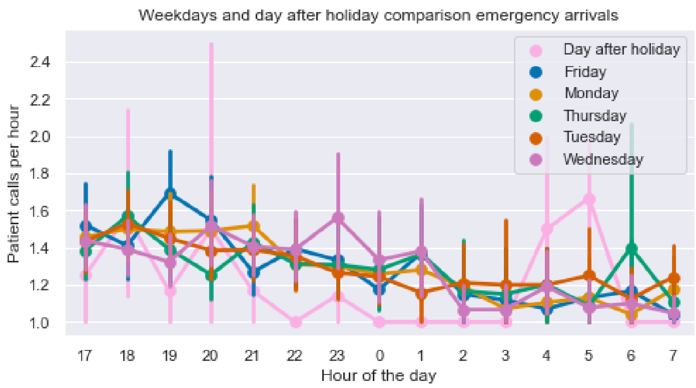

Figure 5.11: Comparison of hourly arrivals weekdays - emergency line

Table 5.8: Hypotheses for Kruskal-Wallis test for comparing arrival patterns between weekdays for normal and emergency line

Hypothesis
H0: No differences in distributions between weekdays for hourly arrivals
H1: Differences in distributions between weekdays for hourly arrivals

weekend. Another reason could be that some general practitioners close earlier on Friday, which leads to more people calling when the out-of-hours departments open. A more elaborate expert and literature interpretation can be found in Section 5.6.

For the emergency line, all days are aggregated for all hours, no differences are significant.

Table 5.9: Results of Kruskal-Wallis test for comparing arrival patterns between weekdays

Arrival line	Significantly different hours	Different group
Normal line - week - Summer/Autumn/Winter	17-23	Friday
Normal line - week - Spring	3, 19-23	Friday
Emergency line - week	-	-

Up and till now, the following groups of data have been found to have significantly different arrival patterns. These groups will be used as demand scenarios in the discrete event simulation model, after the next subsection where it is tested whether there are also significant differences between temperature categories:

Table 5.10: Identified demand scenarios until now

Line	Weekpart	Season	Day
Normal line	Weekend	Summer/Autumn/Winter	Saturday
Normal line	Weekend	Summer/Autumn/Winter	Sunday
Normal line	Weekend	Summer/Autumn/Winter	(Days after) holidays
Normal line	Weekend	Spring	Saturday
Normal line	Weekend	Spring	Sunday
Normal line	Weekend	Spring	(Days after) holidays
Normal line	Week	Summer/Autumn/Winter	Friday
Normal line	Week	Summer/Autumn/Winter	All other weekdays
Normal line	Week	Spring	Friday
Normal line	Week	Spring	All other weekdays
Emergency line	Weekend	All groups	All days
Emergency line	Week	All groups	All days

5.4.6 Differences between temperature categories

As visible in the hypotheses table, Table 5.1, some hypotheses mention temperature as a factor that influences arrivals to medical departments and specifically to an out-of-hours department. In Figures 5.12, 5.13 and 5.14, for all the different groups that have been found up and till now and that are listed just above this section, the pattern is divided into three temperature categories. Then, for each plot, it is checked whether there are statistically significant differences between the arrival patterns for different temperature categories, using the Kruskal-Wallis test. The hypotheses for this test can be found in Table 5.11.

Table 5.11: Hypotheses for Kruskal-Wallis test for comparing arrival patterns between temperature categories for normal and emergency line

Hypothesis
H0: No differences in distributions between temperature categories for hourly arrivals
H1: Differences in distributions between temperature categories for hourly arrivals

For the normal line groups, sometimes a significant difference is found for one or two specific hours of the days between temperature categories. If the difference is only significant for one hour, the data is still aggregated over all temperature categories. For some of the groups, more than one hour was statistically different between the groups, as visible in Table 5.12. This is the case for the Summer/Autumn/Winter group on Sundays and on (days after) holidays, where the Cold category has to be taken separately from the Average and the Hot category and on all weekdays but Fridays, w here the Hot category has to be taken separately from the Average and Cold category.

For the emergency line, no differences are found between temperature categories for the weekends and for weekdays. In the plots, some peaks are visible, especially for hotter days. However, since this happens not very often compared to colder and average temperature days and the data points are very spread out, the difference is not found to be significant.

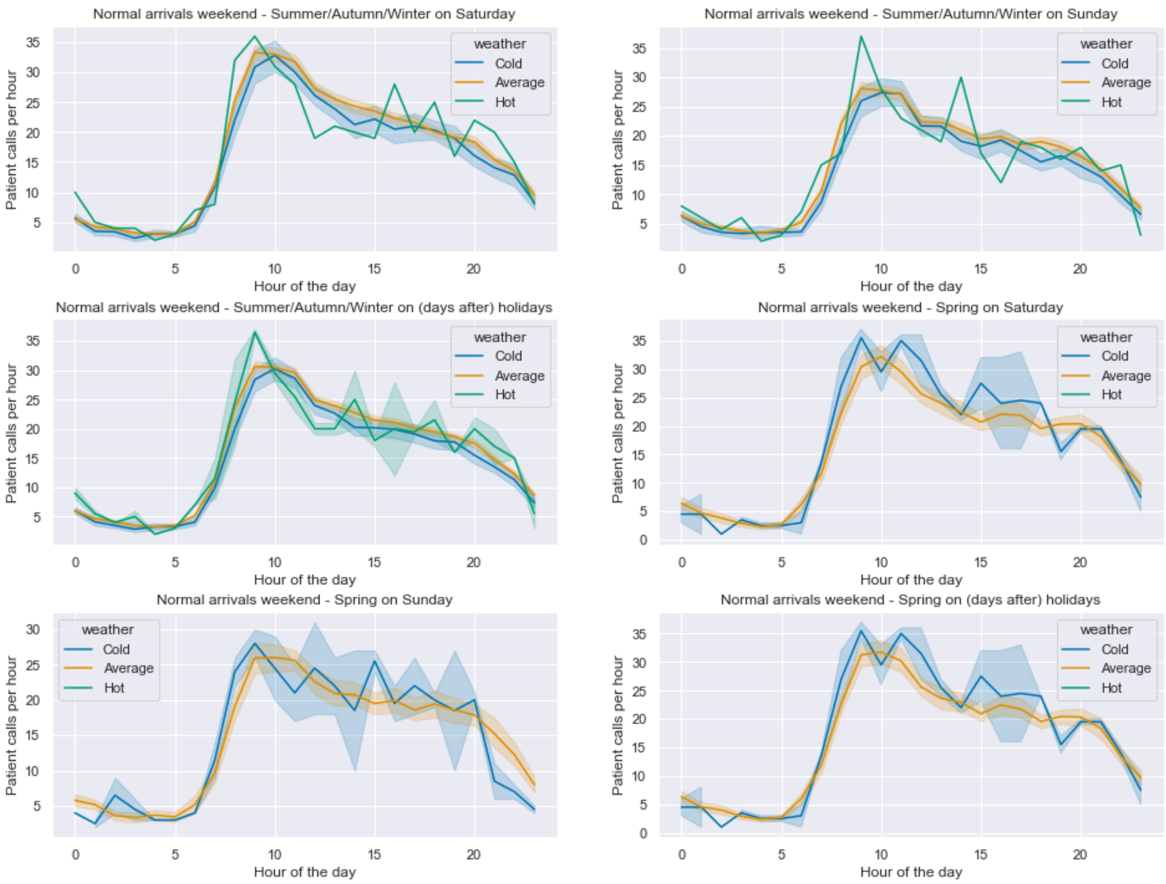

Figure 5.12: Comparison of hourly arrivals between weather types - normal line weekend data groups

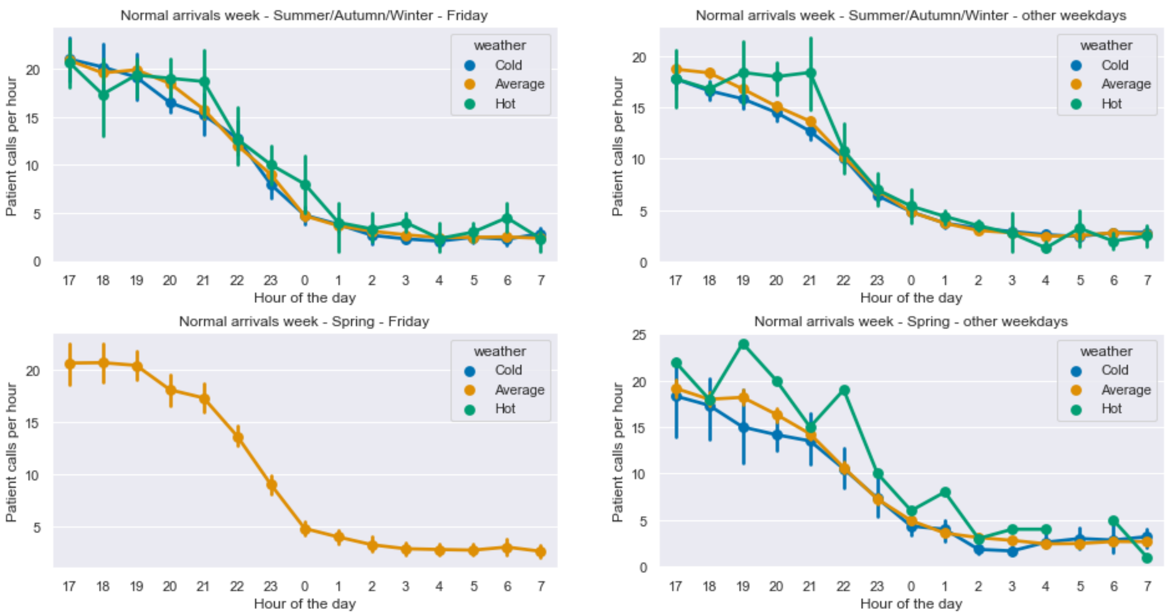

Figure 5.13: Comparison of hourly arrivals between weather types - normal line week data groups

5.5 SUBQUESTION 2: SERVICE TIMES

Similarly to the arrival patterns, the service times of the telephone triage are analyzed to see if temporal, temperature and urgency factors have an impact on service times within the system. If they do, it is important to incorporate this into the simulation model.

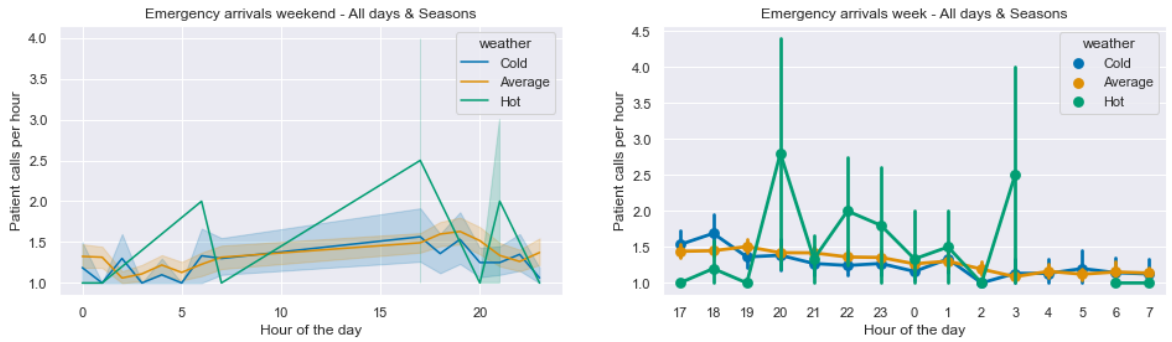

Figure 5.14: Comparison of hourly arrivals between weather types - emergency weekend and week data

Table 5.12: Results of Kruskal-Wallis test for comparing arrival patterns between temperature categories

Arrival line	Significantly different hours	Different group
Normal line - weekend - Summer/Autumn/Winter - Saturday	14	Cold
Normal line - weekend - Summer/Autumn/Winter - Sunday	6, 8, 18	Cold
Normal line - weekend - Summer/Autumn/Winter - (Days after) holidays	6, 8, 14, 23	Cold
Normal line - weekend - Spring - Saturday	-	-
Normal line - weekend - Spring - Sunday	1	Cold
Normal line - weekend - Spring - (Days after) holidays	18	Cold
Normal line - week - Summer/Autumn/Winter - Fridays	-	-
Normal line - week - Summer/Autumn/Winter - All other days	18, 21	Cold/Hot
Normal line - week - Spring - Fridays	-	-
Normal line - week - Spring - All other days	-	-
Emergency line - weekend	-	-
Emergency line - week	-	-

5.5.1 Yearly differences

Before testing for yearly differences in service times in the data, it is checked whether service times in the normal line are different from service times in the emergency line to see whether or not these lines should have separate service times in the simulation model. The patterns are visible in Figure 5.15. It is found that for almost all hours of the day and for every year, the service times are statistically different between the normal and the emergency arrival line. This makes sense, as the nature of a call is very different for the emergency line than for the normal line. Afterwards, it is tested whether there are differences in the service times between the years. This is done because in Figure 5.16, a clear trend upwards is visible for the service times over the past few years. For the simulation model to be valid, the most up to date service times have to be used. The hypotheses for the test can be found in Table 5.13.

Table 5.13: Hypotheses for Kruskal-Wallis test for comparing service times between years for both lines

Hypothesis
H0: No differences in distributions between years for service times
H1: Differences in distributions between years for service times

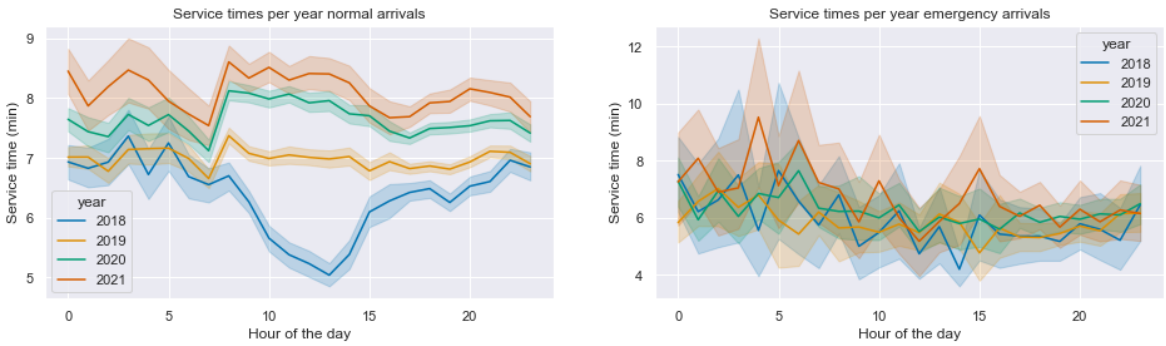

Figure 5.15: Comparison of service time between years

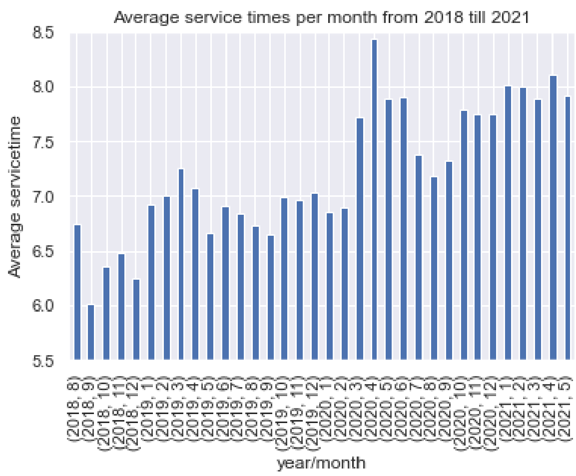

Figure 5.16: Average service times per month, 2018 - 2021

Important to note is that for the statistical test, the data of 2020 and 2021 have been aggregated, because the data of 2021 is not complete yet and covers only a few months. Therefore, three groups are compared: 2018, 2019, and 2020/2021 together. The results of the Kruskal-Wallis test can be found in Table 5.14. It displays the hours of the day in which differences exist between the years, so where the p-values are below 0.05, and the specific group(s) that are different from the rest.

From the test, it can be concluded that only 2020 and 2021 should be taken into account for the normal and emergency line for service times, as these are significantly higher than the service times of the years before. An expert interpretation on the increase in service times over the years can be found in Section 5.6.

Table 5.14: Results of Kruskal-Wallis test for comparing service times between years

Arrival line	Different group
Normal line	2020/2021
Emergency line	2020/2021

5.5.2 Weekend days versus weekdays

It is also tested whether there are differences between weekend days and weekdays when it comes to service times. In Figures 5.17 and 5.18. The hypotheses can be seen in Table 5.15.

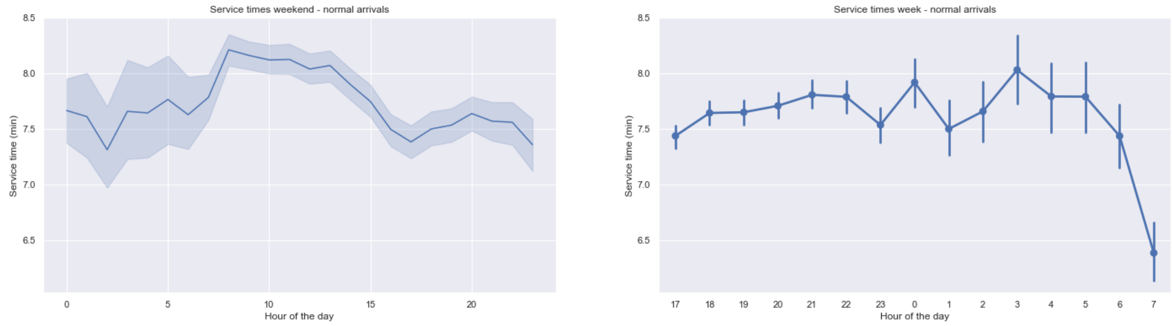

Figure 5.17: Comparison of service time between week and weekend normal line

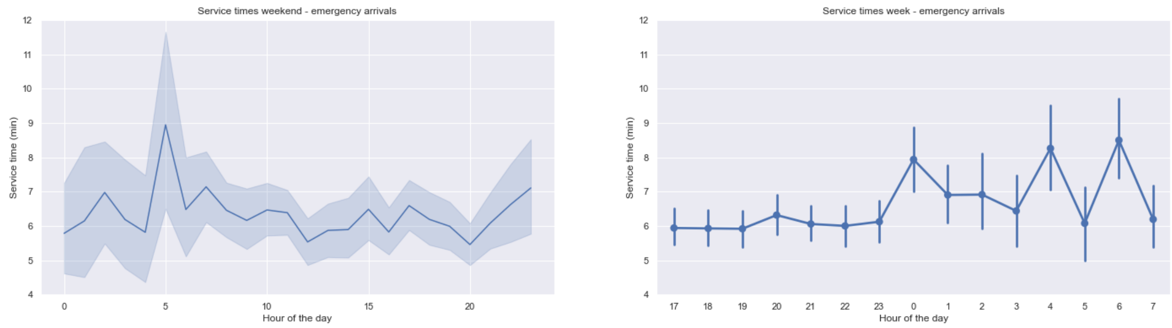

Figure 5.18: Comparison of service time between week and weekend emergency line

Table 5.15: Hypotheses for Kruskal-Wallis test for comparing service times between week and weekend for both lines

Hypothesis
H0: No differences in distributions between week and weekend for service times
H1: Differences in distributions between week and weekend for service times

For the normal line, only a difference can be seen in hour 7 of the day, where service times are lower during the week. This is presumably due to the out-of-department almost closing, and people might be referred to their own general practitioners that open at 8. The choice is made to not make a distinction between week and weekend based on this specific hour, as hour 7 during the week is not where the peak demand in care is visible (see Figure 5.10 for example, where the demand at 7 is either really low for weekdays, or around 10 in weekends)

For the emergency line, in the hours 0, 4 and 17 some differences are visible in service times between weekend and week. Again, it is chosen to not make a distinction between week and weekend as the differences are minor and often, in the emergency line, a few outliers might lead to this difference.

5.5.3 Seasonal differences

in Figure 5.19, the service times are displayed for the four different seasons of the year for both the arrival lines. It is clearly visible that there are some differences between the seasons and the service times, possibly due to different types of calls that are associated with the season and need more or less handling times. Also, it can be seen that service times differ over the course of the day. To verify whether these visible differences are statistically significant, the Kruskal-Wallis test is again performed. The hypotheses are displayed in Table 5.16.

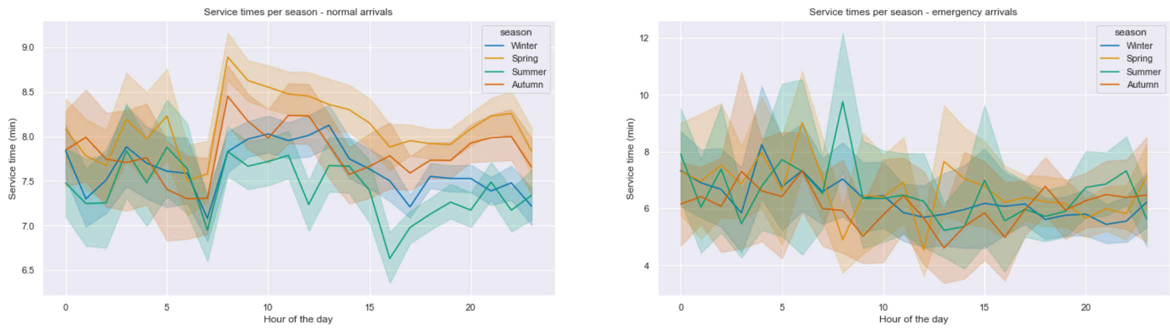

Figure 5.19: Comparison of service time between seasons

Table 5.16: Hypotheses for Kruskal-Wallis test for comparing service times between seasons for both lines

Hypothesis
H0: No differences in distributions between seasons for service times
H1: Differences in distributions between seasons for service times

In Table 5.17, the results of the Kruskal-Test are displayed. For the normal line, almost for all hours there are differences in service times between the seasons. It can be concluded that all seasons have to be taken separately in the simulation model when it comes to service times. For the emergency line, all seasons are similar. For an interpretation on this conclusion, see Section 5.6.

Table 5.17: Results of Kruskal-Wallis test for comparing service times between seasons

Arrival line	**Different group**
Normal line	All seasons
Emergency line	-

5.5.4 Holidays and days after holidays

It is tested whether service times are higher or lower on (days after) holidays compared to normal days. In Figure 5.20, for three of the four seasons for the normal line (there are no holidays in Summer) and for the emergency line the different service times are plotted for (days after) holidays compared to regular days. It is tested whether these service times are significantly different from each other. The hypotheses can be found in Table 5.18.

Table 5.18: Hypotheses for Kruskal-Wallis test for comparing service times between (days after) holidays and regular days for both lines

Hypothesis
H0: No differences in distributions between (days after) holidays and regular days for service times
H1: Differences in distributions between (days after) holidays and regular days for service times

The results of the Kruskal-Wallis test can be found in Table 5.19. For the normal arrival line in Spring, for some of the regular opening hours differences are visible: on days after holidays, conversations tend to take longer than on regular days and on holidays. Days after holidays are therefore taken separately for the normal arrival line in Spring. In Summer, there are no holidays, so no differences can be checked for. In Autumn, no

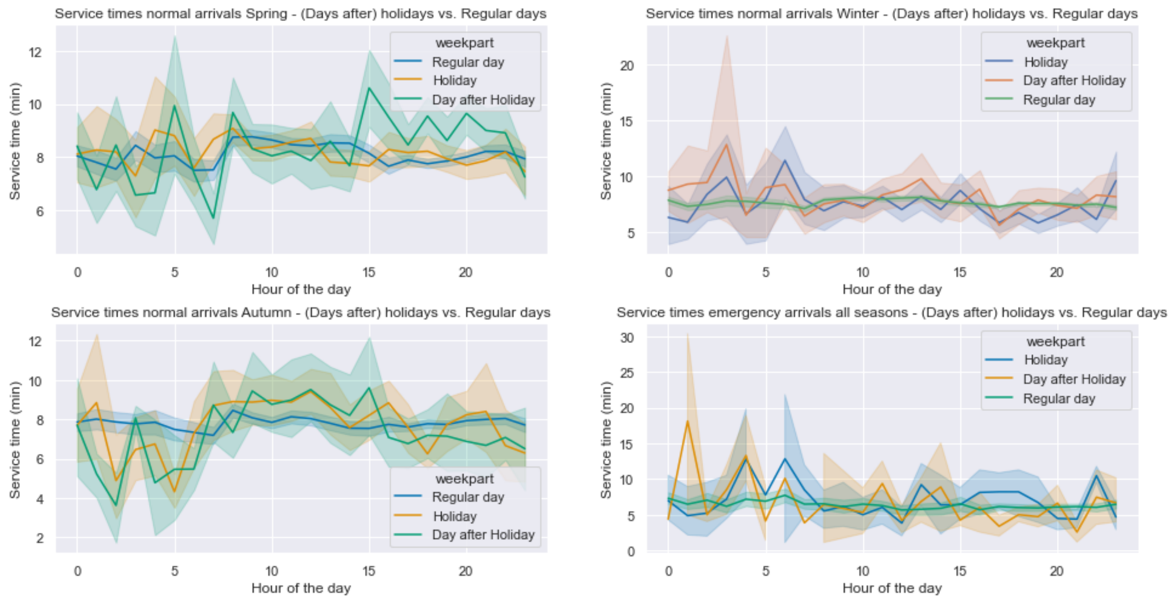

Figure 5.20: Comparison of service time between (days after) holidays and regular days

differences come out of the Kruskal-Wallis test, whereas in Winter and in the emergency line, some differences came out, but these are mostly due to some outlier values on holidays, of which there are not that much and which thus easily manipulate average service times. Therefore, only for Spring as distinction is made between Days after Holidays and other days. For the emergency line,

Table 5.19: Results of Kruskal-Wallis test for comparing service times between (days after) holidays and regular days

Arrival line	Different group
Normal line - Spring	Days after holidays
Normal line - Summer	-
Normal line - Autumn	-
Normal line - Winter	(Days after) holidays
Emergency line - all seasons	(Days after) holidays

5.5.5 Differences between days

Next, the differences between days are checked. In Figures 5.21 and 5.22 it can be seen that no clear pattern can be deduced from the plots, and no specific hypotheses on days that might be different from others when it comes to service times can be made. It is chosen not to statistically check for differences between days, as the variance is high for service times and outliers can happen on any day and presumably have more to do with the urgency level of a patient, than they do with the specific day.

5.5.6 Differences between temperature categories

For all significantly different groups up and till now, the Kruskal-Wallis test is performed to see whether there are significant differences between temperature categories Cold, Average and Hot, similar to the test performed for the arrival patterns. It is found that for the emergency line and for the normal arrival line during Spring, two hours have some differences between the temperature groups, with p-values close to 0.05 but just below. It is

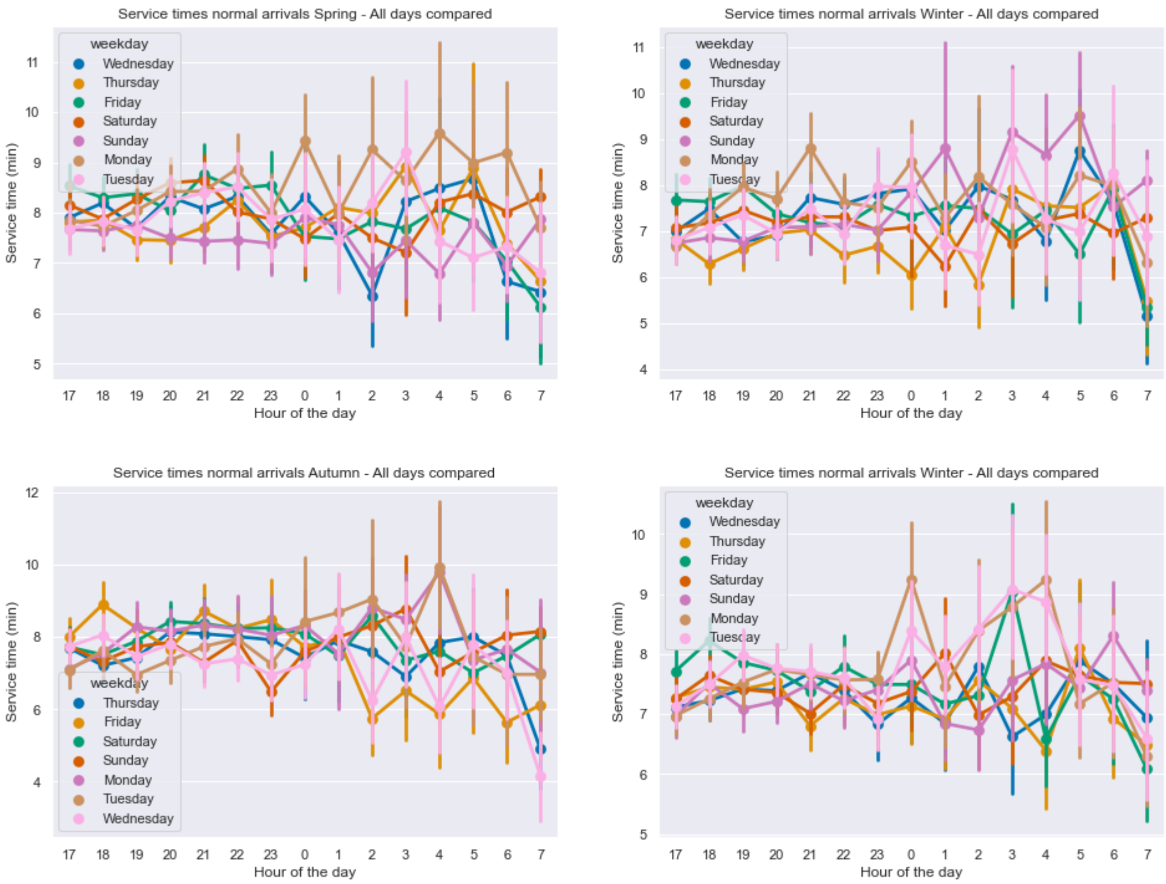

Figure 5.21: Comparison of service time between days

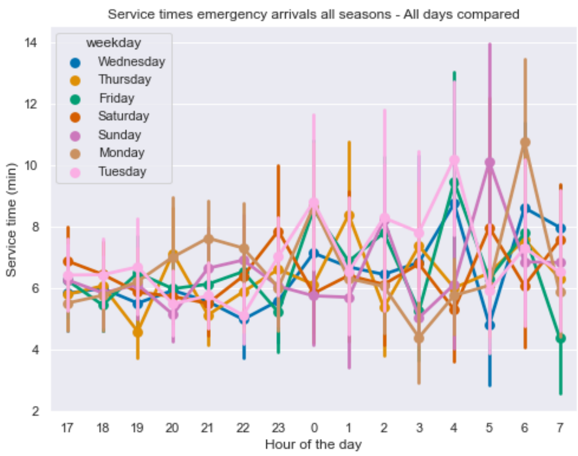

Figure 5.22: Comparison of service time between days

therefore chosen to make no distinction between temperature categories for all groups for service time as again, differences are presumably occurring because of the urgency levels of the calling patients.

5.5.7 Effect of urgency levels

As mentioned a few times in previous sections, the reason why service times might be different between groups of data could be different urgency distributions across the calls. It is checked whether service times are different over different urgency levels. In Figure

5.23, a boxplot with the average service times per urgency level is displayed over the whole dataset. The boxplot shows that the average service times decrease when urgency level decreases (0 is highest, 5 is lowest).

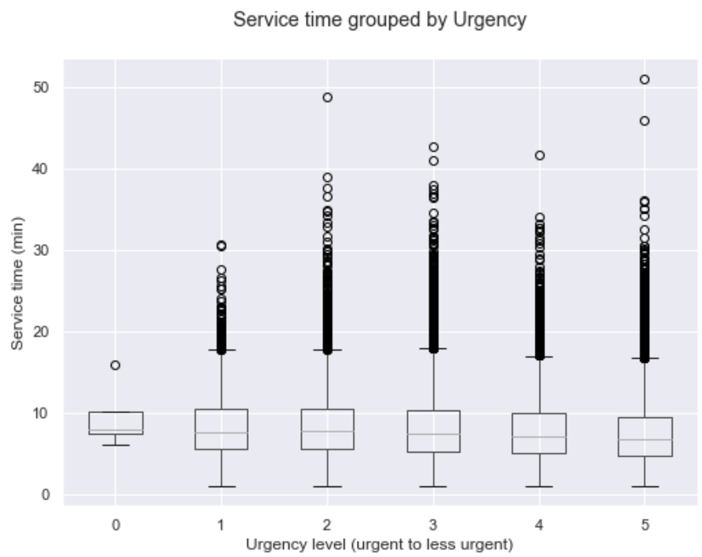

Figure 5.23: Boxplot of service times per urgency level

To confirm that there are differences in service times between urgency levels, the Kruskal-Wallis test is performed on all significantly different data groups for service times. The results are visible in Table 5.20. It can be seen that for most groups, there is a split between the higher and the lower urgency levels. For practical reasons, it is chosen to group the urgency levels for all groups as follows: 0, 1 and 2 form a group, 3 is separate and 4 and 5 are form a group. For the days after holidays group in Spring, no distinction is made.

It is therefore important to know what the urgency of an arriving patient will be, to be able to assign a correct corresponding service time. This is addressed in the next subsection by identifying the probability of an urgency level for a patient over the course of the day. It can also be concluded that because the service times are different per urgency level, the reason why a group is statistically different when it comes to service time (like the seasons) could be due to a high proportion of patients with a specific urgency level. More on this can be read in Section 5.6 when interpreting the found differences between the seasons when it comes to service times.

Table 5.20: Results of Kruskal-Wallis test for comparing service times between urgency levels

Arrival line	Different group
Normal line - Spring	0+1+2, 3, 4+5
Normal line - Spring - Days after holidays	-
Normal line - Summer	0+1+2, 3, 4+5
Normal line - Autumn	0+1+2+3, 4+5
Normal line - Winter	0+1+2, 3, 4+5
Emergency line - all seasons	0+1+2, 3+4+5

5.5.8 Urgency probabilities over the day

In Figure 5.24, of one of the identified demand scenarios the urgency probabilities over the day are displayed. It can be seen that the probability of an urgency level for a patient varies throughout the day. From midnight, so hour 0, it can be seen that the share of higher urgency levels 1, 2 and 3 increases, and the share of lower urgency levels 4 and 5 decreases. At night, the probability that an arriving patient has a high urgency is higher than during the day. In Section 5.6, a literature interpretation of the varying urgency probabilities over the day is given. For all identified scenario's and hours within each scenario, this probability distribution is identified and saved to be used in the model.

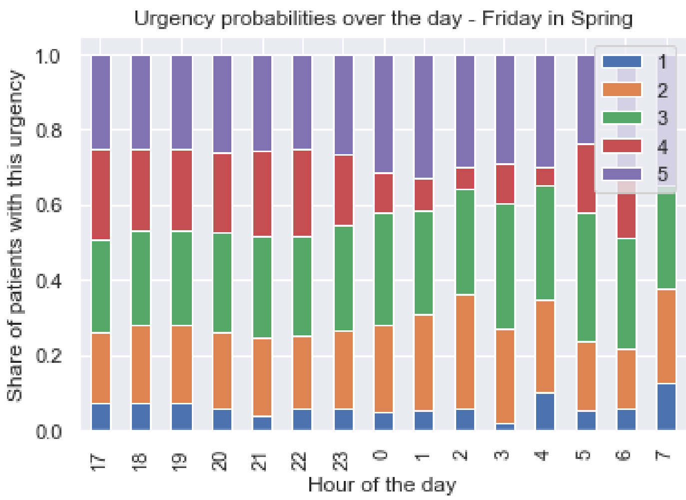

Figure 5.24: Probability per urgency level - Friday in Spring

5.5.9 Aggregate hours for service time

For all identified service time scenarios, it is checked if certain hours of the day can be aggregated when it comes to service time. This would decrease the amount of groups of data and would create interval groups of hours in which the service time can be assumed to be distributed the same. All hours are compared when it comes to service time and compared using the Kruskal-Wallis test. It is found that for a part of the scenario's there are differences between the hours when it comes to service time, and for a part there is no difference between the hours. It is chosen to use the same time interval for each service time scenario for consistent implementation in the model. The three groups of hours that are grouped together for service time are:

1. Hours 7, 8, 9, 10, 11
2. Hours 17, 18, 19, 20, 21, 22
3. All other hours, so 23 - 6 and 12 - 16

An interpretation of these findings can be found in the next section.

5.5.10 Subprocesses: triage or no triage

As mentioned in Chapter 4, a difference exists between the service times of people that receive full triage on the phone and people that do not receive full triage on the phone. It is important to make this distinction in the data, as they are different types of calls and have a different effect on the performance of the system. The distinction between the two subprocesses can be made by two consecutive steps. The first step is to select all the calls that are below 6 minutes in duration, as according to the first department that is the time that a full triage tends to take. Most likely, in the calls that take less than 6 minutes, no full triage was performed. In the next step, from this group of calls below 6 minutes, all calls that have an urgency level are filtered out. These calls did not take 6 minutes, but when they do have an allocated urgency there was still full triage performed. The calls that are left are all calls below 6 minutes that do not have an urgency level allocated. These calls are the calls that belong to subprocess 1 - no full triage. All the other calls belong to subprocess 2 - full triage. A visualization of the distinction between subprocess 1 and 2 is visible in Figure 5.25.

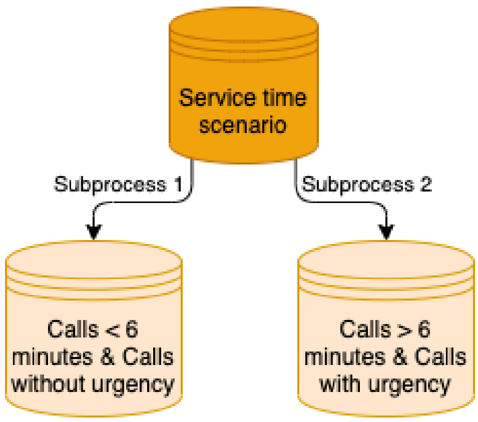

Figure 5.25: Division service time scenario into subprocesses 1 and 2

For each hour of each identified service time scenario (without urgency) the share of patients that goes to each subprocess is calculated. In the model, when a patient arrives, it gets an allocated subprocess according to this share. If it goes into subprocess 1, it gets no urgency level. If it goes into subprocess 2, it gets an urgency level. When the patient reaches the triagist, a servicetime is sampled that matches the subprocess, and for subprocess 2 a servicetime is sampled that also matches the urgency of the patient. In Figure 5.26 it can be seen that the service times for subprocess 1 are significantly lower than the service times for subprocess 2.

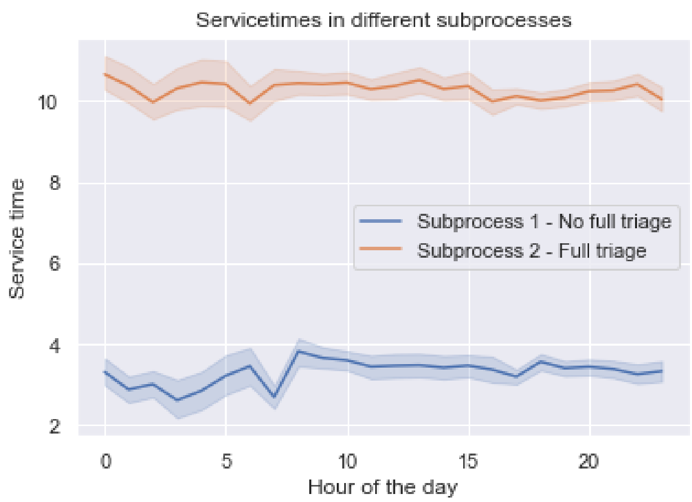

Figure 5.26: Comparison service times subprocess 1 and subprocess 2

5.6 EXPERT AND LITERATURE INTERPRETATION

Based on the analysis of the demand and the service time from the historical data, several patterns and differences stand out. For these interesting conclusions from the analysis, experts working at the first out-of-hours department were consulted. Also, most of the conclusions are discussed in the light of the literature review and of new literature that shines light on similar conclusions in healthcare systems. Sometimes no experts were consulted because conclusions were drawn directly from data the triagists entered. Also, for service times, sometimes a statistical interpretation of a conclusion is given.

First, the conclusions from the demand data are discussed, then from the service time data.

5.6.1 Demand patterns

Week versus weekend day
In Section 5.4.3, it is found that weekdays and weekend days have different arrival patterns. On the one hand because of the different opening hours, 24 hours in the weekends and only on out-of-office hours during the week, but also because of higher peaks in the weekend. In the weekend, in the mornings the amount of calls increases rapidly, whereas during the weekends this does not happen. Also, during the evenings, a little more calls are expected in the weekends.

Expert interpretation
These findings were discussed with people working at the first department. The reason for a rapid increase in calls during the weekend has three reasons: firstly, in the weekends there are morning rounds in nursing homes in which often medical problems are identified. As the general practitioners are not open in the weekends, the out-of-hours department is called. The second reason is the fact that during the week, people wait for their general practitioner to open at 8 AM. In the weekends, the general practitioners will not open at all, so people call earlier in the morning. The last reason is the fact that as the general practitioners will not open again until Monday, the barrier for people to call is lower because if they don't call, they might have to deal with their symptoms all weekend.

Literature interpretation
In literature, as reviewed in Chapter 2, an often mentioned reason for people to call the out-of-hours department is a lack of accessibility to the general practitioner (Keizer et al., 2021). The fact that it is more busy earlier on the day in the weekends supports this claim: the general practitioner cannot be contacted, so people more easily choose to call the out-of-hours department. The claim can also be turned around: during the week, it is less busy in the morning, **because** the general practitioner is reachable from 8 AM. This stops people from calling earlier. It can be concluded that the found differences between week and weekend days in the data are supported by experts and by literature findings.

Saturday versus Sunday and holidays
In Section 5.4.4, it is found that Saturdays are more busy than Sundays. The differences occur from approximately 8 in the morning, and stay present until night time. Also, holidays and days after holidays tend to have a slightly different arrival pattern, in such a way that they cannot be generalized with just a Saturday or a Sunday.

Expert interpretation
Again, these findings were discussed with people working at the department. According to them, the higher demand on Saturdays compared to Sundays is due to, similar to the difference between weekdays and weekend days, the fact that the general practitioner will not open again till Monday. If a health problem occurs on Sunday, more people will wait until Monday morning to call their general practitioner in stead of calling the out-of-hours department. On Saturdays, calling the general practitioner is at least a day and a night away, so more people will call the out-of-hours department with their problem. Another reason is the sports competition of multiple sports that happens mostly on Saturdays, which results in injuries that the out-of-hours department is often called for.

Literature interpretation
The same base in literature can be found for the difference between Saturdays and Sundays as for the difference between weekend days and weekdays, as previously discussed. The accessibility of the general practitioner is the main reason why people tend to call the out-of-hours department more on Saturdays than on Sundays, where they have to wait less time for the general practitioner to be reachable again. The conclusion that holidays and days after holidays are slightly different than normal weekend days is also found in McCarthy et al. (2008), on which the hypothesis was based that holidays and days after holidays are different from a normal day.

Friday versus other weekdays
In Section 5.4.5, it is found that Fridays are more busy than the other weekdays. This trend continues over the course of the evening, and in the night the difference disappears.

Expert interpretation
The increase in demand on Fridays is a known phenomenon for the triagists at the department. They mention that a reason for this is that on Fridays some of the general practitioners close earlier in the afternoon, which leads to people waiting for the out-of-hours department to open at 5 PM. This creates an immediate peak compared to the other weekdays. Similar to the previously discussed findings, on Fridays people know that the general practitioner won't open again until Monday, so the barrier to just call is lower than on a weekday, where you can contact your general practitioner the next morning again.

Literature interpretation
In literature, analyses have been performed on the factors that impact demand in healthcare systems such as emergency rooms or other hospital departments. These analyses come to the conclusion that in an emergency room, the peak in demand is found early in the week and highest on Mondays, and low in the weekends (McCarthy et al., 2008; Walker, Van Woerden, Kiparoglou, & Yang, 2016). This is contrary to what is found to be the trend in this out-of-hours department. This difference can be attributed to the definition of an emergency department versus an out-of-hours department: an out-of-hours department can be seen as an 'emergency general practitioner' which can direct a patient to the emergency room when necessary, whereas an emergency department only handles highly urgent patients that would most likely not call their general practitioner first for their type of problem. At the beginning of the week, people with the types of health issues that an out-of-hours department would handle can call their general practitioners and maybe even make an appointment for the same week and a peak in demand on Monday would therefore not be logical behavior of the system.

Weather influence
It was found that for some of the demand scenarios, the weather has an impact, whereas for some it doesn't. It can be seen when looking at Table 5.21 that on days that are found to be busier than others, there is no difference in demand patterns between temperature groups. On the days that are less busy, for example Sundays or all days of the week except for Friday, the temperature does have an effect.

Expert interpretation
When looking at the hypotheses in Table 5.1, it can be seen that the triagists expect a different arrival pattern during the darker months and that they feel like it is more busy when it is freezing outside. The found demand scenarios confirm their experiences when it comes to cold weather in the Winter as it is found that a cold day behaves slightly different from a normal day. It is not necessarily busier, but the variance in the demand is a lot higher on colder days (see Figures 5.13 and 5.12), which leads to the perception that these days are busier because some of them definitely are, and it is likely that they are easily remembered because the weather was so specific. What they also mention, is that it is rather the types of issues that people call with are varying over temperature categories than the demand.

Literature interpretation
In McCarthy et al. (2008), no impact of temperature is found on the arrivals at an emergency department at an emergency department, whereas in He, Hou, Toloo, Patrick, and Gerald (2011) it is found that more extreme seasonal influences, such as a heat wave, can have an impact on emergency department arrivals. This corresponds to the conclusion from the data analysis that a very hot or cold day can have a different arrival pattern than a regular day. Again, it should be noted that these findings from literature are not done in an out-of-hours department, and that therefore the differences in the conclusions can be attributed to the difference in the use and goal of the department.

Urgency distribution
In Section 5.5.8, in Figure 5.24, it can be seen that the distribution of urgency over the patients varies throughout the day. At night, the probability that an arriving patient has a high urgency is higher than during the day.

Literature interpretation
In McCarthy et al. (2008); Welch, Jones, and Allen (2007), it is found that the acuity (urgency) of patients varies over the day. During the night, a larger proportion of the patients treated at the emergency department has a high urgency compared to afternoon and evenings. This is in line with the conclusion drawn from the data analysis.

5.6.2 Service time patterns

Yearly differences
In Section 5.5.1 in Figure 5.15, it can be seen that since 2018, the average service times have been increasing, in such a way that the past to years have been very different from the years 2018 and 2019.

Expert interpretation
A yearly increase of the service time of a phone call is not a self-evident trend. The department mentions multiple reasons why this increase could have happened. The first one is

that the focus of the conversations has been on help through the phone, sometimes with video, to avoid real life visits at the department. This is in fact an increase in performance of the system as a whole, but a decrease of only the telephone triage system as service times have been increasing because of this focus. The next reason is that service times are different per triagist. New triagists tend to have a longer service time, so the increase could partly be due to a larger share of new triagists in the past years. Lastly, more criteria are added by the year to the triage process which increases service times.

Literature interpretation
The differences in service times between healthcare employees, one of the reasons for an increased service time mentioned by the department because new triagists tend to have longer service times, is a difference that is also mentioned in literature. Another interesting notion is that the average service time in 2021 - around 8 minutes - meets the 7.78 minutes on average that is found in literature for an out-of-hours department in England (Mohammed et al., 2012).

Urgency level
In Section 5.5.7, it can be seen that differences exist in service times between urgency level. The higher urgency levels 1,2 and 3 can be grouped when it comes to service times and the lower urgency levels 4 and 5 can also be grouped. The urgency levels can be grouped similarly for all service time scenarios.

Expert interpretation
For people from the department, it is a recognizable conclusion that the service times are different across urgency levels. They mention that more urgent calls, like when a resuscitation has to be guided through the phone or when a psychiatric patient calls, tend to take longer. This is different from someone who needs brief advise on a small problem.

Literature interpretation
Similar to the department, who mention that psychiatric patients need longer service times than other types of patients, Mohammed et al. (2012) mentions a difference between the length of a mental health or a non-mental health call. The found differences from the data in service times between urgency levels are therefore explained in the same way by on the one hand the department, and on the other hand the literature.

Seasonal differences
When it comes to service times, all seasons are taken as a separate scenario from the data. As opposed to demand, it does not seem self-evident that service times differ across seasons. An explanation might be found in the urgency probabilities between the seasons, because higher urgency patients take more time to be helped on the phone (see Section 5.5.7.

Statistical interpretation
When looking at the urgency probabilities within the demand scenarios, it can be seen that a slight difference exists between the Spring scenarios and the scenarios with the other three seasons combined. The share of Urgency 1 and Urgency 2 patients is higher in Spring in the peak hours of the day, when the peak in service times is also visible in Figure 5.19. This could partly explain why service times might be different across seasons.

Literature interpretation
In literature, no previous research has been performed on temporal influences on service times in out-of-hours care or other types of healthcare systems. Papers that do address temporal influences look at its effect on length of stay at a department or demand and thereby focus more on the underlying problem of a patient that could be due to the season.

5.7 IDENTIFIED SCENARIOS

After interpreting the subconclusions of the data analysis, it can be concluded that temporal and temperature variables, subprocesses and urgency levels have an effect on demand for out-of-hours care and on the service times within. For demand for care, the significant groups are listed in Table 5.21. For the service times of a call, the significant groups are listed in Table 5.22.

Table 5.21: Identified demand scenarios

Line	Weekpart	Season	Day	Weather
Normal line	Weekend	Summer/Autumn/Winter	Saturday	All temperature categories
Normal line	Weekend	Summer/Autumn/Winter	Sunday	Cold
Normal line	Weekend	Summer/Autumn/Winter	Sunday	Average/Hot
Normal line	Weekend	Summer/Autumn/Winter	(Days after) holidays	Cold
Normal line	Weekend	Summer/Autumn/Winter	(Days after) holidays	Average/Hot
Normal line	Weekend	Spring	Saturday	All temperature categories
Normal line	Weekend	Spring	Sunday	All temperature categories
Normal line	Weekend	Spring	(Days after) holidays	All temperature categories
Normal line	Week	Summer/Autumn/Winter	Friday	All temperature categories
Normal line	Week	Summer/Autumn/Winter	All other weekdays	Cold
Normal line	Week	Summer/Autumn/Winter	All other weekdays	Average/Hot
Normal line	Week	Spring	Friday	All temperature categories
Normal line	Week	Spring	All other weekdays	All temperature categories
Emergency line	Weekend	All groups	All days	All temperature categories
Emergency line	Week	All groups	All days	All temperature categories

Table 5.22: Identified service time scenarios

Line	Season	Urgency
Normal line	Spring	Urgency 0, 1, 2
Normal line	Spring	Urgency 3
Normal line	Spring	Urgency 4, 5
Normal line	Spring - Days after holidays	All urgency levels
Normal line	Summer	Urgency 0, 1, 2
Normal line	Summer	Urgency 3
Normal line	Summer	Urgency 4, 5
Normal line	Autumn	Urgency 0, 1, 2
Normal line	Autumn	Urgency 3
Normal line	Autumn	Urgency 4, 5
Normal line	Winter	Urgency 0, 1, 2
Normal line	Winter	Urgency 3
Normal line	Winter	Urgency 4, 5
Emergency line	All seasons	Urgency 0, 1, 2
Emergency line	All seasons	Urgency 3
Emergency line	All seasons	Urgency 4, 5

In Appendix A, the same data analysis and scenario identification as performed in this Chapter is performed for the second available data-set from a second out-of-hours department. It is visible that many of the patterns that are visible for the data of the first department are similar to the patterns in the second department data: demand profiles have a similar shape and are only scaled a bit differently due to a different amount of patients that the second department serves, and it can be seen that many of the interesting results as discussed in Section 5.6 are also visible for the second department, such as the difference between weekends and weekdays, the difference between Saturdays and Sundays, the difference between Fridays and other weekdays, and the difference between mornings during the week and mornings during the weekend. The notion that the patterns in the data are similar is an indication that interventions that might in the end lead to waiting times reductions - which are identified in Chapter 7 - have a good chance of having that same effect other departments.

5.8 CONCLUSION SYSTEM DATA ANALYSIS

When predicting and modelling demand and service times in out-of-hours healthcare, one should definitely consider temporal, temperature and urgency factors. Demand and service times do not only vary over hours of the day, but also over days in the week, seasons in the year and over the years. Also, for the systematically busier days, temperature variations do not seem to make much of an impact while they do for the less busy days. Lastly, urgency levels impact the time it takes to help a calling patient. Looking back at the hypotheses made in Table 5.23, it can be concluded that all hypotheses on temporal and temperature variables are true for some parts of the data. The hypothesis that states that service times will be higher with higher urgency levels is also true. Many of the notions seen in literature for different types of healthcare systems and notions made by the triagists can be confirmed.

Table 5.23: Hypotheses from triagists and from literature

Hypothesis	Source	Conclusion
Temporal variables have an impact on arrival pattern	Hamrock et al. (2013); Marcilio et al. (2013); McCarthy et al. (2008)	Confirmed
During holidays, the arrival pattern is different	McCarthy et al. (2008)	(Partly) confirmed
In the 'darker' months, the arrival pattern is different	Triagists first department	(Partly) confirmed
When it freezes, it is more busy	Triagists first department	(Partly) confirmed
On days after a holiday, it is more busy	McCarthy et al. (2008)	(Partly) confirmed
The higher someones urgency level, the higher the service time	McCarthy et al. (2008); Mohammed et al. (2012)	Confirmed

5.9 DATA DISTRIBUTIONS AND DATA SAMPLING

To be able to use the identified scenarios, it is necessary to use the distributions of the data from these scenarios to draw samples from within the simulation for each calling patient.

For this purpose, a theoretical distribution can be used or the empirical distribution of the data can be computed. In literature, certain theoretical distributions are proposed for different processes. For arrival processes, often a Poisson process is assumed where the inter arrival times follow an exponential distribution (Tiwari et al., 2016). For service time processes in systems like a call center, often a good fit to a lognormal distribution is found (Gualandi & Toscani, 2018). In this chapter, it is tested whether these notions from literature fit the data of the out-of-hours department and a decision is made between a fitted theoretical distribution or the empirical distribution of the data. Doing this accurately results in a model that can be verified and validated after implementation in Chapter 6.

To have a good image of what data we are dealing with, multiple goodness of fit tests are compared and used. When testing if a sample comes from a certain theoretical distribution, there are three common procedures to follow (Razali & Wah. Y.B., 2011), from which the first and the last are used in this chapter:

- Graphical methods such as histograms, boxplots and QQ-plots
- Numerical methods such as skewness and kurtosis indices
- Formal tests such as the Shapiro-Wilk test for normality, the Kolmogorov-Smirnov test, the Anderson-Darling test and the chi-square test for discrete distributions.

Firstly, graphical methods are used to assess which distribution visually fits best. To thoroughly do this, it should be done hourly for each of the significantly different groups of data. Next, some statistical goodness-of-fit tests are performed and its results and limitations are discussed.

5.9.1 Distribution Inter Arrival Times and Arrival Rate

Hourly plots and visualizations were made to visually assess if the data fits a theoretical distribution as proposed by literature accurately. To give an image of the distribution of all the data, on which later distribution fittings were based, the histogram and a histogram where multiple distributions are fitted for inter arrival times are visible in Figures 5.27 and 5.28.

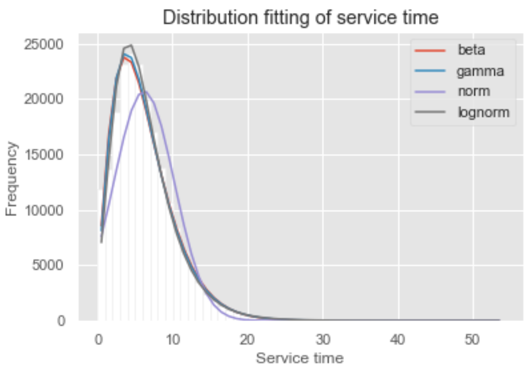

Figure 5.27: Distribution fitting on all data inter arrival times

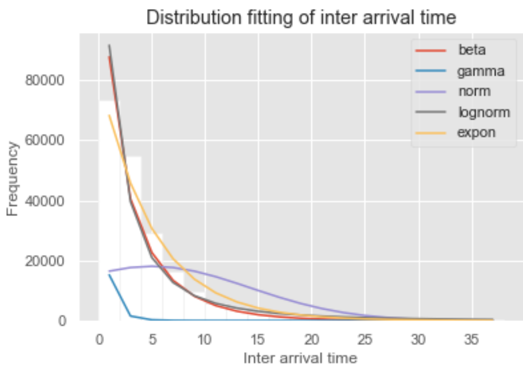

Figure 5.28: Distribution fitting on all data inter arrival times

GRAPHICAL INSPECTION In the plots, it can be seen that visually, an exponential distribution fits the data rather well compared to the other distributions, which suits the definition of a Poisson process (Tiwari et al., 2016). Based on these plots, probability plots were computed hourly for all data groups. As the data used for analysis of the groups is hourly arrival count data, in stead of the exponential distribution - which would be suited for inter arrival times - the Poisson distribution is used for fitting. As there are 15 significant demand groups of which 6 are weekday arrival patterns, this means that 10*24 + 6*15 = 330 plots were computed to see whether any of them visually deviates from a Poisson distribution. Two probability plots are displayed in Figures 5.29 and 5.30: the first one shows a good fit with many samples, the second one also shows a good fit, but there were not that many data points available. This can be the case for example at night, when sometimes nobody arrives. It can also be seen that because the data is discrete as it is count data, the points lay exactly on the lines of the plot. In this case, a probability plot is used because of this discreteness: in Python it is not possible to fit discrete distributions with a statistical package which is necessary for the QQ-plot. The probability plot accepts parameter input for lambda for a Poisson distribution, which is possible with the means of the available data. The only difference with the QQ-plot is that the scaling is based on the theoretical distribution, rather than on the data for the QQ-plot.

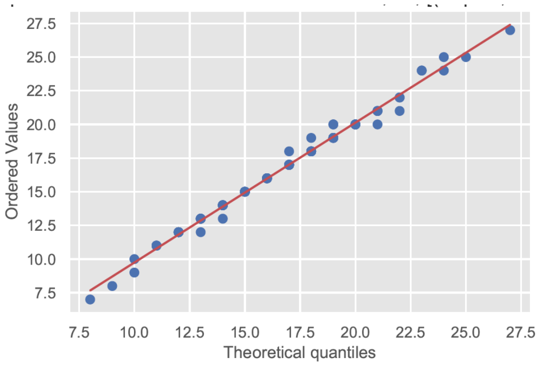

Figure 5.29: Probability plot with good fit

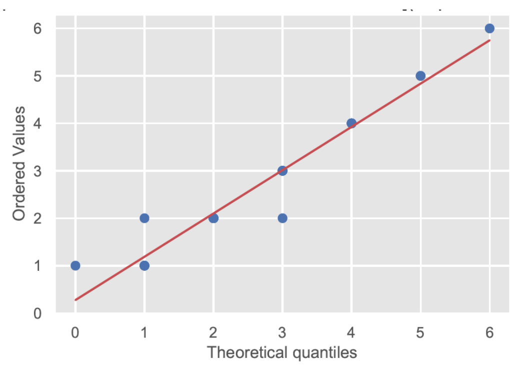

Figure 5.30: Probability plot with good fit - small sample

CHI-SQUARE GOODNESS-OF-FIT TEST Because the count data is a discrete variable, the chi-square goodness-of-fit test is used to see whether the sample data follows a specified distribution (null hypothesis), and to see whether the conclusions drawn from the graphical inspection can also be drawn based on a statistical test. The normalized observed frequencies of the count data are compared to expected frequencies of a Poisson distribution, using the probability mass function, with as lambda the mean of the data sample. The test is, similar to the plot, performed for all hours of the different demand scenario groups. For all hours in all scenarios, the p-value is higher than 0.05, which indicates that it is likely that the sample data follows a Poisson distribution.

KOLMOGOROV-SMIRNOV AND ANDERSON-DARLING TEST To ensure that the arrival process really follows a Poisson process, it is also tested whether all separate hours of the data follow an exponential distribution. It is found that for 95% of the hours in the data, the inter arrival times follows an exponential distribution according to the Kolmogorov-Smirnov test, and even more, around 98% according to the Anderson-Darling test.

CONCLUSION Based on the visual inspection of the probability plots and the histograms and based on the chi-square goodness-of-fit test, it can be concluded that all hours within all demand scenarios follow a Poisson distribution rather well over the whole range of the data. When running tests on the inter arrival times of all hours in the data, it can be concluded that based on these tests, 95 to 98% of the hours follow an exponential distribution. Based on these statistical tests, the arrival process of patients can be seen as a Poisson process as seen often in literature for similar systems and processes. In the simulation model, implemented in the next chapter, it is chosen to use empirical distributions derived from the data for inter arrival times, in stead of the fitted exponential distribution. This is to account for outlying values in the data that are not captured when fitting distributions, which in this system have a large impact on performance metrics.

5.9.2 Distribution Service Times

To give an image of the distribution of the whole data, on which later distribution fittings were based, the histogram and QQ-plot for service times are visible in Figures 5.31, 5.33

and 5.32. Important to note is that for the distribution fitting to be accurate, all service times have been decreased by one minute, as that was the threshold for a call to be included in the data. For the emergency data, this threshold amount was 2 minutes, as many calls lasted just a bit more than one minute which indicates that these were not real emergency calls, but were stopped after a bit more than a minute to be referred to the normal line. By doing this, the data starts at zero which will increase the fits of the distributions that often start at 0. If it chosen to use the theoretical distributions in the simulation model, this will be accounted for. If empirical distributions are used, there is no need for decreasing values for a better fit.

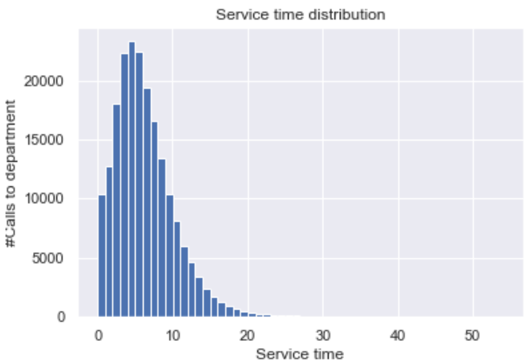

Figure 5.31: Histogram of service times all data

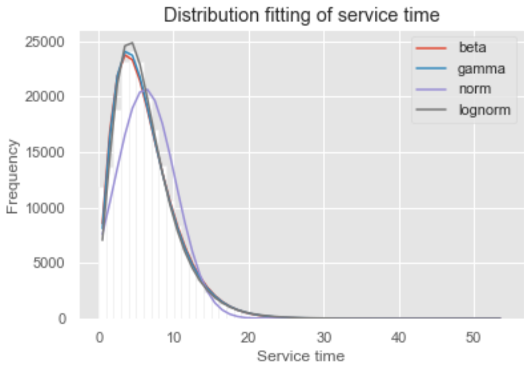

Figure 5.32: Distribution fitting on all data service times

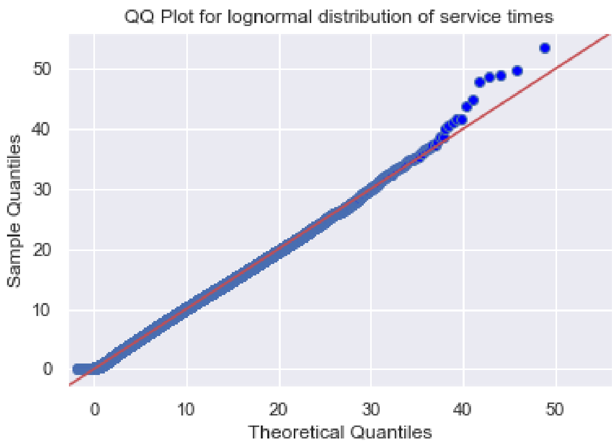

Figure 5.33: QQ-plot service times all data for lognormal distribution

GRAPHICAL INSPECTION In the plots, it can be seen that visually, a lognormal distribution fits the data rather well compared to the other distributions, which is in line with what is often found in literature when the distribution of service times in call center systems is analyzed (Gualandi & Toscani, 2018). However, in the tails, some outliers are visible on the upper side of the 45-degree line, which means that the fitted distribution will sample lower values for these outliers. These histograms and QQ-plots were computed hourly for all data groups to see whether the outliers in the tail would also be present there and in the same way for the lognormal distribution. As there are 16 significant service time groups, this means that 2*16*24 = 768 plots were computed to see whether any of them visually deviates from a lognormal distribution. 3 interesting QQ-plots from these 288 histograms and QQ-plots in total are displayed in Figures 5.34, 5.35 and 5.36. The first displays a QQ-plot where there are outliers that the fitted distribution will underestimate, the second one has a good fit in the tails, and the third has outliers in the tails that the fitted distribution will overestimate.

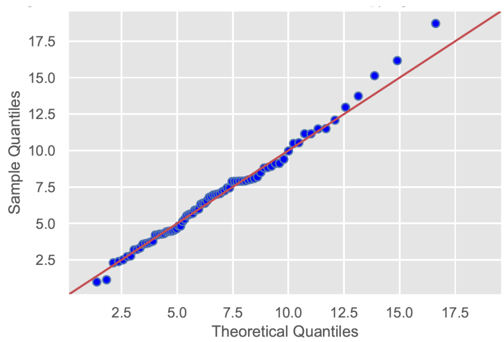

Figure 5.34: QQ-plot underestimating tails

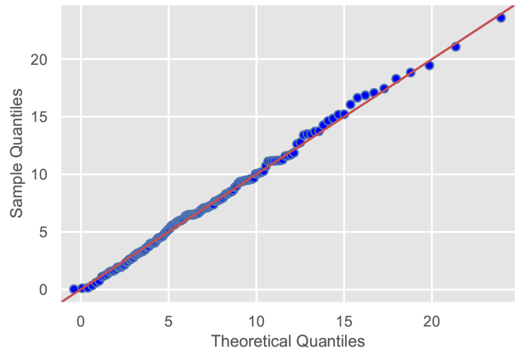

Figure 5.35: QQ-plot with good fit in tails

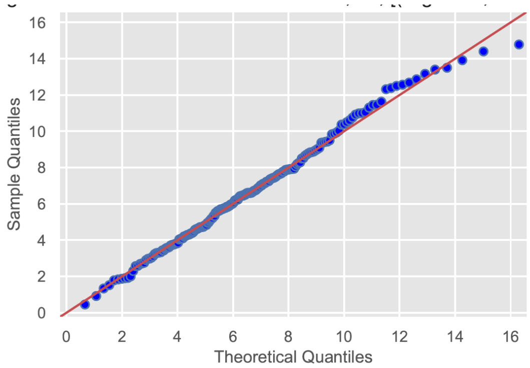

Figure 5.36: QQ-plot overestimating tails

ROOT MEAN SQUARE ERROR To see whether these outlier tails form a problem, the RMSE, root mean squared error, is computed for the fitted distribution compared to the data for all hours in the 6 service time groups. The Root Mean Square Error is a good metric for the goodness-of-fit of a hypothetical distribution to a data sample, as it computes the root of squared errors between the sample and the fitted distribution, that gets bigger when the tails of the distribution perform poorly and there are large outliers. The range of the RMSE for servicetime can be seen in Figure 5.37, where it is visible that most of the RMSEs are between 0 and 1, which indicates good performance, and few are above. For some hours in the data, there is some discrepancy between the fitted lognormal distribution and the sample, which is mostly due to one or two very big outliers when looking at the QQ-plots.

KOLMOGOROV-SMIRNOV AND ANDERSON-DARLING TEST Another test performed to assess the goodness-of-fit of the lognormal distribution for servicetimes is the Anderson-Darling and the Kolmogorov-Smirnov test. Both tests check whether a sample follows a theoretical distribution (null hypothesis). The Anderson-Darling test is a modification of the Kolmogorov-Smirnov test, and weighs the tails more than the KS-test, which tends to be more sensitive in the center of a distribution. As in some of the QQ-plots, there seem to be some problems with the tails, the Anderson-Darling test is a good choice to account for this. It should be noted that when the sample size is small, the Kolmogorov-Smirnov test tends to have less power - the ability of a test to find a difference between a sample and a theoretical distribution if there is one - than the Anderson-Darling test (Razali & Wah. Y.B., 2011). If the sample size is large, a statistical test like the two used here is better capable of finding differences and can reject the null hypothesis even though the distributions are

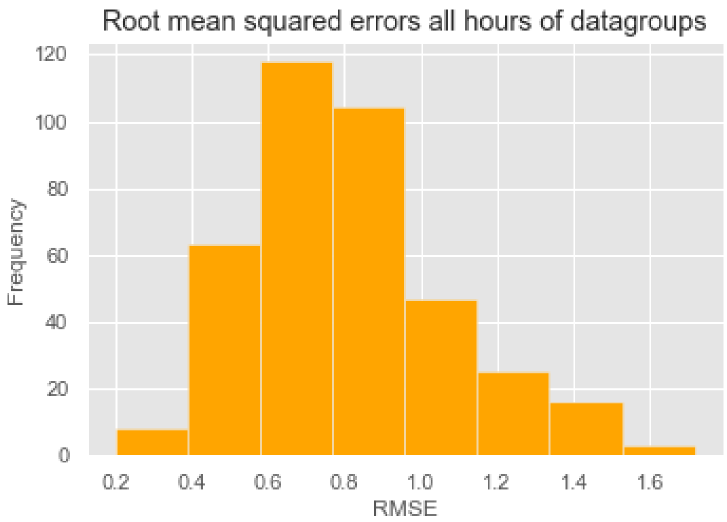

Figure 5.37: Root mean square errors when fitting lognormal distribution

very similar. In this case, the sample sizes vary from 1600 to 11000, in which range Razali and Wah. Y.B. (2011) reports no differences in power between the tests.

When running both of the tests on all the hours within the data groups for the service times, only in two hours a p-value below 0.05 is found, and only for the Anderson-Darling test. This means that for only two of the 16*24 = 384 hours a difference is found. It can be concluded that according to these tests, the lognormal distribution is a good choice for all data groups.

SUBPROCESSES IN TRIAGE SYSTEM As mentioned in Section 5.5.10, the system consists of two subprocesses when it comes to a phone call. The first subprocess is a phone call where no full triage is performed and a patient thus gets no allocated urgency. The second subprocess is where the full triage is performed and a patient gets an allocated urgency. The division in the data between these subprocesses is based on information from the first department that an average full triage conversation takes at least 6 minutes and based on the absence of an urgency level for a part of the calls. Because of this division, a fitted theoretical distribution on all the data would not accurately display what actually happens in the system. The use of empirical distributions of the two parts of the divided data in each hour of each service time scenario is therefore the most accurate way to follow the data and leave no information out.

CONCLUSION Overall, the lognormal distribution is the best fitting theoretical distribution on the data of the out-of-hours department. The RMSE's computed for servicetimes do not often increase a value of 1 and are mostly related to a few outliers, which on the one hand shows that the lognormal distribution is the best fitting theoretical distribution, but also that it comes with problems in the tails. Also, statistical tests are used that come to the conclusion that the lognormal distribution fits the data well. This should however be approached with some caution as sometimes they do not account for important problems in distribution fitting, such as problems in the tails.

All in all, the notion from literature that service times in call center systems follow a lognormal distribution can be confirmed, as besides the tails, it is the best fitting theoretical distribution on this data. However, because of problems in the tails and because the service time data actually consists of two different processes: no full triage or full triage, it is chosen to compute the empirical distributions of the different service time scenarios for the simulation model as implemented in Chapter 6.

5.9.3 Monte Carlo Sampling

The empirical distributions are used to sample from in the simulation model, conceptualized in Chapter 4 and implemented in Chapter 7. The model will be run multiple times to account for the stochasticity of the distributions and to account for extreme demand and servicetime scenarios, to be able to foresee what the impact on the system will be in these cases. In Figure 5.38, a potential Friday arrival pattern is visualized by sampling 25 times from the empirical distributions for each hour. This corresponds to running the model 25 times. The 95% confidence intervals are included and it is visible that most of the samples fall within this interval. When running the model, these plots are also made for waiting times, amount of patients in the queue, service times, urgency distributions for both arrival lines.

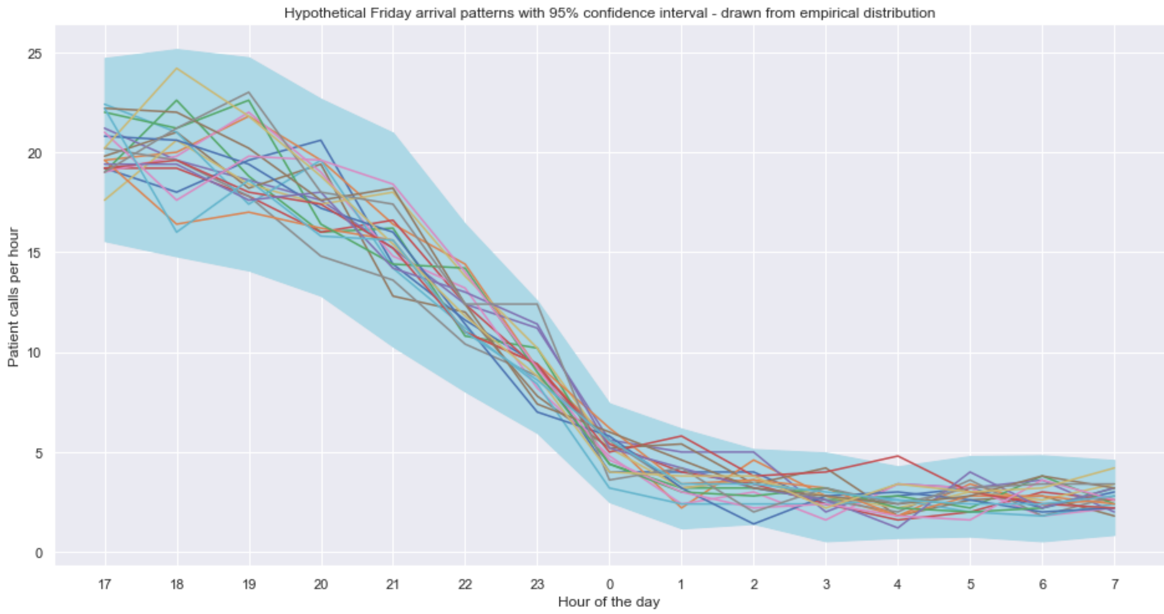

Figure 5.38: Hypothetical Friday 25 samples from hourly Poisson distribution

5.10 CONCLUSION

In this chapter, the different demand and service time scenarios where identified by systematically analyzing the data on different temporal, weather and urgency factors. It was found that these factors all have an effect on demand or service times to some extent and if a factor had a significant effect, a separate scenario for that specific factor value was created by grouping the data, for example on seasons, days of the week or on weather conditions. Next, it was concluded from statistically analyzing the distributions of the data that even though theoretical distributions could be fit on the data, reality would be better simulated in the model if empirical distributions are used. This is therefore used in the rest of the research.

By ending this Chapter, subquestion 2 is answered. When the model is implemented in Chapter 6, subquestion 3 can be answered in that same chapter, after which the interventions can be identified - subquestion 4 - in Chapter 7.

6 MODEL IMPLEMENTATION AND EXPERIMENTS

Based on the conceptualization, the analysis of the system data and the analysis of the distributions of the data, the model is implemented with the use of the programming language Python: an open source language used for many means and also a great tool for simulation modeling. This chapter describes the used tool and package for the simulation, it describes the flow of the model, the inputs, processes and outputs, and it describes the validation and verification of the model. Many of the discrete event simulation concepts used when for example discussing the model flow in this chapter have been discussed in Chapter 2 in Section 2.3. This chapter also experiments with various system changes to see its effects on system performance: waiting times, norm performance but also idle time of personnel. This concludes the chapter with an answer to subquestion 3, giving room to answer subquestion 4 in Chapter 7.

6.1 DISCRETE EVENT SIMULATION IN PYTHON

When implementing a discrete event simulation model in Python, one can choose between different packages, all with their own advantages. The most known one is *SimPy*, but many process interaction methods that are necessary for discrete event simulation models, like activation and deactivation or holding of components or concepts such as queues, monitors and states are not available in *SimPy*. A package that does include these concepts and was developed to complement the missing pieces of *SimPy* is the *salabim* package (van der Ham, 2018). The model is therefore implemented with use of the salabim discrete event simulation package, combined with other statistical and data packages *NumPy*, *SciPy* and *pandas* and visualization packages *matplotlib* and *seaborn*.

6.2 MODEL FLOW

In Figure 6.1, the model is visualized. The grey boxes display inputs into the model. The orange squares display a process, the orange ovals display the start and the end of the simulation and the white-orange diamond shaped figures are decisions within the system.

At the start of a simulation run, certain inputs have to be supplied to the model. These inputs are the amount of times the model should run, the amount of triagists (servers) in the model, whether it is a weekend or week run and the scenario in which the model should run. When the model starts, it immediately starts generating patients with a certain inter arrival time. The patients are *components* in the simulation model. The inter arrival time is deduced from the empirical arrival rate distribution for the scenario and the hour that the model runs in. There are two types of generators: an emergency patient generator and a normal patient generator with different arrival rates. When a patient is generated, it is checked whether or not this patient will receive full triage on the phone, or if the patient will just be serviced for a short while without going through the whole triage process. This decision is based on historical data which gives the model an empirical distribution over the two subprocesses. If the patient receives full triage, it will also be allocated an urgency level and it will enter the queue. If the patient does not receive full triage, it will not be allocated an urgency level but it will also enter the queue. If the patient is an emergency patient, the patient will enter the front of the queue since these patients have priority ("Last in First Out" principle, LIFO). If a patient is in front of the queue, it will be the first to be serviced by a triagist when it becomes available. A triagist is modelled as a *resource* in the simulation model, with a certain *capacity*, which corresponds to the amount of triagists working at that moment. The patient leaves the queue and is serviced with a certain service time. This service time is sampled from the empirical distribution of service times for that specific scenario, hour and urgency level. If the time horizon of the model is reached, the simulation stops.

6.2.1 Model inputs

In Figure 6.1 it can be seen that the model has a few inputs at the start of the simulation. These inputs are listed in Table 6.1. The demand and service time scenarios have to be typed in by the user of the model. Note that the amount of scenarios in the Table deviates from the amount of scenarios found in Chapter 5, because the emergency scenario is automatically selected based on the choice of demand scenario and the choice of weekend or week. The emergency scenarios are therefore not present to choose from.

The model simulates a day in the out-of-hours department in the chosen scenario. If weekend is set to False as input, the model will run for 15 hours, starting at 5PM when the out-of-hours department opens and ending at 8AM when it closes. If weekend is set to True, the model will run for 24 hours. The run will then start at midnight.

Table 6.1: Input values out-of-hours simulation model

Input	Possible values
Number of runs	>0
Number of triagists	>0
Weekend?	True or False
Demand scenario	1 of 13 scenarios
Service time scenario	1 of 5 scenarios

The rest of the inputs are empirical cumulative distribution functions that are used to sample from during the simulation. These inputs are listed in Table 6.2 and correspond to the grey boxes in Figure 6.1. It should be noted that all listed distributions have 2 versions: a normal patient version and an emergency patient version. The model has non-stationary inter arrival time distributions that vary hourly within the chosen scenario. The same goes for service time, which only does not vary hourly, but within three groups of hours in which the service times are the same, as identified in Section 5.5.9.

Shifts

In the real out-of-hours system, there are standard triagist shift schedules for weekdays Monday to Thursday and a separate schedule for Friday, Saturday and Sunday. In Table 6.3, these shifts are visible. These standard schedules are used for validation of the model.

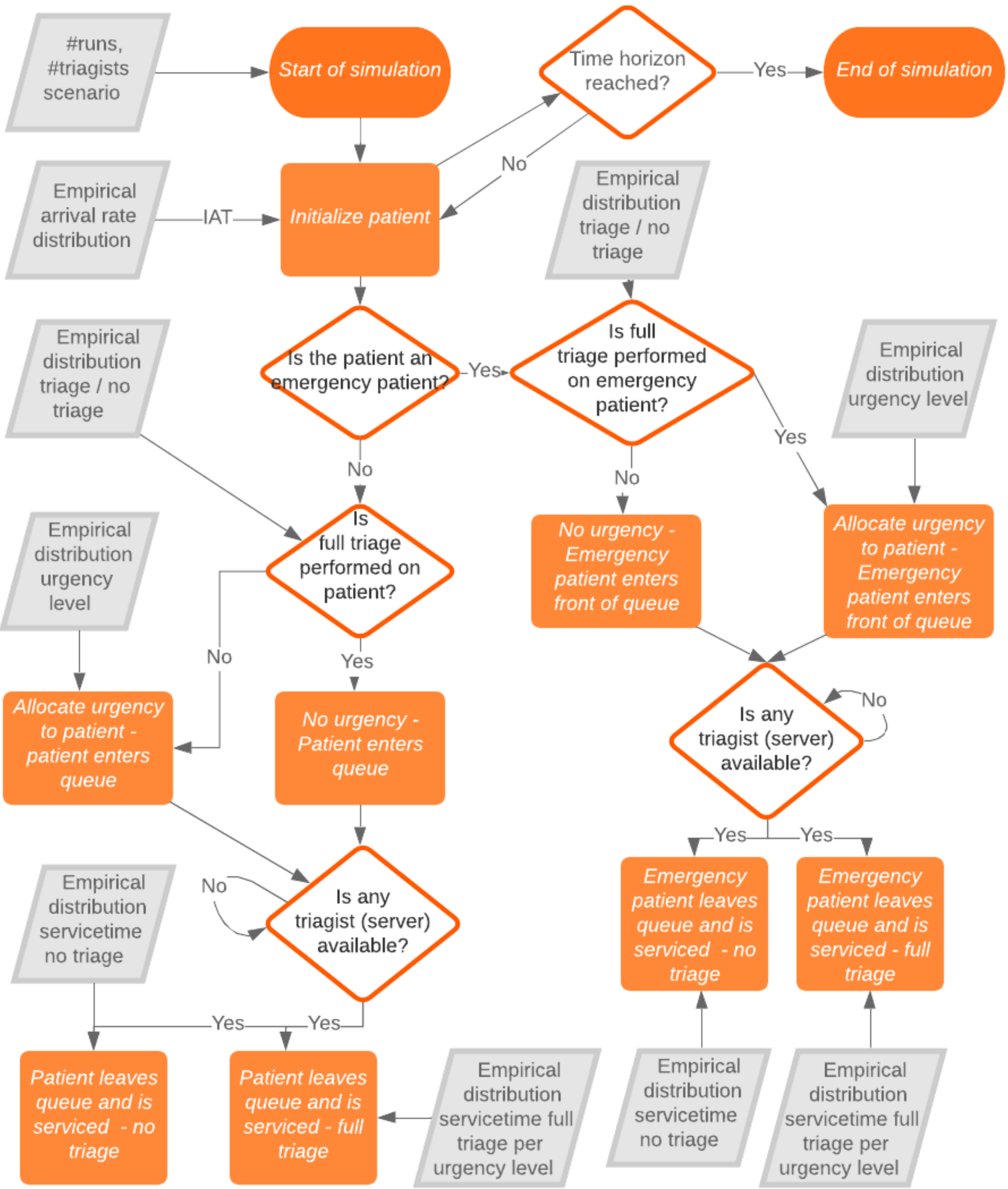

Figure 6.1: Flowchart of DES model implemented in Python

Table 6.2: Input distributions out-of-hours simulation model

Empirical distribution
Inter arrival time patients
Distribution patients over subprocess 1 and 2
Distribution patients subprocess 2 over urgency levels
Service time of patients from subprocess 1
Service time of patients from subprocess 2, different distribution per urgency level

More on these shifts and the way they are modelled can be found in the waiting time validation of the model, in Section 6.4.2.

Table 6.3: Triagist schedule out-of-hours department

Shift	Triagists
Weekday	17-18: 4 18-18:30: 3 18:30-21: 4 21-22:30: 3 22:30-8: 2
Friday	17-18: 4 18-18:30: 3 18:30-21: 4 21-22:30: 3 22:30-8: 2
Saturday	Day: 6.5 Evening: 3,5 Night: 2
Sunday	Day: 5 Evening: 3,5 Night: 2

6.2.2 Model outputs base model

A model run has 10 outputs. These outputs can be used for analysis and for visualization. The visualizations can be directly created in the same file as the simulation file to inspect what has happened in the model run. A list of the outputs is found in Table 6.4. With these outputs the entire performance of the system can be analyzed by looking at the utilization of triagists, waiting times, norms, demand and service times.

Table 6.4: Outputs from running simulation model out-of-hours department

Output name	Output format	Contents
Availability of triagists	List	For each run, the availability of triagists at discrete time steps
Occupancy of triagists	List	For each run, the occupancy of triagists at discrete time steps
Capacity of triagists	List	For each run, the capacity of triagists at discrete time steps
Length of stay in the queue	List	For each run, the length of stay in the queue of patients at discrete time steps
Patient results	Dataframe	ID, Urgency, Arrival minute, Arrival hour, Waiting time, Service time, End time of call, Total time in system, runID, performance norms 1&2
Emergency patient results	Dataframe	ID, Urgency, Arrival minute, Arrival hour, Waiting time, Service time, End time of call, Total time in system, runID, performance norm 3
Demand profile	Dataframe	Hour of the day with run ID, total arrival count in that hour,
Emergency demand profile	Dataframe	Hour of the day with run ID, total arrival count in that hour,
Performance on regular norms	Dataframe	Performance on norm 1 in % for each hour in each run, Performance on norm 2 in % for each hour in each run
Performance on emergency norm	Dataframe	Performance on norm 3 in % for each hour in each run

In Figures 6.2, 6.3, 6.4 and 6.5 some of the visualizations that display performance from 100 normal model runs of the Saturday Winter scenario are displayed to show what types of outcomes the model can generate. All Figures have a 95% confidence interval displayed. Figure 6.2 visualizes the waiting times over the day compared to the 2 minute norm that has to be met in 75% of the calls, Figure 6.3 displays the occupancy of triagists over the day, and Figures 6.4 and 6.5 visualize the performance on norms 1 and 2 each hour, in this case indicating that the normal amount of triagists does not give good performance on the norms, especially during the peaks

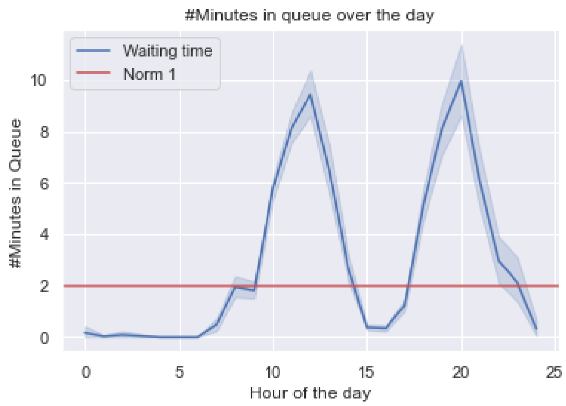

Figure 6.2: Waiting time over the day base model

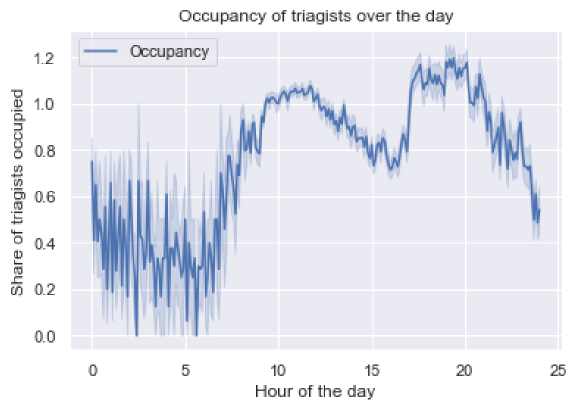

Figure 6.3: Triagist occupancy over the day base model

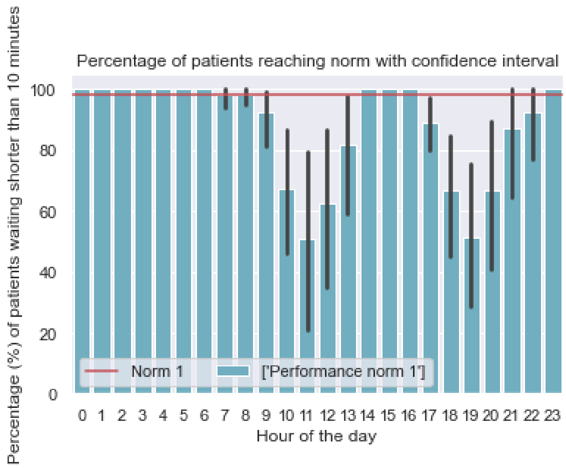

Figure 6.4: Performance on norm 1 base model

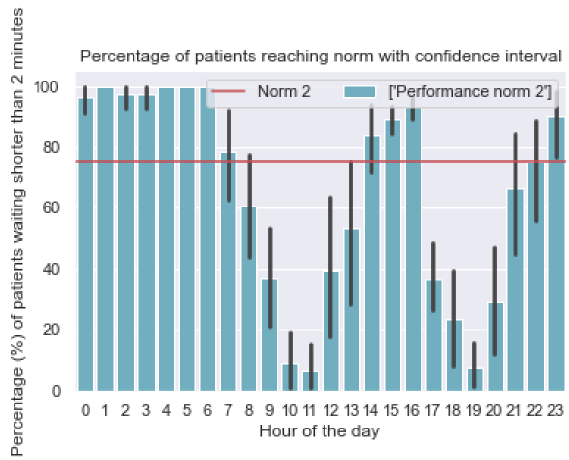

Figure 6.5: Performance on norm 2 base model

6.2.3 Warm-up time

In discrete event simulation, often a warm-up time is used to get the system in a state that is considered 'normal'. When the warm-up time is over, the simulation starts data collection at a point where normal conditions apply. For the out-of-hours department, a warm-up time is not necessary for weekdays as the state that the system is in when the simulation starts, without patients and no queue, reflects the state that the real system is in when the out-of-hours department opens at 5PM in the afternoon: empty. For weekend days and holidays however, a warm-up time from the last hours of the previous day might be necessary to get the system in the state that it was in when the previous day ended, so for example from Friday to Saturday or from Saturday to Sunday. A comparison is made between a simulation with a warm-up time from the last three hours of the previous day and a simulation without a warm-up time with two triagists during the night, the standard occupation in the base model. In Figure 6.6 the result of simulating with warm up time is visible when it comes to the waiting time of patients, and in Figure 6.7 it is visible for a simulation without warm-up time. When comparing the two figures, it can be seen that they are very similar and that the first hours of the day are the same. This means that adding the previous hours of the day before to the simulation as warm-up time does not have an effect on the performance of the system. This is due to the peaks in demand taking place mostly during the daytime and not at night: the queue tends to be empty at night. It is therefore concluded that a warm-up time is not necessary in the out-of-hours telephone triage system: it will increase run times but will not add value to the results.

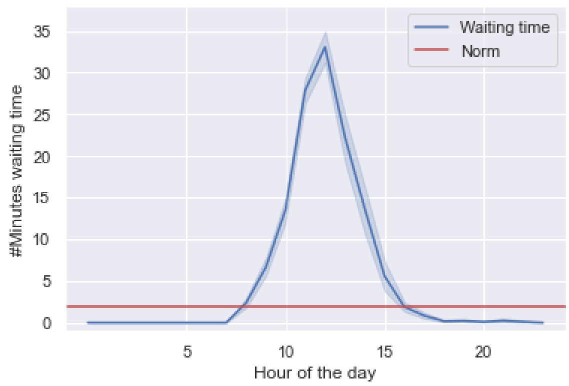

Figure 6.6: Waiting time over a weekend day with warm up time

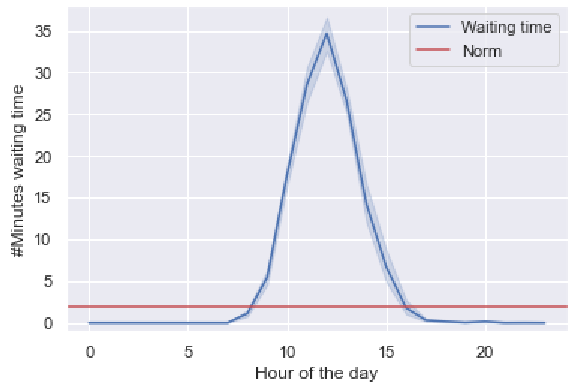

Figure 6.7: Waiting time over a weekend day no warm up time

6.3 MODEL VERIFICATION

In the model verification, it is checked if the implementation of the model is according to the conceptualization of the model, as described in Chapter 4 and complemented with conclusions and scenarios from the data analysis as described in Chapter 5. By walking through the model, its components and processes, it can be checked for bugs and errors and it can be checked whether the conceptual model is adhered to (Sargent, 2011).

6.3.1 Model component verification

Patient & Emergency patient generator

First of all, the patient generator and the emergency patient generator are verified. These components create the patients throughout the simulation. After each patient, they 'hold' for a sampled amount of time: the inter arrival time. From Chapter 4 and 5, it became clear that the demand for care at an out-of-hours department varies within different scenarios, but also varies hourly within the different scenarios. This is implemented in the model by hourly changing the empirical distribution for demand during the simulation, based on the running scenario. To verify that this notion is correctly implemented into the model, the demand profile is visualized for the normal patient generator and for the emergency patient generator, respectively in Figures 6.8 and 6.9. It can be seen that the demand pattern varies over the day and therefore the notion that demand is different per hour and sampled from a different distribution is verified.

Patient

Next, the patient and the emergency patient component itself are verified. From these components it is most important that they accurately get their attributes allocated for data collection. Also, they should go trough the processes in a correct and chronological order: they need to arrive, get a subprocess allocated, if applicable an urgency level, then they need to enter the queue and if they are in front of the queue they need to be serviced. Only

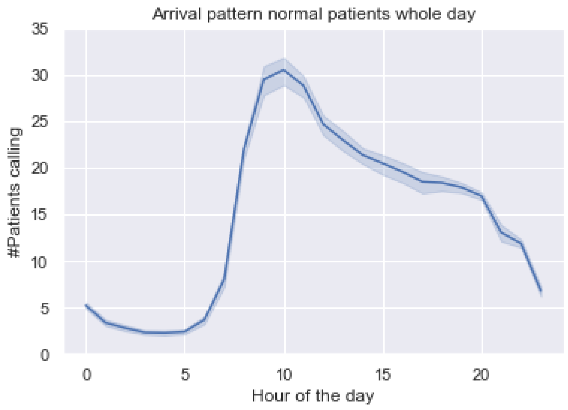

Figure 6.8: Demand profile normal patient generator for verification

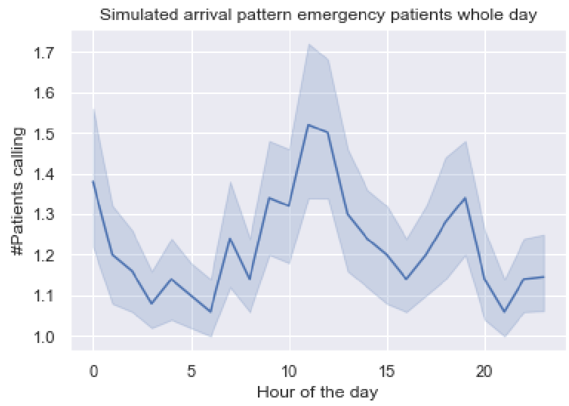

Figure 6.9: Demand profile emergency patient generator for verification

after all that has happened, they should leave the system. If this is not the case, the outputs of the model are incorrect and results cannot be used by out-of-hours departments. In the conceptualization the different input variables that should be available in the model are listed. From these inputs, the following are associated with a patient: arrival time, service time, subprocess, urgency level and performance on the waiting time norm. Next to these, the waiting time and the time that a call stopped are stored. The model is run within the Saturday scenario, and both the data of a normal and of an emergency patient are inspected to verify the values of all the patient data and to verify the chronological order.

patient_ID	4.000000
Urgency	3.000000
Start_Call	60.571429
Start_Call_Hour	1.000000
Waiting_Time	5.928571
Service_Time	3.500000
Stop_Call	71.000000
TimeInSystem	10.428571
runID	0.000000
Norm1	1.000000
Norm2	0.000000

Figure 6.10: Normal patient data from simulation model

empatient_ID	20
Urgency	2.0
Start_Call	1060
Start_Call_Hour	17
Waiting_Time	0.0
Service_Time	12.0
Stop_Call	1073.0
TimeInSystem	13.0
runID	0
Norm3	True

Figure 6.11: Emergency patient data from simulation model

In Figures 6.10 and 6.11, the data associated with one normal patient and one emergency patient from a model run are displayed. Based on these figures, it can be concluded that a patient goes through the model chronologically and that the simulation model accurately displays patient data: the start of the call is for both patients before the end of the call. The time in system is the difference between these two values, and the service time has to be lower or equal to the time in the system. The waiting time and the service time may should be equal to the time in the system minus 1, as the handling time that the triagist needs for the patient after the call (1 minute) is included in the time in the system. Next, it is checked whether the Norm variables work as they should: they are 1 or True if the norm is met, and 0 or False if the norm is not met. For the normal patient, it can be seen that Norm1 is indeed met: the waiting time was below 10 minutes, and that Norm2 is indeed not met: the waiting time was above 2 minutes. For the emergency patient, it can be concluded that Norm3 is indeed met: the waiting time was below 30 seconds.

To also verify that every patient is serviced by a triagist, it is checked whether there are patients with a service time of zero. This is neither the case for normal patients or emergency patients.

Lastly, as visualized in Figure 4.1 in Chapter 4, emergency patients have priority in the queue and are placed in front when they arrive. To verify whether this is correctly implemented in the model, the trace of the queue is followed to see what happens to an emergency patient when calling the out-of-hours department in the model. In Figure 6.12 the trace at the time of arrival of an emergency patient is displayed, and it can be seen that even though many other patients were already in the queue and the capacity of the triagists is only 4, the emergency patient is immediately serviced (it claims 1 of the 4 triagists).

```
name=triagists
capacity=4
requesting component(s):
  patient.206          quantity=1
  patient.207          quantity=1
  patient.208          quantity=1
  patient.209          quantity=1
  patient.210          quantity=1
  patient.211          quantity=1
  patient.212          quantity=1
  patient.213          quantity=1
claimed_quantity=4
claimed by:
  emergencypatient.18  quantity=1
  patient.202          quantity=1
  patient.203          quantity=1
  patient.205          quantity=1
```

Figure 6.12: Trace at arrival of an emergency patient

It can be verified that the patient and emergency patient component behave as indicated in the conceptualization and data analysis chapter.

Triagist
The triagists in the simulated model are modelled as a resource with a certain capacity. As can be seen in the conceptualization and in Table 6.3, the capacity of the triagists varies throughout the day because of the different shifts that have to be scheduled. It has to be verified that the capacity also varies throughout the day in the simulation model. If this is not the case, no efficient allocation of triagists can take place with the model, which is one of the main reasons of creating the model in the first place. In Figure 6.13, the capacity throughout the day for a Weekday can be seen. It is visible that the capacity deviates throughout the day, according to the way it is inputted into the model. It can therefore be verified that the amount of triagists varies throughout the day.

6.3.2 General verification

Subprocesses and urgency verification
In the model, certain attributes are given to the patients. It is found in the Patient verification that the order and the values of these patients are valid, but it should also be verified whether indeed the patients are split up in two different subprocesses correctly by the model and whether the empirical distributions sample in the right way. It should also be verified whether a patient that is assigned to the second subprocess gets an urgency level allocated between 0 and 5. If this is not verified, the service times in the model are not

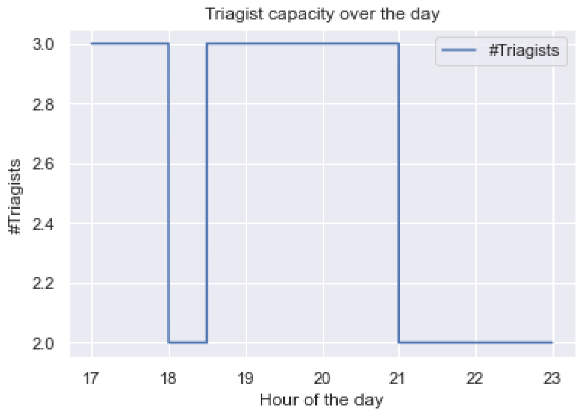

Figure 6.13: Triagist capacity over the course of a simulation

accurate as they depend on subprocess and urgency level. In Figure 6.14, the histogram that counts the amount of patients for each urgency level is displayed. It can be seen that of each urgency level patients are present and that there are also patients in the system that did not receive an urgency level: these were the patients that were assigned to the first subprocess where they do not get full triage. The distribution over the urgency levels is also similar to the empirical distribution that it is drawn from: many patients have urgency levels 3, 4 and 5, and very few people have urgency level 0 and 1. Only around 10% of the patients gets no urgency, which matches the empirical distributions (see Figure 6.31 in the validation, where the low points at 0.1 represent the share of patients in subprocess 1).

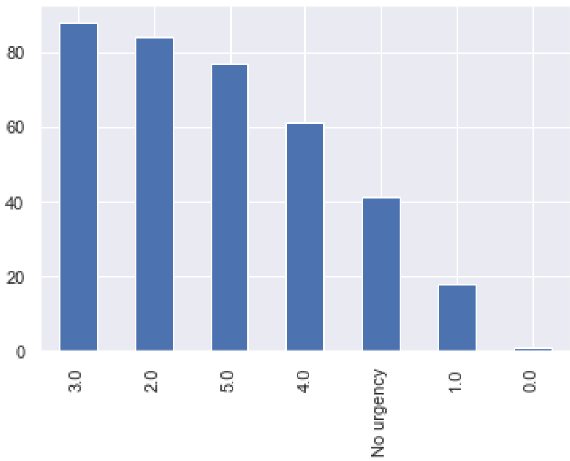

Figure 6.14: Amount of patients per urgency level over 1 simulation run

Service time verification

Similar to the subprocess and urgency levels, a service time is allocated to a patient. This service time is different per subprocess and urgency level. In the paragraph above it was verified that the patients correctly get subprocesses and urgency levels assigned, with the right probabilities. It should next be verified that for each of these groups, a different service time distribution is used to sample from and that the service times are thus not the same for each group. In Figure 6.15 it can be seen that the service times are slightly different per urgency level, which corresponds to what was found in Chapter 5, and that the service time for a patient without urgency so without full triage is a lot lower. It can be concluded that service times are accurately assigned to their corresponding subprocess and urgency level and sampled from different distributions.

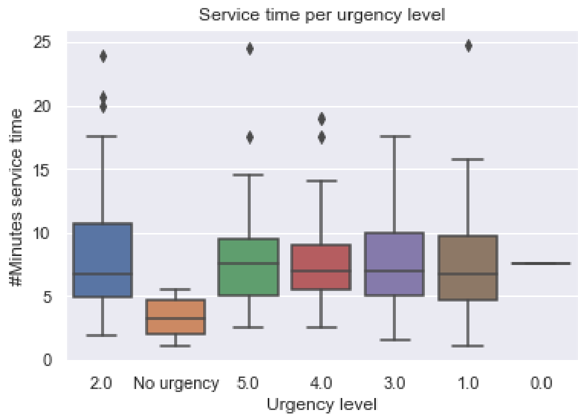

Figure 6.15: Service times of patients per urgency level over 1 simulation run

6.4 MODEL VALIDATION

When validating a model, it is checked whether the results of the model come close to the observed results from reality, taking into account the intended use of the model. To achieve this, multiple validation methods are used (Sargent, 2011). First, a sensitivity analysis is conducted to verify if changes in input and internal parameters result in plausible model outputs. Second, results of running the simulation model are compared to historical data from the out-of-hours department that the model inputs are derived from. Third, it is tested whether the extreme values of the input variables have plausible effects on the outcomes of the model. Lastly, it is checked whether the literature and expert interpretation of the data and the system in Chapter 5 are valid for the simulation model as well. It should be noted that for the validation the following two scenarios are chosen to run in the simulation model: the Saturday scenario in Autumn, Summer or Winter for a weekend day, and the Friday scenario in the same seasons for a weekday.

6.4.1 Sensitivity analysis

The first validation technique is sensitivity analysis. For the simulation model, the input and internal variables that are not directly derived from the data are varied across a certain range and it is checked whether the outputs of the model are plausible with the changed input or internal variable. Also, by running a sensitivity analysis, the most sensitive variables to the model outputs are identified for which it is highly important that their values are accurate. The input and internal variables that are changed and the range within which they are changed and run in the Saturday scenario can be found in Table 6.5. The sensitivity analysis is conducted by running the model itself various times, but also by running it using the Exploratory Modelling Workbench. This is a Python package that allows for open exploration and optimization with a model (Kwakkel, 2017). The package creates experiments with different combinations of uncertain values (the handling time) and policy values (the capacity of triagists) and runs many of them to explore the behavior of the model. For the sensitivity analysis, 1000 experiments were run.

In Figure 6.16, the results of the sensitivity analysis with the model itself can be seen with 95% confidence intervals, displaying the average and maximum lengths of stay in the queue for different day and evening triagist capacities. It can be seen that the system outputs are very sensitive to the amount of triagists during the day and during the night, and that the relationship is not linear. Often, adding one extra triagist (for example from 3 to 4 during the day) gives way more waiting time reduction than increasing from 4 to 5. It can therefore be concluded that a trade-off exists between capacity of triagists and length of stay in the queue: the higher the capacity of triagists, the lower the length of stay in the queue.

Table 6.5: Varied variables and its range in sensitivity analysis

Variable	Range
Amount of triagists day	3 - 6 (steps of 1)
Amount of triagists evening	1 - 4 (steps of 1)
Amount of triagists night	1 - 3 (steps of 1)

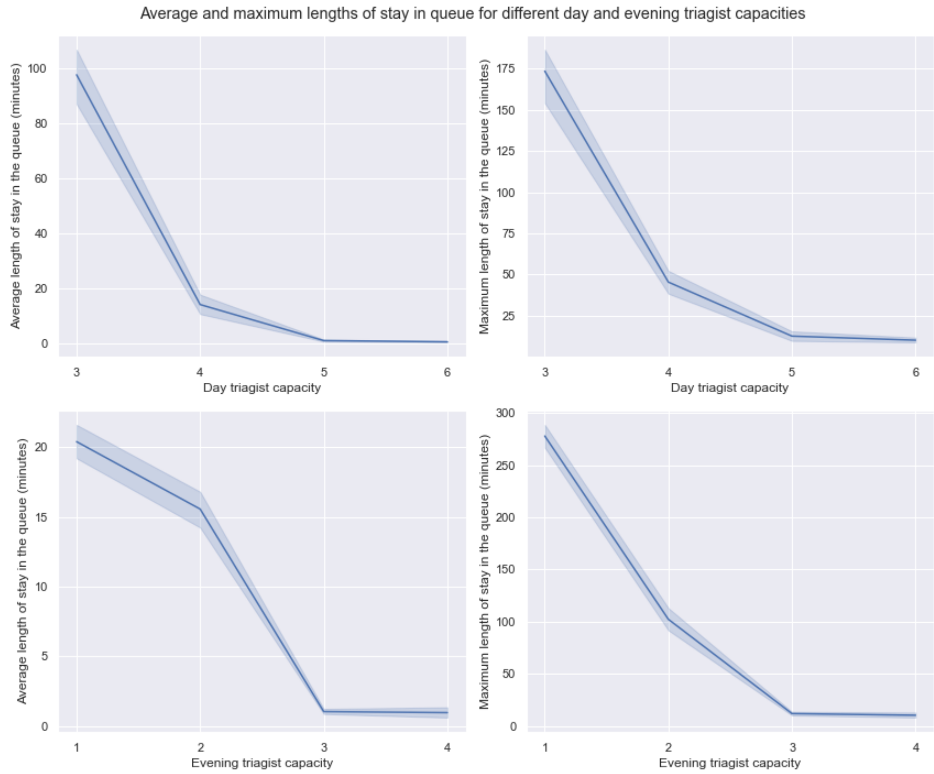

Figure 6.16: Results sensitivity analysis triagists day

In Figures 6.17 and 6.18, the results of the sensitivity analysis using the Exploratory Modelling Workbench are displayed. In Figure 6.17 the relationship between waiting times and triagist occupancy are displayed. It can be seen that similar to the trade-off seen in the previous Figures, also a trade-off exists between waiting times and occupancy of triagists, which is not linear: occupancy very quickly increases to 1 once the waiting times start increasing. This indicates that there is not much room for manoeuvring between meeting the norms and keeping waiting times low and large waiting times where the norms are not met.

The metrics that are visible in Figure 6.18 are *occ_avg*, the average triagist occupancy, *los_avg* and *los_max*, the average and maximum length of stay in the queue (waiting times in minutes). The capacities *cap_day*, *cap_evening* and *cap_night* stand for respectively the day, evening and night capacity of triagists. The Figure displays the importance of variables in the model. It can be seen in Figure 6.18 that the same conclusions can be drawn from running the model with the EMA-workbench as when running the model itself multiple times with different day and evening capacities: the model outcomes for waiting times are really sensitive to the amount of triagists during the day and also the evening. Also, this figure indicates that the model outcomes are not sensitive to night triagist capacity. This makes sense, as there are often very few patients during the night and increasing capacity will not give better performance.

The results of the sensitivity analysis emphasize the notion that the allocation of workforce, the triagists, is very important within out-of-hours care, especially during the day. A balance should be found between capacity of triagists and the length of stay within the queue. More on capacity optimization and on other system changes that might help reduce waiting times can be read in Section 6.6.

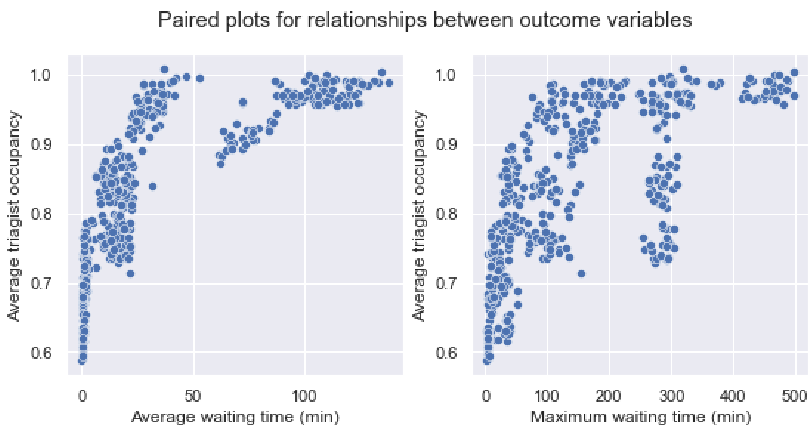

Figure 6.17: Sensitivity analysis with EMA workbench - paired plots between outcomes

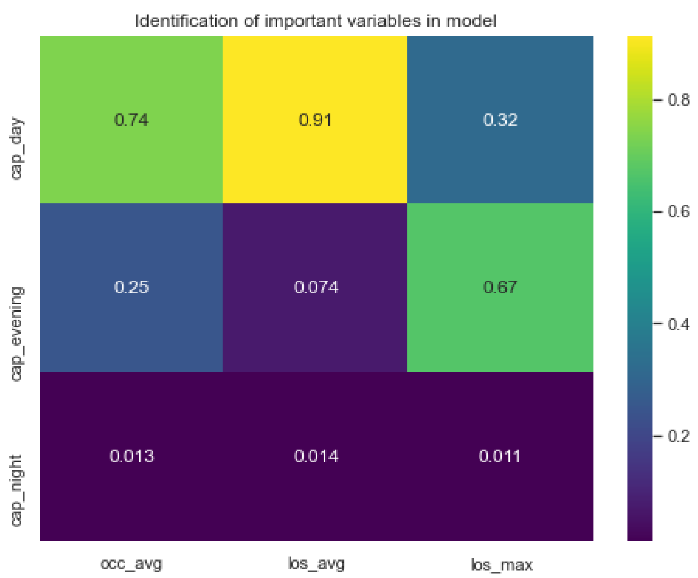

Figure 6.18: Sensitivity analysis with EMA workbench - identification important variables

6.4.2 Historical data validation

To test whether the model accurately simulates the telephone triage system in an out-of-hours department, historical data from different processes within the system is compared to the simulated processes from the model to validate whether the behavior of the model is similar to the behavior in reality.

Normal demand profile validation
The demand profile of patients over a day of a specific scenario simulated by the model is compared to the real demand profile from historical data of the system.

In Figures 6.19 you can see the demand profile of a Saturday in Autumn, Winter or Summer, obtained by running the simulation model 50 times, compared to the observed demand profile for a weekend day from the data that consists of approximately 100 days of this type. The 95% confidence intervals are displayed in the figures to account for the stochasticity and variability of the model: patient arrivals are not always the same. Visually, it can be seen that the patterns are very similar and the confidence intervals mostly overlap. To mathematically validate the similarity, the correlation between the observed and the simulated means of the data points is displayed in Figure 6.22. If all points lie perfectly on the red line, the simulated data and the observed data are exactly the same and have a correlation of 1. The R-squared value when using a simple regression model where the observed values are the predictor values and the simulated values are the response values is 0.998, which indicates that the observed means can represent a large portion of the variance in the simulated means and thus that the simulated demand accurately represents the system (and the other way around). Note that in this Figure only the means per hour are displayed, while in reality there are many data points associated with each hour which leads to a confidence interval around the mean. These are visible in the left Figure.

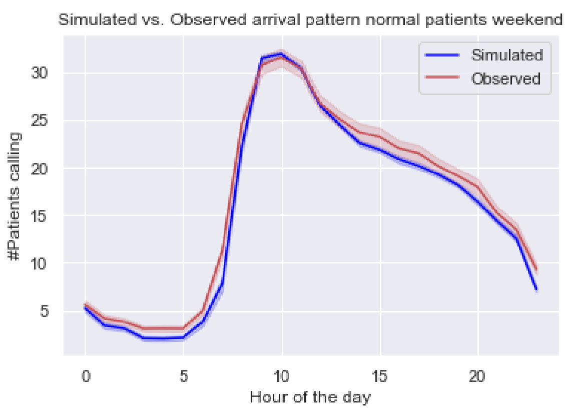

Figure 6.19: Simulated vs. Observed demand profile Weekend day

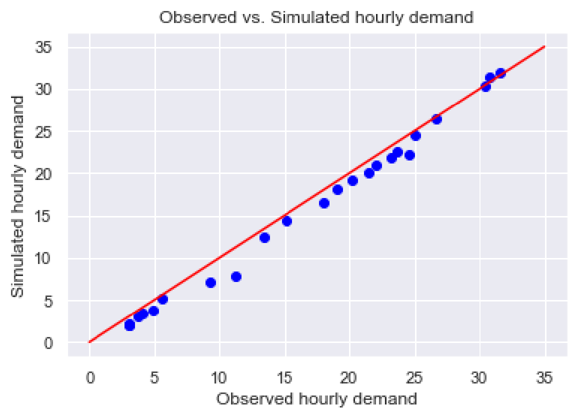

Figure 6.20: Correlation observed vs. simulated demand profile weekend

The same is done for weekdays, visible in Figures 6.21 and 6.22 for a Friday in Autumn, Winter or Summer. The department opens at 5PM, and the model runs until 8AM Saturday because it treats Friday as a weekday. The whole of Saturday can be simulated by running a weekend scenario, as just done in the weekend validation of the demand profile right above. The model is again run 50 times, and the observed pattern consists of It can be seen that again, the patterns are very similar. To mathematically verify this, again the R-squared value of the observed values versus the simulated values is computed and found to be 0.997. Again, this indicates that a large portion of the variance in the simulated means is explained by the observed means and it can be deduced from the Figure that the correlation between the two samples is nearly 1.

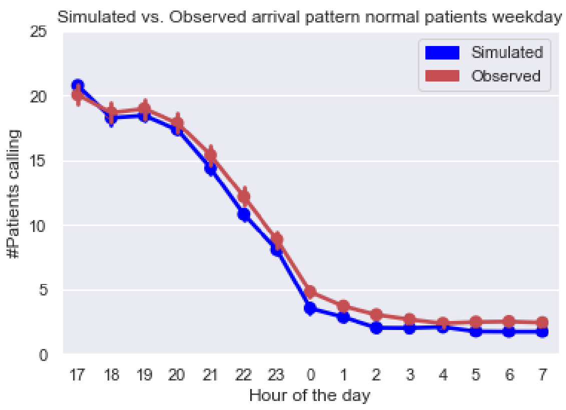

Figure 6.21: Simulated vs. Observed demand profile Weekday

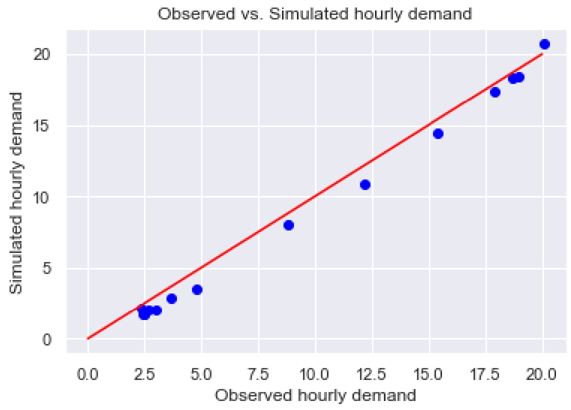

Figure 6.22: Correlation observed vs. simulated demand profile Weekday

It can be concluded that the simulation model accurately simulates normal patient arrivals.

Emergency demand profile validation
For the emergency demand profiles the same type of validation is executed. In Figure 6.23 a comparison between the observed and simulated emergency arrivals on a Saturday in Autumn, Winter or Summer is displayed. It can be seen that the variability in the emergency patients seems high, but on a small scale. An outlying amount of patients arriving has a large effect on the graph and its variability. When looking at Figure 6.24 it can be seen that the correlation is positive, but has a high variance. The R-squared is 0.645, which indicates that the correlation is not perfect and the variance of the simulated values cannot be represented completely by the observed values. It is chosen to accept this because of the small scale in which the emergency arrivals fluctuate and because of the stochasticity of the model.

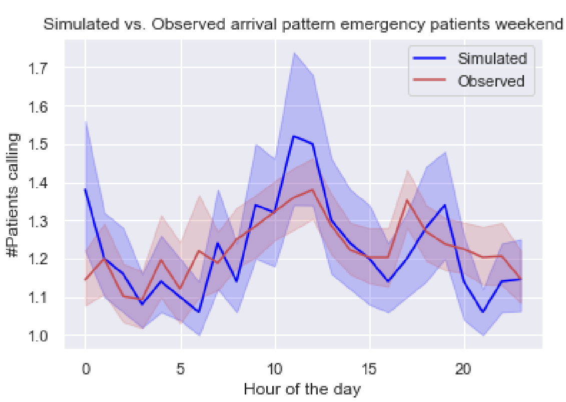

Figure 6.23: Simulated vs. Observed demand profile Weekend emergency patients

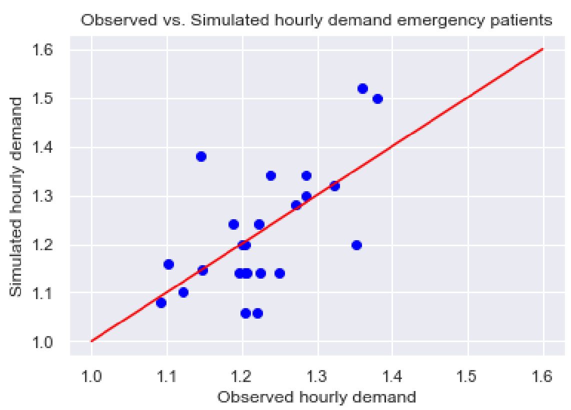

Figure 6.24: Correlation observed vs. simulated demand profile Weekend emergency patients

For the weekday emergency demand profiles, on a Friday in Autumn, Winter or Summer, again the observed data is compared to the simulated data. Similar conclusions can be drawn from Figures 6.25 and 6.26, although the variance is less high for the weekdays than for the weekends, supported visually since the data points are closer to the red line and by the R-squared value of 0.851 which is significantly higher than the R-squared value of 0.654 for the emergency line in the weekends.

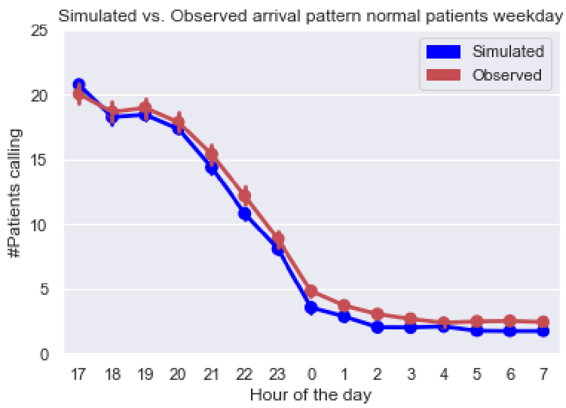

Figure 6.25: Simulated vs. Observed demand profile Weekday emergency patients

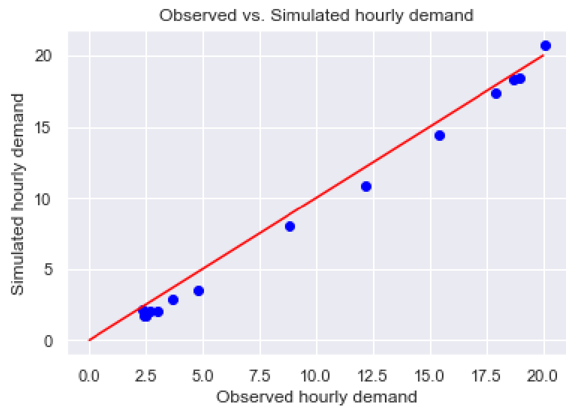

Figure 6.26: Correlation observed vs. simulated demand profile Weekday emergency patients

It can be concluded that the simulation model accurately simulates emergency patient arrivals when accepting the fact that outlying values can cause great variability in the overall demand profile. This is something that is useful in a simulation model, since it should also account for more extreme cases of multiple emergency arrivals in an hour.

Service time validation

The same is done for service times of the normal arrival line, visible in Figures 6.27 and 6.28 for the Winter scenario for urgency levels 0, 1 and 2. The model is again run 50 times. It can be seen that again, the patterns are very similar. For the second Figure, only three points are visible, because for the service times 3 hourly groups were identified in Section 5.5.9 and these were thus aggregated when checking how well the model simulates service times. From these hour groups, the means are computed for the observed and simulated data and compared to each other in the correlation plot. It can be seen that the data points are close to the red line. The R-squared value is found to be 0.956, again a high value that indicates good explanation of the simulated data by the observed data.

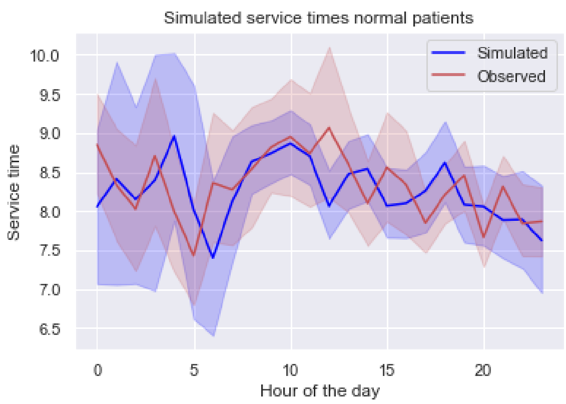

Figure 6.27: Simulated vs. Observed service times

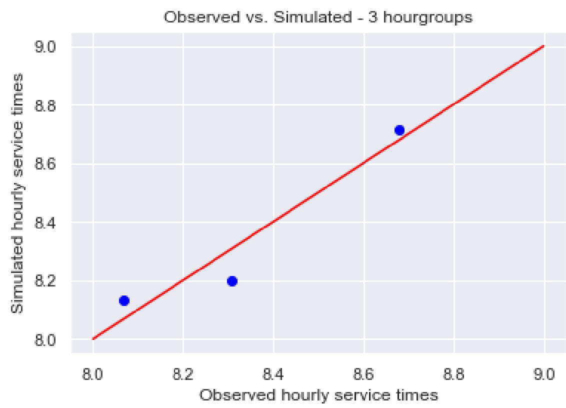

Figure 6.28: Correlation observed vs. simulated service times per hourgroup

For the emergency line, the observed and simulated service times of patients are compared in Figure 6.29. It can be seen that, similarly to the demand profiles of the emergency line, the variability is high because of the low amount of patients arriving to the emergency line and the effect that an outlier can have on the pattern. In Figure 6.30, it can be seen for the three hourly groups of service times that they are close to the red line: the difference is very low on the scale. The R-squared value is 0.923, indicating a high correlation between the two samples.

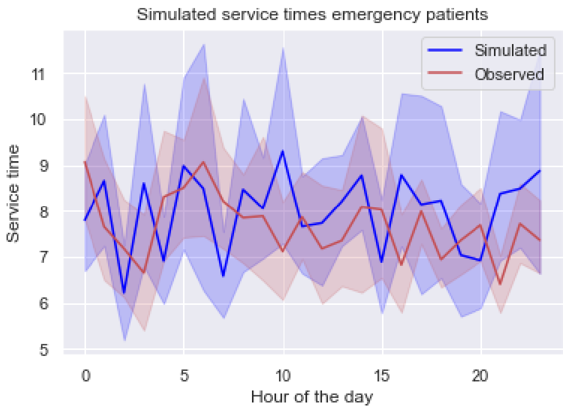

Figure 6.29: Simulated vs. Observed service times emergency patients

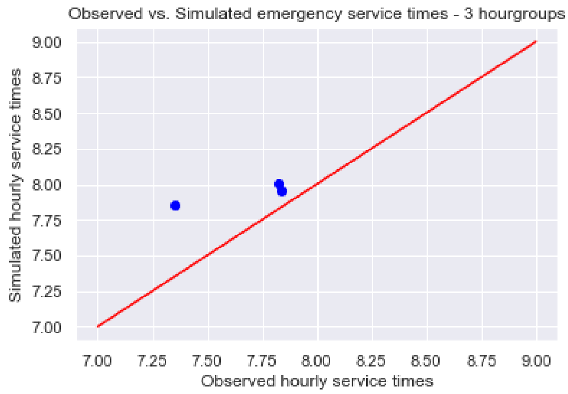

Figure 6.30: Correlation observed vs. simulated service times per hourgroup emergency patients

It can be concluded that the service times are accurately simulated by the model. Again, for the emergency line, higher variability can be seen due the lower amount of emergency patients arriving, which accurately displays the uncertain situation in reality.

Subprocesses validation

Arriving patients are divided into two subprocesses, no full triage and a full triage conversation. This division is based on the data, as explained in Section 5.5.10. This distribution differs per hour in which the patient arrives, aggregated to three groups as service time is the same within these three groups. To validate whether the model divides the patients into the two subprocesses similar to the real division from the data, the correlation between the simulated and observed data was displayed in Figure 6.31. There are 6 datapoints: for each of the two subprocesses there are three hour groups. They are again all below 1 as they represent a probability of getting a certain subprocess assigned. It can be seen that the observed probability of being assigned to a certain subprocess is very similar to the simulated probability, both for the lower probabilities (the no full triage subprocess) and the higher probabilities (the full triage subprocess in which most patients end). The R-squared value supports this with a value of 0.999.

Urgency distribution validation

For the distribution of urgency levels over the patients that receive full triage, also a comparison between the observed distribution and the simulated distribution is made. The data from a Saturday in Autumn, Winter or Summer is used. In Figure 6.32 the correlation between the observed and simulated distribution is displayed. The numbers are below 1, as they stand for the probability that a patient will get a certain urgency. The R-squared value of a simple linear regression gives an R-squared value of 0.94, an indication of a good fit and a high correlation can be deduced from the Figure.

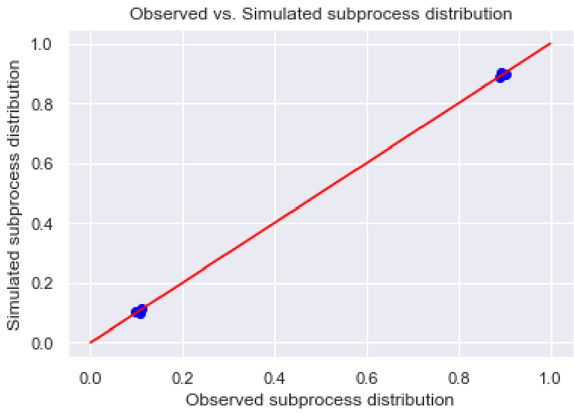

Figure 6.31: Correlation observed. vs. simulated subprocess distribution

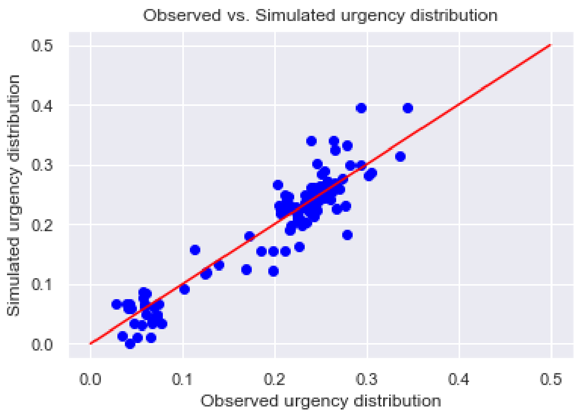

Figure 6.32: Correlation observed. vs. simulated urgency distribution

Waiting times validation
In Figures 6.33 and 6.34, the observed and simulated waiting times of a Cold weekday are displayed. In these Figures, the standard deviation is plotted in stead of the confidence interval to show the spread of waiting times especially in the observed data, as there are so many values that the confidence intervals are really narrow, and a statistically significant difference will always be found between some simulation runs and all the recorded data. All values of the simulated waiting times lay close to the means of the observed waiting times. The same is visible for the observed and simulated waiting times of a Saturday, visible in Figures 6.35 and 6.36.

It should be noted that to come to this validation the model was first run with the standard schedule of triagists as visible in Table 6.3. When running the model with this schedule, it was found that it is too optimistic to assume that the scheduled triagists are always available and taking phone calls non-stop, immediately when a patient enters the queue. Modelling the model that way results in way better system performance than the data indicates. This deviation in historical performance data from the performance that one would expect with the scheduled triagists when running the model can be due to breaks of triagists, longer handling times than the average, dealing with other types of calls and problems in the department, not immediately picking up the phone when a triagist is available and a caller is in the queue and might be related to other factors that come with the fact that the system is a system in which humans operate. A waiting time of exactly zero, which the simulation model reaches often when capacity is high enough, is in reality almost never reached.

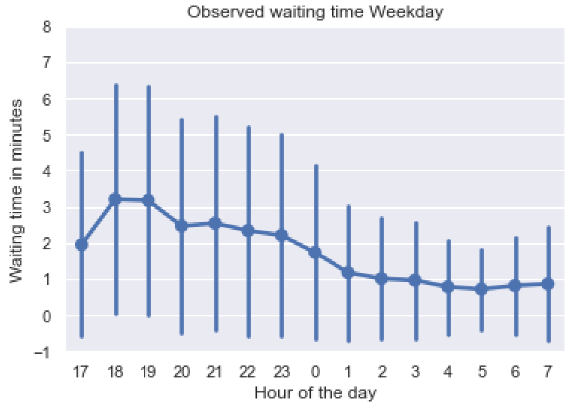

Figure 6.33: Observed waiting times cold weekday (Mon-Thur)

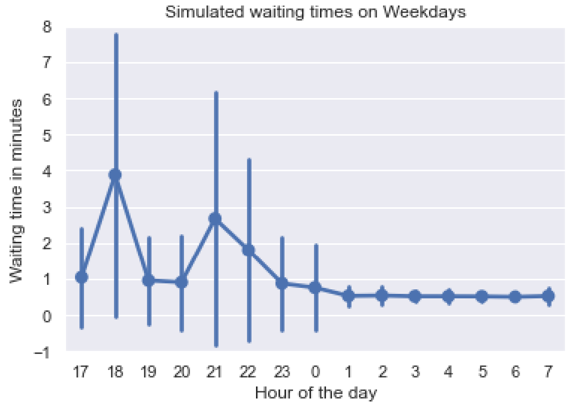

Figure 6.34: Simulated waiting times cold weekday (Mon-Thur)

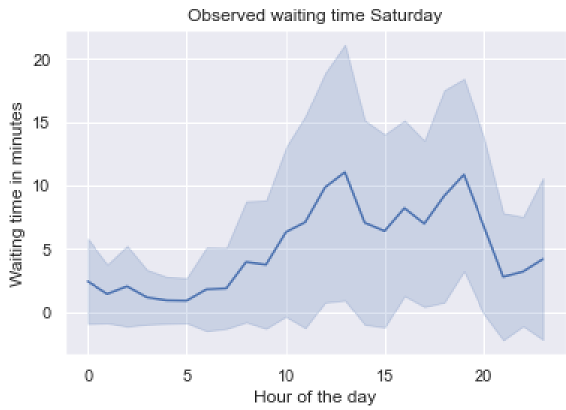

Figure 6.35: Observed waiting times weekend (Saturday)

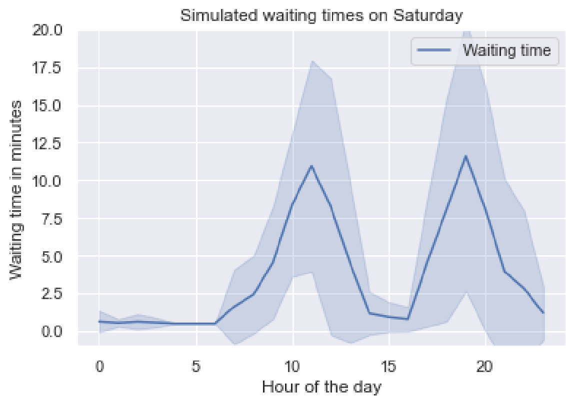

Figure 6.36: Simulated waiting times weekend (Saturday)

To account for these factors, many combinations of handling times, triagist capacities and delays before picking up the phone have been tested. It is found that when reducing the scheduled capacity by 1 during the week, using a delay before picking up the phone of half a minute and using a handling time after a call of 1 minute, the system operates as indicated in the data and as is visible in the displayed figures. The same is done for the weekends: the constantly available triagist capacity is reduced by 2 during the day, so it is 4 or 5 depending on the run, and the evening capacity is reduced by 1, sampled between 2 and 3. The night capacity stays 2 for the weekends and weekdays.

6.4.3 Extreme conditions validation

The model is tested for extreme conditions. The first case that is tested is when there are no triagists available during the simulation, to see whether indeed no patients are serviced and the queue only becomes longer. When running this situation in the model, none of the patient or emergency patient outputs are filled as that happens after being serviced by a - here non-existent - triagist. The total length of stay in the triagist queue is also non-existent, since there is no capacity of triagists. The second case is when there are no patient arrivals, to see whether indeed no patients enter the queue and are serviced by the triagists. When running this situation, no patient enters the queue as the length of stay monitor of the triagists queue is completely empty. Also, the capacity of the triagists is always at its fullest. It can be concluded that these extreme conditions have the expected outcomes in the simulation model.

6.4.4 Face validation

Face validation is the validation of the model by people with knowledge about the system. The behavior of the system and conclusions from analyzing its data are interpreted by literature and by people working within the out-of-hours field in Section 5.6, and since the model accurately displays the historical data, these interpretations are valid for the model outcomes.

6.4.5 Second department validation

In Appendix A, the observed waiting times from the data-set of the second out-of-hours department are compared to the simulated waiting times within an identified scenario. It can be concluded that the model also accurately simulates waiting times for another data-set. This indicates that the model has a general use for out-of-hours departments with telephone triage systems like the one that these two departments have and that change of data and input scenarios does not change the accuracy of the model as the model was not built around the data, but around the concepts of the system as identified in Chapter 4.

6.5 Conclusion

With the demand and service times scenarios defined, the empirical distributions for all input and internal variables identified and the model implemented, verified and validated, the model can serve a practical day-to-day use for out-of-hours departments by foreseeing the performance of the system as it is now when it comes to waiting times, people in the queue, handling and service times and performance on the set norms. In the following sections, the model is used for the identification of input changes or internal system changes that lead to the reduction of waiting times. The promising system changes are used for the exploration and identification of system interventions - subquestion 4 - in Chapter 7.

6.6 Results: system experiments for waiting time reduction

In this Section, subquestion 3 is answered: different system changes are experimented with to see their effect primarily on the performance of the out-of-hours telephone triage system when it comes to waiting time and therefore norm performance, but also on idle time of triagists and thus patient and employee satisfaction and efficient use of resources, visualized as the second box in Figure 6.37. In Table 6.6, the attempted system changes in this chapter are listed, with the internal system property that is changed. All these experiments had as a primary goal to reduce the waiting times compared to normal model performance. For each of the experiments, a statistical analysis is run to see whether waiting time performance is actually different for system changes compared to the normal waiting time performance without any changes (visible in for example 6.36). In every subsection, one of these system changes is experimented with for the Saturday scenario and it is explained more thoroughly how exactly the system is changed for this experiment in the simulation model. For three other scenarios all experiments are also run, this is visible in Appendix B.

To reach those system changes that prove promising for waiting time reduction in this section, potential system interventions are identified and discussed in Chapter 7. All combinations of values within the variable ranges are run between 10 to 100 times to account for the stochasticity of the model. All attempted internal model changes and additions were programmed in Python and run using the model, in some cases in combination with the Exploratory Modelling Workbench (Kwakkel, 2017).

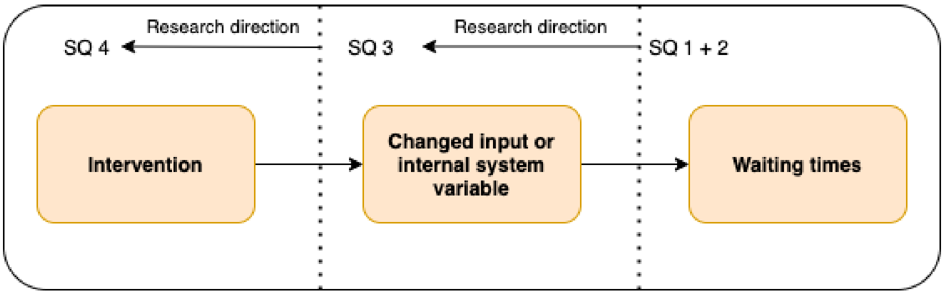

Figure 6.37: Intervention strategies

Table 6.6: Attempted system changes to achieve waiting time reduction

Name of section	**Varied or added input or internal variable**
Shift lengths and capacity optimization	Amount of triagists and length of shift in hours (2 variants)
Decrease in low urgency calls	Low and no urgency patient demand decreased by different percentages
Change in general demand	Empirical demand time distributions increased/decreased by certain percentage
Demand shifting	Peak demand shifted to less busy moment by certain amount of patients
Change in service times	Empirical service time distributions increased/decreased by certain percentage

6.6.1 Shift lengths and capacity optimization

One of the system variables that can be changed to potentially reduce waiting times is the scheduling of triagists and the duration of their work shift. The main question here is whether allocation of the triagists based on the real time performance of the system, constrained by a certain amount of hours that triagists have to work - helps in the reduction of the waiting times at the department. This optimization is tested for different lengths of working shifts to come to a grounded conclusion on the trade-off between working shift length, triagist capacity and system performance. It should be noted that the way the optimization is performed and the thresholds that are chosen greatly influences the results of these optimization runs. Changing the thresholds to increase capacity slightly can change the outcomes of the optimization. The results of the optimization variants, explained below, should therefore be taken into account with that notion.

The optimization has two variants:

1. Increases and decreases in the capacity of triagists can only take place after at least the shift length.
2. Increases of the capacity can take place every hour, decreases can only take place after a triagist has worked the shift length.

This approach was chosen because one can imagine that it might be possible to add an extra triagist hourly, but once a shift is started it is also finished and not stopped early.

The optimization works as follows: the amount of triagists has a minimum value of 1 and a maximum of 7 (in the week version visible in Appendix B this amount is 4 as it is less busy during weekdays). The shift length varies from 1 hour to 8 hours. In the real system, a shift now takes 4 to 8 hours. These conditions are the same for variants 1 and 2. For variant 1, the capacity of triagists can only be increased or decreased after a certain amount of hours have passed since the start of the previous shift, the shift length. For variant 2, this constraint only applies to the reduction of the capacity of triagists, increases are possible every hour if any of the norms are *almost* not met in the past hour in the simulation. If this is the case and the optimization allows for capacity changes based on the shift length, the capacity is always at least increased by 1, plus the amount of people in the waiting line divided by the amount of patients the present triagists can approximately help in an hour, which is assumed to be 5 per hour. This is an assumption based on the amount of high urgency calls triagists can service in an hour, in reality they might help a bit more people in an hour (see Figure 5.26 for average service times of people with an allocated urgency). This means that if there are 25 people waiting and there are 5 triagists, $1 + 25/(5*5) = 2$ triagists are added to the capacity with a maximum of 7. A reduction of the capacity takes place when the shift length allows for it and when the occupancy of the triagists is lower than 0.8 on average in the past hour in the simulation.

There are different scenarios in the model. The capacity and shift length optimization are tested on several scenarios as they behave differently as concluded in Chapter 5 and confirmed by expert and literature interpretation in Section 5.6. The busy Saturday scenario is displayed in this chapter, a busy weekday and less busy days are also simulated with shift optimization: these are displayed in Appendix B.

VARIANT 1 Busy weekend scenario - Saturday
In Figures 6.38 and 6.39 the capacity and the occupancy over the day can be seen when optimizing capacity under different work shift lengths. It can be seen in the left figure that the shorter the work shift, the more the system shifts between triagist capacities because it is allowed to shift more often when a shift length is short. It is also visible that the shorter shift lengths tend to increase capacity to the maximum earlier in the day than the longer shift lengths. This is supported by the occupancy graph, which shows that the occupancy changes more for the lower shift lengths than for the higher shift lengths, because they can increase and reduce their triagist capacity more often according to occupancy. Because the longer shift lengths shift to maximum capacity later, their occupancy stays at the maximum of 1 for a longer time because it was not possible yet to increase capacity at the time when it was already necessary and therefore it takes more time to handle the peak demand and the queues.

In Figures 6.40 and 6.41 it can be seen that the shorter the shift length, the lower the length of stay in the queue is. It can therefore be concluded that if it is possible to shift the capacity of the triagists earlier, the model tends to do that early at the start of the peak which leads to prevention of a really high peak: shift lengths 1 to 3 hours do not reach the really high waiting times that are reached by shift lengths 4 to 8 hours. If it is only possible to shift 1 or 2 times in the day, it is harder to accurately let the capacity follow the behavior of the system which leads to often too little or too much capacity. It can also

Figure 6.38: Capacity of triagists over the day with different shift lengths - Optimization variant 1

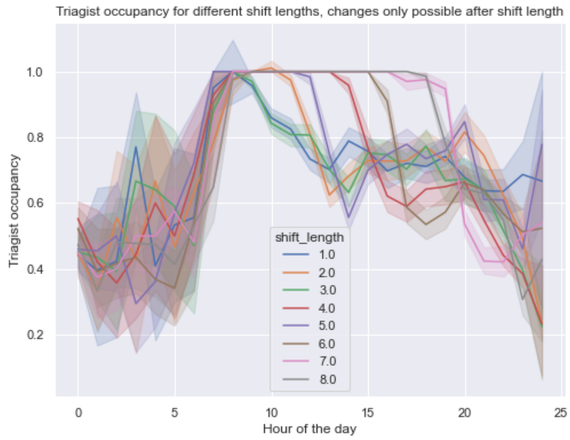

Figure 6.39: Occupancy of triagists over the day with different shift lengths - Optimization variant 1

be seen that the performance on the norms rapidly decreases after the shift length goes beyond 3 hours.

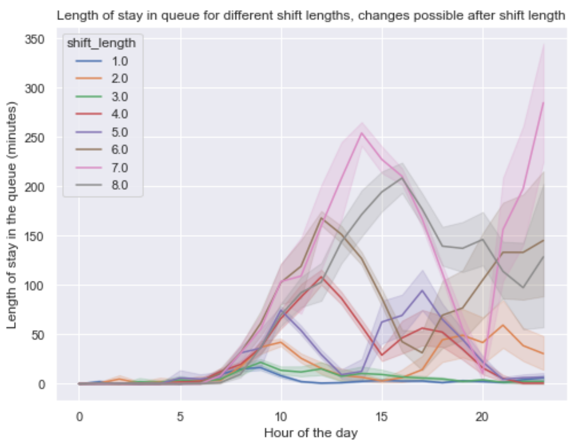

Figure 6.40: Length of stay in the queue over the day with different shift lengths - Optimization variant 1

Figure 6.41: Norm performance over different shift lengths - Optimization variant 1

For variant 1, it can be concluded that it only yields satisfactory performance on the norms and the waiting times when a shift length of 1 or 3 hours is adapted. When comparing the performance of these three shift lengths to the performance of the system now, as visualized in Figures 6.42 and 6.43, it can be seen that shift lengths of 1 and 3 hours perform better in some hours than the current situation, but for 3 hours it sometimes performs slightly worse. Over the whole day, it can however be seen that the peaks in the late afternoon are flattened, reducing pressure on triagists and increasing norm performance. The 2 hour shift is performing worse: this is due to the way the optimization is programmed: the model can probably respond better after 3 hours than after 2 hours in this scenario, leading to lower waiting times when the shift length is set at 3 hours.

STATISTICAL TESTING To statistically verify this notion, the Kruskal-Wallis and posthoc-Dunn test were performed (explained in Section 5.3) to see if the waiting times for the shift lengths are statistically different from the normal performance. Visually, this seems to be

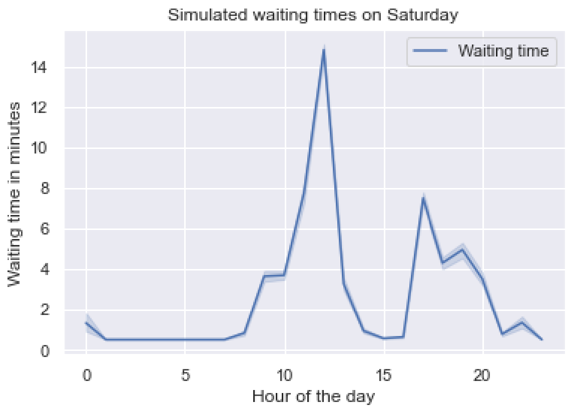

Figure 6.42: Length of stay in the queue regular Saturday real system

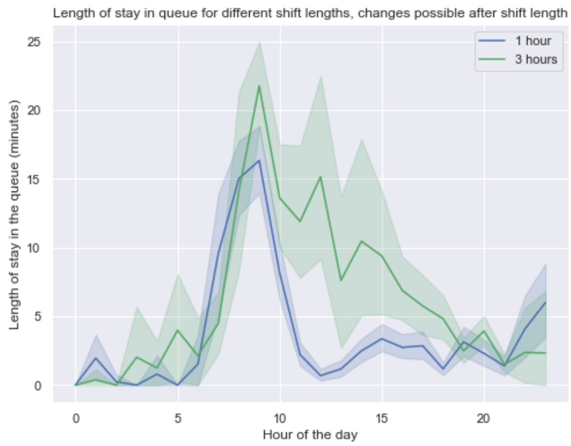

Figure 6.43: Length of stay in the queue for shift lengths 1, 2 or 3

the case for almost all shift lengths as they perform way worse, except for shift lengths 1 and 3 that perform similar or better especially after the first peak, as the waiting times with shift length 1 never increase 5, where in the normal model performance they increase to 8. When running the Kruskal-Wallis test, it is found that all hours starting from 5 in the morning to hour 21 in the evening are statistically different with a p-value below 0.05. When comparing the normal model performance with the performance when shift lengths are 1 hour pairwise with the posthoc-Dunn test, it is also found that the same hours between those two groups are statistically different with a p-value below 0.05. Up and till hour 14 this is visible in Figures 6.42 and 6.43, as the peaks come at a slightly different moment of the day. After hour 14, it indicates that a shift length of 1 statistically significantly performs better than normal model performance. These statistical tests are also performed for the other scenarios, briefly summarized in Appendix B.

CONCLUSION It can be concluded that a shift length of only 1 hour has a statistically significant reducing effect of around 50% on waiting times compared to normal model performance in the afternoons in the weekends, which also goes for the second data-set. For Fridays, the performance of a shift length of 1 is better for the whole of Friday. During weekdays however, the performance of a low shift length is statistically the same or a bit worse than the normal schedule. This can be explained by the way the optimization works, and the fact that during weekdays it is already not that busy and optimizing therefore does not have the intended effect. If waiting times are not that long, an optimization can be seen as redundant. More on possibilities to implement such a short shift length on busy days this can be read in Chapter 7 and 8.

VARIANT 2 The same analysis has been performed when running the second variant of the shift length and capacity optimization in the model, where only reduction of capacity is constraint after the shift length. Again, it was performed for 4 scenarios, 3 of which are visible in Appendix B.

Busy weekend scenario - Saturday

In Figures 6.44 and 6.45, the capacity and the occupancy over the day can be seen when optimizing capacity under different work shift lengths for variant 2. It can be seen that the triagist capacity lays very close to each other for all shift lengths when increases can take place every hour for all shift lengths: this makes sense as this is not constrained between shift lengths. Again, the maximum of 7 is often reached very early, but capacity reduction also takes place a lot earlier as more shifts end every hour when it is possible to start shifts every hour. This is supported by the occupancy graph where only for a short while an occupancy of 1 is reached for all shift lengths.

Figure 6.44: Waiting time over a weekend day with warm up time

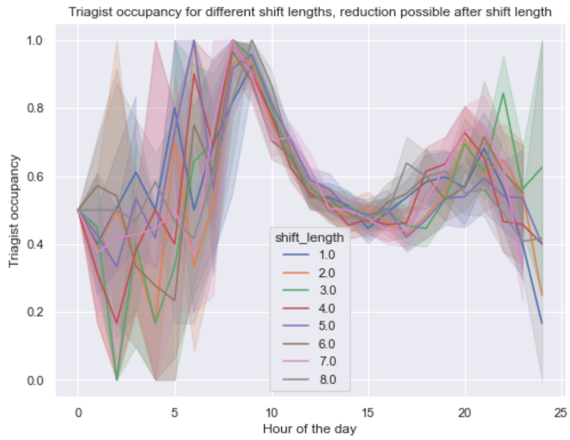

Figure 6.45: Waiting time over a weekend day no warm up time

In Figures 6.46 and 6.47 it can be seen that the performance on the waiting times increases greatly compared to Variant 1, where waiting times of 200 are reached. The shift lengths are not that different anymore when it comes to waiting times. It can be seen in the right figure that in variant 2 shift length does not have the big effect on the norm performance that it has for variant 1.

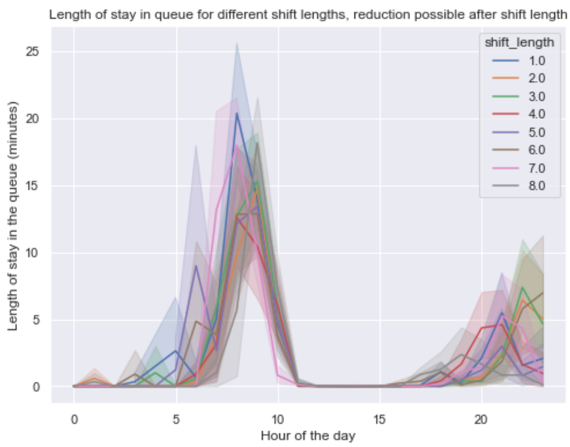

Figure 6.46: Waiting time over a weekend day with warm up time

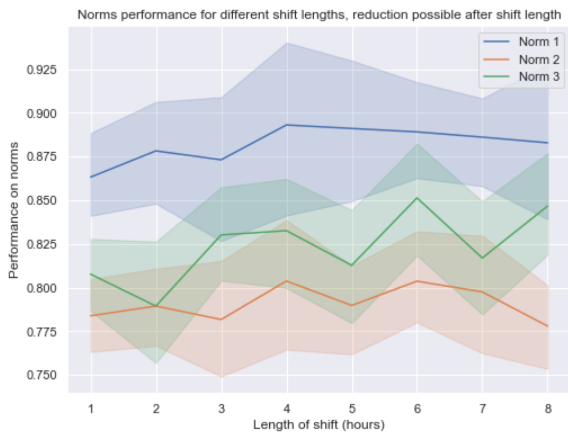

Figure 6.47: Waiting time over a weekend day no warm up time

For variant 2, it can be concluded that satisfactory results are achieved for most of the shift lengths when it is possible to increase capacity every hour, but it is only possible to decrease capacity after a triagist has worked the full shift. When comparing the performance of the length of stay in the queue to the situation now, which was already displayed in Figure 6.42, it can be seen that some of the shift lengths achieve better results than the current system. Similar to variant 1, it can also be seen that to achieve this, 7 triagists are needed. However, the amount of time that 7 triagists are necessary tends to be lower in variant 2 than in variant 1, indicating that the system responds better to increases in demand.

STATISTICAL TESTING To statistically verify that some shift lengths in this variant perform better, the Kruskal-Wallis and posthoc-Dunn test were performed (explained in Section 5.3) to see if the waiting times for the shift lengths are statistically different from the normal performance. Visually, for almost all shift lengths a reduction in waiting times seems to take place, either in the first or second peak or in both. Also, the confidence intervals of almost all shifts partly overlap, indicating that they are statistically the same. When running the Kruskal-Wallis test, it is found that when comparing all shift lengths and the normal performance at the same time, all hours starting from 8 in the morning to hour 21 or 22 in the evening are statistically different from each other with a p-value below 0.05. When comparing the normal model performance with the performance of all shift lengths pairwise with the posthoc-Dunn test, it is also found that the same hours between those two groups are statistically different with a p-value below 0.05. This can be confirmed visually: the first peak in the normal performance is moved to a slightly earlier point in the day and has approximately the same height or slightly higher than normal performance for all shift lengths, indicating not that much of a difference in total waiting times. However, the second peak of the day is lower, indicating a statistically significant waiting time reduction for most shift lengths. This is beneficial compared to variant 1, as longer shift lengths are possible for triagists, which is easier to implement. These statistical tests are also performed for the other scenarios, briefly summarized in Appendix B. More on possibilities to implement optimized start moments of shifts can be read in Chapter 7 and 8.

CONCLUSION It can be concluded that with this type of optimization, it is beneficial for the system to increase triagist capacity every hour for the busy Friday and Saturday scenarios, where higher shift lengths are possible. This could be interesting for the out-of-hours department, as shifts of 4 or 8 hours might then still be possible. A strategy has to be defined on the best shift length that triagists are comfortable with and that leads to acceptable performance, and on the best way to have triagists on standby for a certain amount of hours. The less busy scenarios like a normal weekday have approximately the same performance when optimizing, therefore for these days it is not necessary. More on the implementation of this system change can be read in Chapter 7 and 8.

6.6.2 Decrease in low urgency calls

A next possibility for the reduction of waiting times is the reduction of low urgency calls. Patient calls with subprocess 1, no full triage, or with urgency level 5 in subprocess 2 are reduced from the total demand by a certain percentage: from 0% reduction to 50% reduction. The idea for this change in demand comes from the notion from (Verzantvoort et al., 2018) that 65% people using the "Moet ik naar de dokter?" app would follow the advise of the app, thus reducing the amount of low urgency calls, as some people don't call the department anymore with their problem when the app indicates that it is not necessary. Testing this in the simulation model can give quantitative evidence for implementing such an intervention to reduce waiting times. Again, this change of the system was run for 4 scenarios, 3 of which are displayed in Appendix B.

In Figures 6.48 and 6.49, the amount of low urgency calls (U5) and no urgency calls (subprocess 1) handled by the triagists in a normal model situation are compared to the amount handled by triagist when a certain percentage of the low urgency calls is reduced. In the first Figure, the low urgency call reduction is **high**, therefore the amount of low or no urgency patients helped by triagists is a lot lower. In the second Figure, the reduction is **low**, and the amount of patients helped by triagists is almost similar to the total amount of low or no urgency patients.

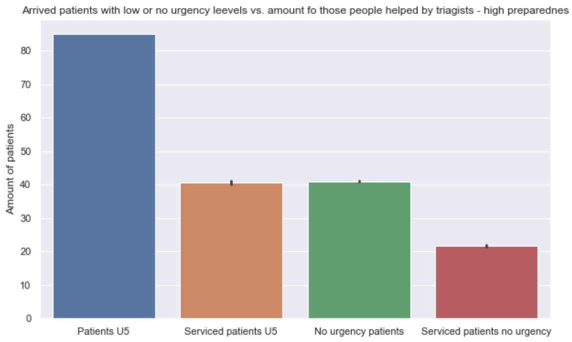

Figure 6.48: Total patients with low and no urgency level compared to serviced patients by triagists - high reduction

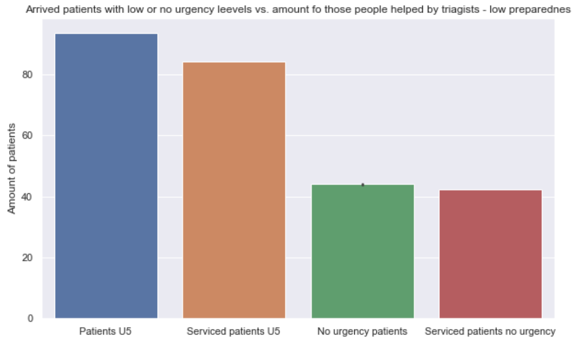

Figure 6.49: Total patients with low and no urgency level compared to serviced patients by triagists - low reduction

In Figure 6.50, it can be seen that the higher the low urgency call reduction, the lower the lengths of stay in the queue are: waiting times are reduced by at least 20% starting from a low urgency call reduction of just over 10%. This is due to the decrease in low urgency patients that have to be serviced by triagists. This indicates a positive effect on the norm performance and on the length of stay in the queue before being serviced compared to the normal system performance.

STATISTICAL TESTING To statistically verify this notion, the Kruskal-Wallis and posthoc-Dunn test were performed (explained in Section 5.3) to see if the waiting times for the different low urgency call reductions are statistically different from the normal performance when there is no reduction. Visually, this seems to be the case especially in peak hours, as the confidence intervals mostly do not overlap with the normal performance. When running the Kruskal-Wallis test, it is found that for all hours where waiting times are not 0, so hours 9 up and till 23 excluding hour 14 and 15, the performance of the reductions can be seen as statistically different from the normal performance as the Kruskal-Wallis p-value is below 0.05 for all these hours, and comparing the normal performance with the 10% reduction pairwise with the posthoc-Dunn test also gives p-values below 0.05 for hours 9 till 22 again excluding 14, 15 and also 10. When looking at the figure, this indicates that for hours 16 up and till 22, a small reduction of 10% already yields a statistically significant reduction of waiting times of 20%. These statistical tests are also performed for the other scenarios, briefly summarized in Appendix B.

CONCLUSION It can be concluded that a 10% reduction of low urgency calls statistically significantly reduces waiting times in the weekend days in the second peak of the day from hour 16 till 22 by at least 20%. In Appendix B, the results of decreasing low urgency demand in other scenarios are displayed. In these scenarios, similar effects as in the Saturday and Sunday scenario can be seen: in the Friday scenario, a small reduction even leads to more than 20% waiting time reduction. For weekdays, a larger reduction of 20% is necessary to have a significant effect due to the fact that it is less busy on those days. When implementing a low urgency call reduction with the data and scenarios of the second dataset as analyzed in Appendix A, a 20% reduction is found to be necessary for a waiting time reduction especially in the afternoons. This indicates that low urgency call reduction has

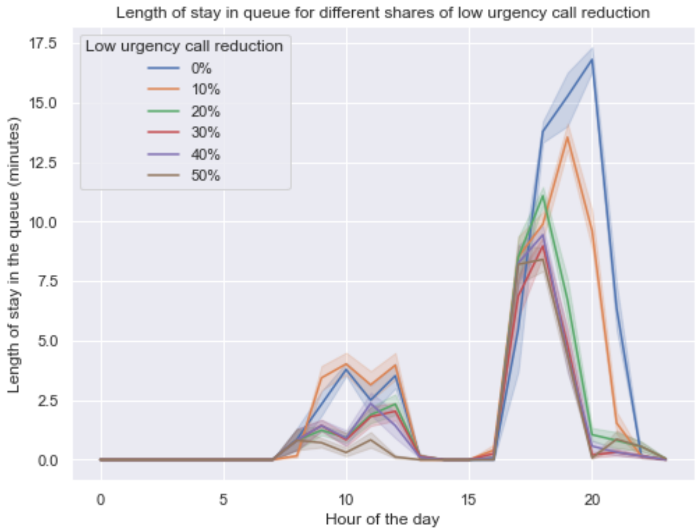

Figure 6.50: Waiting time with different levels of low urgency call reduction

similar results over scenarios and has a similar effect on another out-of-hours department with the same type of system.

6.6.3 Change in general demand

Next to only reducing low urgency demand, it is interesting to see what an overall reduction or increase in demand does to the waiting times in the system. If a small reduction leads to less needed triagist capacity and better norm performance, it is interesting to look at interventions which could lead to demand reduction. In Figures 6.51 and 6.52, the effect of certain fractions of increase and decrease in demand (*value_increase*) on the waiting time in minutes and on the demand are displayed. It can be seen that when the demand is decreased, the waiting times and the demand go below the waiting times of the normal system (where *value_increase* is 0), and it can thus be seen that a decrease of demand leads to better system performance: already a small decrease of 10% leads to a waiting time reduction of over 50%. It can also be seen that only small amounts of extra demand lead to way worse performance of the system, in such a way that the simulation stops running because of the length of the queue, which never solves over the course of the day.

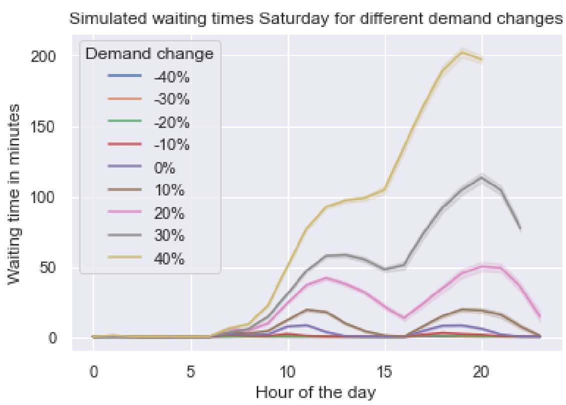

Figure 6.51: Waiting time with different levels of demand increases and reductions

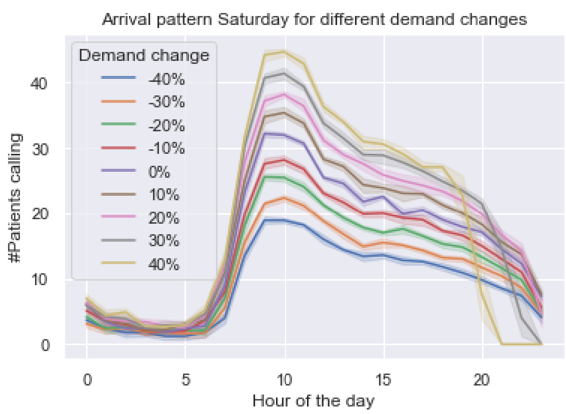

Figure 6.52: Demand with different levels of demand increases and reductions

STATISTICAL TESTING To statistically verify this notion, the Kruskal-Wallis and posthoc-Dunn test were performed (explained in Section 5.3) to see if the waiting times for the different demand reductions are statistically different from the normal performance when there is no reduction. Visually, this seems to be the case especially in peak hours, as the confidence intervals do not overlap with the normal performance. When running the Kruskal-Wallis test, it is found that for all hours where waiting times are not 0, so hours 0 up and till 20, the performance of the reductions can be seen as statistically different from the normal performance as the Kruskal-Wallis p-value is below 0.05 for all these hours, and comparing the normal performance pairwise with the 10% reduction variant with the posthoc-Dunn test also gives p-values below 0.05 for hours 8 up and till 22: indicating that in the peak waiting times the reduction is statistically significant. These statistical tests are also performed for the other scenarios, briefly summarized in Appendix B.

CONCLUSION It can be concluded that small increases or reductions in demand have a statistically significant disruptive or very positive effect on the system. Increases can lead to multiple hours of waiting time which is absolutely unacceptable in a healthcare system, but on the other hand, a small demand reduction could lead to the flattening of waiting time peaks and should therefore be aimed for with the use of an intervention: small decreases of 10% lead to a reduction of at least 50% of waiting times for most hours of the day in all scenarios.

6.6.4 Demand shifting

It might well be the case that a reduction of demand is hard to realize with an intervention, while it might be possible for triagists to shift (low urgency) demand to another time of the day, especially on weekend days when the department is open all day. This is implemented in the model by taking a few callers, varied from 0 to 5, from each hour in peak demand and moving those to a time of day where there is less demand but which is close to the time the patient called in the first place. For a Saturday, this peak demand is between 9 and 11 AM (this can be seen in for example Figure 6.52) and the callers are moved to the demand between 13 and 15 PM, when it is less busy. In Figures 6.53 and 6.54 the effect of demand shifting can be seen for various amounts of shifted patients. It is clearly visible that the waiting time peak when no shifting takes place, so when *shift* is 0, is higher than when patient shifts take place and the peak moves to the indicated time between 13 and 15 PM. A shift of 1 patient can already reduce waiting times by 20%.

STATISTICAL TESTING To statistically verify this notion, the Kruskal-Wallis and posthoc-Dunn test were performed (explained in Section 5.3) to see if the waiting times for the different demand shifting amounts are statistically different from the normal performance when there is no demand shifting. Visually, this seems to be partly the case especially in peak hours, as for some of the shifts the confidence intervals do not overlap with the normal performance without shifting (shift = 0 in the Figure). When running the Kruskal-Wallis test, it is found that for the hours where you can also visually verify that the confidence intervals don't always overlap, so hours 7 up and till 21, excluding 16 where waiting times are the same for all demand shifts, the performance of demand shifting can be seen as statistically different from the normal performance as the Kruskal-Wallis p-value is below 0.05 for all these hours. Comparing the normal performance pairwise with the variant where 1 patient is shifted per hour in peak demand with the posthoc-Dunn test gives p-values below 0.05 for hours 10, 11 and 12 and hours 18, 20 and 21 indicating that in the

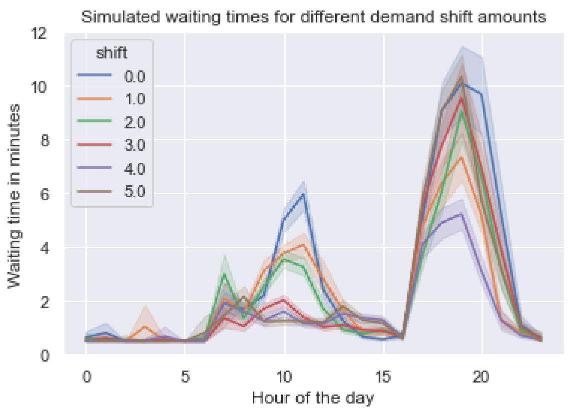

Figure 6.53: Waiting time with different levels of demand increases and reductions

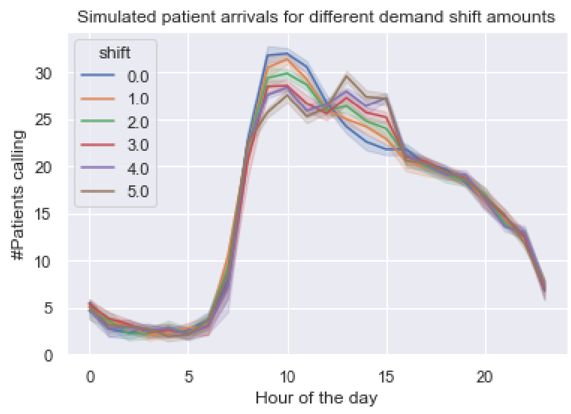

Figure 6.54: Service times with different levels of demand increases and reductions

peak waiting times shifting 1 patient per hour to less busy periods is statistically significant. This result is also visible in the Figure, where you see that in those hours, the orange line is below the blue line and the confidence intervals don't overlap. Doing this pairwise comparison for all other shifts yields similar results for the earlier hours, but does in some cases not have the same results for the later hours, for example for shifts of 2 or 3 patients. It is therefore important that when implementing demand shifts, the simulation model is used to see what amount of shifting will work best in that specific scenario and to be confident that a waiting time reduction will take place. These statistical tests are also performed for the other scenarios, briefly summarized in Appendix B.

CONCLUSION It is very promising to see that it can be statistically proven that a small amount of patient shifting can lead to a waiting time reduction of already 20%. This is beneficial for norm performance, but also for the workload on triagists. In Appendix B, demand shifting is tested for other scenarios and for a scenario of the other data-set. It is concluded that for the busy Saturdays for both departments, demand shifting has the most promise to reduce waiting times by at least 10-20%, possibly increasing to at least 50% for higher amounts of shifted patients. For the weekdays, a waiting time reduction can be seen when shifting a patient (up to 50% for weekdays, around 10% for Fridays), but its effect is less than for weekend days. This could be due to the fact that demand is not that low in the hours that the patients are shifted to, but shifting to a later moment leads to shifting patients to the night, something that is not desirable. It should be noted that demand shifting is very sensitive and when implementing it with an intervention, the simulation model needs to be used to see beforehand what amount of shifting will work best in that specific scenario and to be confident that a waiting time reduction will take place: there can be a difference already between shifting 1 or 2 patients. Moving too many patients might only lead to another time of the day where the waiting time peak takes place, in stead of flattening the peaks and spreading workload.

6.6.5 Change in service times

Lastly, it is analyzed whether changing service times have a beneficial effect on system performance to see whether interventions leading to this reduction have to be identified. In Figures 6.55 and 6.56 it can be seen that increases of the service time lead to very disruptive effects on the waiting time in the system, similar to what happens when demand increases in Section 6.6.3. Reductions in service time lead to better performance than the system now: a reduction of 10% already leads to a waiting time reduction of 50%. It is also interesting to see that the variation of service times is higher during the night than during the day in the right figure. This is due to the higher share of high urgency levels calling at night, whose conversations tend to take longer. Also, no pressure to finish a conversation lays on the triagists shoulders as there are almost never waiting times during the night.

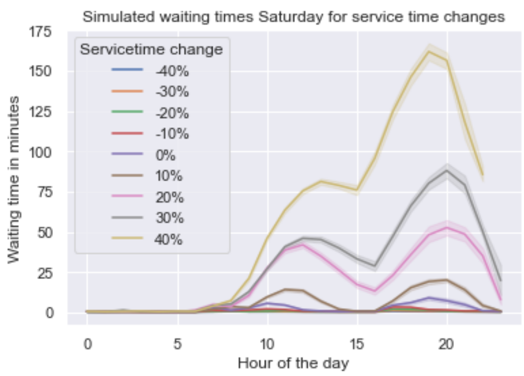

Figure 6.55: Waiting time with different levels of service time increases and reductions

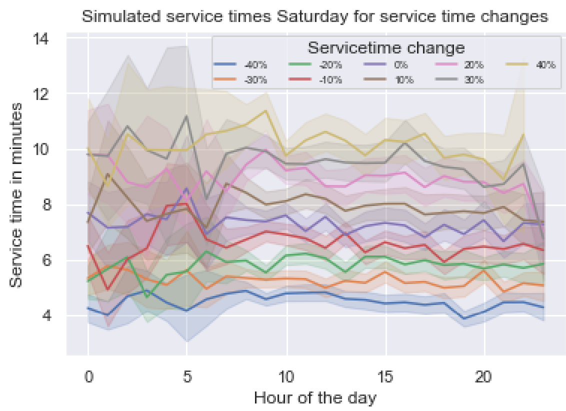

Figure 6.56: Service times with different levels of service time increases and reductions

STATISTICAL TESTING To statistically verify the notion that a small service time reduction already gives a big waiting time reduction, the Kruskal-Wallis and posthoc-Dunn test were performed (explained in Section 5.3) to see if the waiting times for the different service time reductions are statistically different from the normal performance when there is no reduction. Visually, this seems to be the case especially in peak hours, as the confidence intervals almost do not overlap with the normal performance. When running the Kruskal-Wallis test, it is found that for all hours where waiting times are not 0, so hours 7 up and till 22, the performance of the reductions can be seen as statistically different from the normal performance as the Kruskal-Wallis p-value is below 0.05 for all these hours. When comparing the normal performance pairwise with the 10% reduction variant with the posthoc-Dunn test also gives p-values below 0.05 for hours 8 up and till 22: indicating that a small service time reduction is already significantly different from the normal model performance. These statistical tests are also performed for the other scenarios, briefly summarized in Appendix B.

CONCLUSION It can be concluded that an increase of service times should be avoided as it disrupts the waiting time performance of the system. This is a very important implication, as service times have been increasing the past years (see Figures 5.15 and 5.16 for this increase) which is now quantitative proof for the increased pressure that triagists feel. Also, it makes the increases that are tested here possible scenarios for the future and something that should definitely be avoided. On the other hand, only a small decrease of 10% leads to a reduction of at least 50% of waiting times for most hours of the day in all scenarios.

6.6.6 Overall conclusion

Overall, it can be concluded that some of the system changes, after testing for statistically significant differences from normal model performance, prove promising for model performance:

- For the shift and capacity optimization option, it was found that shift lengths of 1 hour prove promising in variant 1 especially in weekends and on Friday, reducing waiting times with up to 50% in the weekend afternoons and a reduction of the second peak on Friday of also up to 50%. The longer shift lengths prove promising for variant 2 for the busy Fridays and Saturdays, giving a waiting time reduction of 10% that can increase to around 50% in the Saturday scenario when a shift takes over 4 hours.

- For low urgency call reduction, it was found that it proves promising for demand and waiting time reduction: a low urgency reduction of 10% on a Saturday or Sunday already gives a 20% reduction of waiting times in the afternoon peak, which can increase to a 30% reduction when low urgency calls are also reduced by 30%. For Fridays, the waiting time reduction is even bigger, around 50%, with a low urgency reduction of 10%. For weekdays and for the second department, an increase of 20% is necessary for statistically significant waiting time reductions of around 20%. In conclusion, a 10 or 20% reduction has a significant effect in all scenarios.

- For the demand and service time change options, it was found that small increases in demand and service time have disruptive effects on the system and need to be avoided, but similarly, small decreases of 10% lead to a reduction of at least 50% of waiting times for most hours of the day in all scenarios.

- For the demand shifting option, it was found that shifting already 1 patient from peak demand on Saturdays for both departments can lead to a waiting time reduction of at least 10-20%, possibly increasing to at least 50% for higher amounts of shifted patients. For the weekdays, a waiting time reduction can be seen when shifting a patient (up to 50% for weekdays, around 10% for Fridays), but its effect is less than for weekend days. It should be noted that using the simulation model is very important when implementing an intervention that leads to demand shifting, as waiting times are very sensitive to it and shifting too many patients only leads to shifting in stead of reduction of waiting time peaks.

With the identification of these options, subquestion 3 is answered. In Chapter 7, it was explored which (qualitative) interventions could lead to these changes and conditions within the system.

7 RESULTS: SYSTEM INTERVENTIONS

In this chapter, the last phase of the research is addressed. In Figure 7.1, it is visible that in subquestion 3 (answered in Chapter 6) system changes that lead to waiting time reductions are identified. This leaves the question what has to happen within the out-of-hours department's power - or out of its power - to reach those system changes: the interventions. They are identified and discussed in this chapter based on consultations with three out-of-hours departments in the Netherlands, one of which this research was not performed for. Also, interventions are identified by looking at the identified system improvements and policies in the literature review and by looking at new literature. Figure 7.1 is extended to display all interventions and system changes at the end of this chapter, in Figure 7.3.

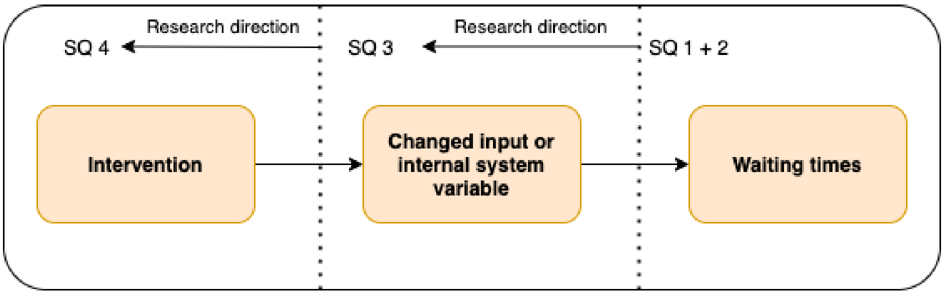

Figure 7.1: Intervention strategies

7.1 SUBQUESTION 3: DESIRED SYSTEM CHANGES

Below, the system changes that were found to reduce waiting times, identified in Section 6.6, are listed. In the next section, the possible interventions which could lead to those system changes are identified and discussed.

1. Shift and capacity optimization option: low shift lengths for variant 1
2. Shift and capacity optimization option: all shift lengths for variant 2
3. Low urgency demand reduction
4. Demand and service time change: prevention of increases, incentives for decrease
5. Demand shifting: shift small amounts of patients from peak demand to less busy period

7.2 SUBQUESTION 4: POSSIBLE INTERVENTIONS

System interventions that might lead to these desired conditions or system changes are identified in this Section. First, interventions that could potentially lead to the identified system changes that reduce waiting times are discussed for each of those conditions. Next, a visualization of how all interventions interplay with the desired system changes and waiting time reductions is displayed. It should be noted that the simulation model can only test and quantify the results of internal system changes, as done in Section 6.6 and that the real effect on the out-of-hours system can only be tested after implementation by the departments. Next, the decision arena in which the out-of-hours system operates is defined, as well as other stakeholders which influence the performance of the out-of-hours system. The identified interventions are placed within this decision arena to get an overview on what is within the power of the department itself and what isn't. A short conclusion is given on the two types of identified interventions and how out-of-hours departments can use them.

In Table 7.1, an overview of the interventions discussed in this section and the source that they come from is given. This can be from literature, from expert consultations or is a contribution made during this research, inspired by literature in other research areas or by knowledge of the system itself.

Table 7.1: Sources of interventions

Intervention	Source
Financial (dis)incentive	Literature (Anantharaman, 2008; M. J. Giesen et al., 2017; Keizer et al., 2016)
Increased overall health literacy	Literature (Anantharaman, 2008; Van den Heede & Van de Voorde, 2016)
	Literature/Expert consultations
	(Heutmekers et al., 2017; Van den Heede & Van de Voorde, 2016)
Case management	(Jansen et al., 2018; Keizer et al., 2017)
Increased general practitioner accessibility	Literature/Expert consultations (Keizer et al., 2017, 2021)
Flexible schedule policy	Literature/contribution from this research (Boyle, Beniuk, Higginson, & Atkinson, 2012)
Retrieval of patient information in the queue	Expert consultations
Extra audio fragment in peak demand	Expert consultations
Possibility for triagists to work from home	Expert consultations
Separate telephone lines for nursing homes and home care	Expert consultations

7.2.1 Shift and capacity optimization variant 1

It was found that when optimizing triagist capacity in variant 1, so where triagists can only be changed after a certain shift length (visible in Section 6.6, waiting times are reduced when triagists can change shifts every 1 to 3 hours for almost all scenarios. Waiting times increase when the shifts take longer than that, as the optimization cannot increase capacity often enough to handle changes in waiting times and norm performance anymore.

Intervention: flexible schedule policy or working from home
An intervention is therefore needed which implements a new shift policy with shifts of 1, 2 or 3 hours and which uses the simulation model with the optimization settings to create the schedules. It is questionable whether an intervention like this would be accepted by out-of-hours departments, as it is not at all beneficial for employees to work such short shifts and to have irregular schedules: similar scheduling models have been proposed in emergency departments, where they often trigger resistance because of the personal lives of staff, reducing employee satisfaction (Boyle et al., 2012). When speaking to the three out-of-hours departments, it was discussed that they are interested to enable working from home as a triagist somewhere in the future. This could be an interesting way to implement shorter shift lengths that takes away some of the objections triagists might have against them, such as travel time.

7.2.2 Shift and capacity optimization variant 2

For variant 2, where capacity can be increased hourly but reduction can only take place once a shift has ended, it was found that for any shift length the performance of the system increases and that it is similar over all shift lengths. To implement variant 2, shift lengths of 4 to 8 hours can be kept, similar to the system now. A new shift policy has to be implemented within the system to accommodate the possibility to hourly increase triagist capacity and create schedules accordingly. The willingness of triagists to start at different times of the day than they are used to has to be identified, and schedules have to be made accordingly with the use of the simulation model. It might however be easier to implement than variant 1 as triagists know that they will work a shift with a length that they are used to and working from home is not yet possible at the out-of-hours departments. Also, it has to be checked financially whether it is possible to use optimized triagist capacities over the whole range of the day, especially when the waiting times have shrunk and the optimized capacity is not needed anymore.

7.2.3 Low urgency demand reduction

It was found that reducing low urgency demand can greatly reduce the waiting times and the amount of low or no urgency patients serviced by triagists. To achieve low urgency demand reduction, at first an app or website like "Moet ik naar de dokter?" ("Should I see a doctor?") comes to mind. However, when speaking to the three out-of-hours departments, it became clear that they feel like it keeps the amount of low urgency calls the same if not higher, due to the website and app being so conservative and still sending almost all patients to the out-of-hours department. Even though in (Verzantvoort et al., 2018) it was found that 65% of people who tried the website or app would consider following the advise of the app or website, and in M. J. Giesen et al. (2017) it was found that online advise has a high potential to reduce unnecessary use of out of hours services, if they are too conservative the demand will not decrease. The solution seems to make these apps less conservative, but this is a risky choice when talking about medical cases and peoples lives and can raise ethical concerns. The observation on the effectivity of this website for demand reduction that the out-of-hours departments make and how it might be improved for the better has to be verified with further research.

Intervention: financial (dis)incentives

Another intervention to reach low urgency demand reduction is the implementation of financial (dis)incentives for the use of healthcare services. In Anantharaman (2008) it was found that financial incentives can have an impact on the non-urgent use of emergency department care if it costs patients more than going to a primary health service. In the Netherlands, they are both free, so a small disincentive could potentially lead to a reduction in use of out-of-hours care. In the literature review, in M. J. Giesen et al. (2017) it was found that presenting a cost-overview before making use of the department is an incentive for less use of the out-of-hours system, and in Keizer et al. (2016) it was found that co-payment, letting people pay a certain part of the costs for using the out-of-hours department is also an incentive to reduce the use of out-of-hours care. Financial incentives are implemented relatively easy, without the need for internal system changes. However, increasing the costs for healthcare is not up to out-of-hours departments alone. To implement financial incentives, cooperation is required with other healthcare institutions and insurance companies.

7.2.4 Demand and service time changes

It was found that increases in demand and service times are disruptive for the system, while reductions can lead to flattening of the waiting times peaks and better performance on norms. In literature, some interventions on healthcare systems are identified that led to reduction of demand for healthcare. Potentially interesting ones to out-of-hours departments are discussed here.

Intervention 1: Financial (dis)incentives
The first intervention for demand reduction is, similar to low urgency demand reduction, financial (dis)incentives as described right above.

Intervention 2: Case-management
A second intervention is active case-management. In emergency departments, it is estimated that 1-5% of patients comes to the department often (Van den Heede & Van de Voorde, 2016), and active identification and case-management of those people can have positive effects on the use of emergency department care when thoroughly and continually performed. In the literature review in Heutmekers et al. (2017), it was also found that their are certain groups, for example with intellectual disabilities, who call out-of-hours departments more often with low urgency calls. Case-management for these types of patients is therefore beneficial for the system. The consulted out-of-hours departments also indicate that this is something they do on a small scale as they do have data on how often a patient calls and that can therefore be done on a larger scale, for example in collaboration with the regular general practitioner of the patient. This can result in better healthcare service to the patient and to a reduction of their use of the out-of-hours department.

Intervention 3: Increase of overall health literacy
Another intervention identified in literature, but focused on the reduction of emergency care use, is large scale public health education campaigns, which in Singapore led to a decrease in non urgent emergency department use from 57% to 18% in a time frame of 12 years (Anantharaman, 2008). Similarly, in Van den Heede and Van de Voorde (2016), it was found that large scale education on health literacy and the use of healthcare facilities could potentially have a large effect on emergency department use. In the literature review of this thesis, similar education recommendations were identified. It was recommended to better educate people on the purpose of out-of-hours care and to improve overall health literacy of fragile groups (Jansen et al., 2018; Keizer et al., 2017). This intervention requires cooperation between healthcare institutions and governmental organizations depending on the scale of the campaign. A smaller campaign held by out-of-hours departments and the general practitioners could be a start, but it should be considered that in Van den Heede and Van de Voorde (2016) it is stated that educational interventions help better when among other interventions: it was found that one-time educative measures such as a booklet did not help in the reduction of the use of an emergency department. This could indicate that one-time measures might also not be of great use for the reduction of the use of out-of-hours care, but combinations with other interventions such as case-management to educate targeted people or by creating a large scale educational healthcare campaign like in Anantharaman (2008) could potentially help reduce demand.

Intervention 4: Separate telephone lines for nursing homes and home care
Another possibility to reduce service times is to implement a separate telephone line for nursing homes and home care. The consulted departments indicate that they do not have a

clear image of what percentage of the calls are of this kind, but they know that it accounts for quite a portion of the calls when it is already busy. In cooperation with nursing homes and home care, a separate staff member specialized in this type of care and with knowledge of the patient population can then be assigned to handle these types of calls in a certain time frame of the day, possibly reducing the amount of time it takes to handle them.

Intervention 5: Increased accessibility of regular general practitioners
A last intervention that could reduce demand for out-of-hours care and which is often mentioned in literature is increasing the accessibility of regular general practitioners (Keizer et al., 2017, 2021). This is an intervention that is mostly up to the general practitioner department, but where the consulted out-of-hours departments feel like they can also play a cooperating role. People with a lower socioeconomic status often have problems with contacting their own general practitioner and find it easier to contact the out-of-hours service. General practitioners have to make sure that they are easily accessible for patients, for example by having broad opening times for the scheduling of an appointment in stead of the two hours a day that it is for some departments, leading to people waiting to call the out-of-hours department all day. Also, new patients should be contacted to give them an introduction on how the regular general practitioner department can be accessed. This intervention can be initiated by the out-of-hours departments with a few of the general practices they serve, to see if it affects the amount of patients of those practices that call the out-of-hours department.

Intervention 6: Retrieval of patient information in the queue
Interventions to reduce service times are not often found in literature, but are among the most important system changes to achieve: the past few years, service times in the analyzed out-of-hours department has been increasing, which was quantitatively found to have a disruptive effect on the waiting times (see Section 6.6. One of the interventions that could possibly lead to service time reduction is the registration of patient information while they are in the queue to reduce administrative tasks during and after a call. One of the three consulted out-of-hours departments is planning to implement registration of patient information in the queue soon, by asking for peoples social security number, as people often do not have it ready yet when calling. This takes much time from a triagist during a call. Another thing that could be interesting is to start measuring which of the calls were video calls, as these often take more time and also often prevent a real visit to the department, which was not researched in this thesis.

7.2.5 Demand shifting

It was found that shifting a few patients from peak demand to a less busy period a few hours later reduces the peak waiting times and creates more stable demand and therefore workload for the triagists over the course of the day.

Intervention: extra audio fragment in peak demand
An intervention to achieve demand shifting is the addition of an extra audio fragment in peak demand to the fragment that is played in the queue, informing patients that they are calling in peak hours of the day and that they should call back later when demand is lower if it can wait. The systems are already in place, and adding audio fragments based on demand conditions are, according to the consulted out-of-hours departments, definitely possible in the current systems. The departments however indicate that it is key that the patients call back themselves, so that this task does not also lay on the shoulders of the

triagists, as that resulted in stressful situations in the past. Also, boundaries have to be set for the use of demand shifting: a minimum amount of people have to be in the queue, triagist capacity at the time of the day that the demand is shifted to should be taken into account and it should not be busy at the time of the day where the demand is shifted to. The simulation model can be used to foresee what different amounts of shifted patients will do to waiting times.

7.2.6 Decision arena

In Figure 7.2, the inner ring shows the interventions which can be implemented by the out-of-hours department itself. The two outer rings show interventions that can be implemented by general practitioner departments and by the government. Some interventions overlap between stakeholders, indicating the involvement of all these stakeholders.

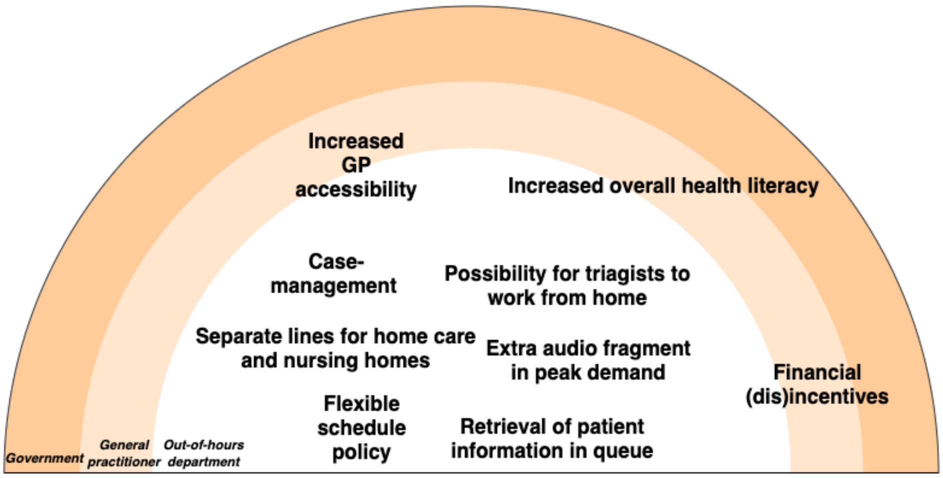

Figure 7.2: Decision arena: Influential decision layers on out of hours care with useful interventions

7.3 CONCLUSION

Overall, it can be concluded that different kinds of interventions are possible to reach the system changes that were identified in Section 6.6, visible in 7.2. In 7.3, the flow of impact from interventions, identified in this chapter, towards the system changes that in their turn reduce waiting times is visualized. It should be noted that self triaging with a website or app is not included in this figure as the out-of-hours departments indicated that it is unclear whether it leads to low urgency demand reduction and therefore to waiting time reduction as intended.

The identified system interventions are divided into three categories: quick win interventions that out-of-hours departments can implement themselves and whose effectivity can be measured almost directly, long term interventions often focused more on the behavioral change of people for which cooperation with other stakeholders is necessary and measurements need to take place over a longer period of time, and interventions that are not practically implementable based on a consultation with three out-of-hours departments. In Table 7.2 all system changes with the corresponding interventions and its category are found.

By ending this Chapter, subquestion 4 is answered which concludes the research. Next, the discussion and conclusion are found, in which the implications from this research are discussed, the main research question is answered and recommendations for further research are made.

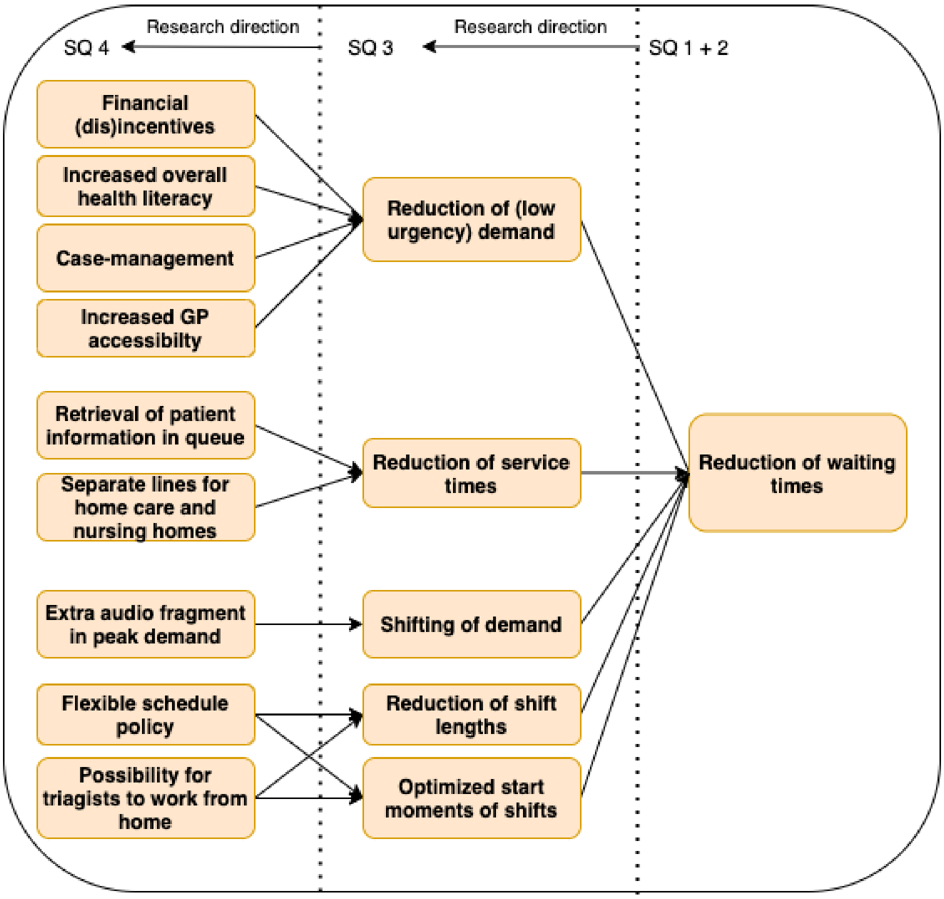

Figure 7.3: Effect of interventions on system changes and waiting time reduction

Table 7.2: System changes and interventions that can make them happen

System change	Intervention	Category
Reduction of shift lengths	Flexible schedule policy	Not practically implementable
Reduction of shift lengths	Possibility for triagists to work from home	Long term intervention
Optimized start times of shifts	Flexible schedule policy	Quick win
Optimized start times of shifts	Possibility for triagists to work from home	Long term intervention
Shifting of demand	Extra audio fragment in peak demand	Quick win
Reduction of (low urgency) demand	Financial (dis)incentive	Long term intervention
	Increased accessibility and knowledge of primary healthcare system	Long term intervention
	Monitoring of frequent users of out-of-hours care	Long term intervention
	Separate telephone lines for nursing homes and home care	Long term intervention
	Use of website or app for self-triaging	Not practically implementable (too conservative)
Service time reduction	Personal patient information retrieval in queue	Quick win

The practical implications of these interventions can be found in Section 8.2.2.

8 DISCUSSION

This chapter covers different reflections on the performed research. The data analysis, the identified scenarios, the modelling approach, the simulation model and its results are reflected on by discussing (model) assumptions and limitations as well as the transferability of the results. Also, the practical and literature implications of the results are discussed.

8.1 CRITICAL RESEARCH ASSUMPTIONS & LIMITATIONS

When performing modelling research, assumptions are made throughout the whole process, sometimes leading to model limitations. In this Section, the assumptions and limitations of the results of the system data analysis as conducted in Chapter 5 and those of the model as implemented and used in Chapters 6 and beyond are discussed and reflected on.

8.1.1 System data analysis

During analysis of the system data where different demand and service time scenarios were identified, several assumptions were made. Many of these assumptions resonate throughout the rest of the research, as almost all of the model input and internal variables are derived from data.

Assumption 1 A first assumption is that the telephone triage system of an out-of-hours department only receives calls from patients, where in reality also a portion of the calls is within the department itself, to or from other healthcare departments such as nursing homes or home care or for example to ambulances or visiting general practitioners.

Limitation 1 This means that it can be the case that a triagist in reality is not available to take patient calls as it has to take a call from for example a nursing home, while in the system that might appear to be the case.

Assumption 2 The second assumption is that calls that lasted shorter than 1 minute for the normal arrival line and shorter than 2 minutes for the emergency arrival line were excluded from the data. These calls account for quite a bit of demand but can often not be seen as a real call: many entries had no waiting times and no service times, indicating that a patient was actually never there. It could however be the case that some of the real system demand and waiting time data is lost because of this assumption, impacting the identified demand and service time scenarios.

Assumption 3 The third assumption was made when merging the data of two systems from the out-of-hours department. One is the system in which the calls are logged, the other one is the system in which the characteristics of the call are logged among which is the urgency level. This was done to give each logged call an urgency level for further analysis. However, it was found that not all calls have an indicated urgency level. This could be due to the fact that a caller simply did not get an allocated urgency, for example when no full triage was performed. It could also be due to missing data or human mistakes when entering data and there is no way to distinguish between the two. It is therefore assumed that all calls in the data where the urgency level is missing are 'no triage' calls, unless they have service times higher than 6 minutes: this is the number indicated by the department that it takes at least to perform a triage and it is therefore used as the border between no triage and full triage calls.

Assumption 4 The fourth assumption is directly derived from the third assumption: the triage process is divided into two subprocesses: process 1 where no full triage takes place, and process 2 were full triage takes place. Of the data, all calls below 6 minutes that do not have an urgency level are seen as no triage calls, and all calls that do have an urgency level OR are above 6 minutes in duration are seen as full triage calls. This assumption could lead to the allocation of some patients to another subprocess in the simulation model than they would have been allocated to in reality because they don't meet the restrictions set by these assumptions. This can lead to a lot lower or higher service times for these patients.

Assumption 5 The fifth and last assumption made in the system data analysis is that sometimes groups of data are aggregated when for one or two hours of the day the statistical test indicates differences. This is done because often these hours are during the night where demand is low and no waiting times emerge. Also, the groups of data that this happens for tend to be small, with only a few calls in them and the amount of scenarios would increase rapidly if a new scenario was created for each deviating hour of a small subgroup of data. This assumption can have led to slight simplifications of scenarios that could have been divided on one or two hours on for example weather conditions.

8.1.2 Model

During implementation, verification and validation of the model, several critical assumptions were made, some of them originating in the assumptions made in the system data analysis. Several resulting limitations are identified.

Stochasticity
Limitation One of the most important limitations of the model is that it is a stochastic model, indicating that results from the model run tend to be a little different each time it is run. This called for a large number of runs, which takes a long time specifically when complex simulations are run. When using the model to draw conclusions from, many runs need to be performed to come to a valid confidence interval of the result metrics.

Variable availability of triagists
Assumption A first critical assumption in the model partly comes from the fact that only inbound patient calls are present in the data. Because of this absence, it is assumed in the model that triagists only handle patient calls, where in reality they take more types of calls, they speak to each other and the present general practitioner, have breaks and might have to help somewhere else in the department. To account for this variability, a 30 seconds

delay for each patient before they can be serviced is assumed in the model to account for some of the variability in the availability of triagists. Various time frames were tested for this number, and 30 seconds gave model performance that was most similar to historical data performance. This was compared by looking at the performance during the nights, where there is always enough capacity but still patients sometimes have to wait for some time. When validating the model, it was however still found that when running the model for the standard amount of scheduled triagists at the department, it performs way better than the actual system does. A reduction of the triagists by around 30% in the model had to be implemented compared to the original capacity to reach the historical performance of the system.

Limitation It can therefore be concluded that triagists are, understandably, not available at full capacity for the whole duration of a simulation. This is a limitation when running the model and it should be taken into account when reading system results on triagist capacity. It is important to be cautious about these human aspects of the system which cannot be derived from data. Also, when using this simulation model for other out-of-hours departments, the efficiency of their triagists has to be validated beforehand and cannot be assumed to be the same as for the two out-of-hours departments in this research.

Urgency allocation
Assumption The second critical assumption in the model is that patients get a subprocess and an urgency level allocated *before* they are serviced by a triagist. In reality, the triage process itself is designed to result in an urgency level and the necessary steps to take. However, as it was concluded in the system data analysis that urgency levels have an impact on service times, it is necessary for the simulation model to know beforehand what kind of service time it has to allocate to a patient. If the urgency level is high, these service times tend to be higher than if the urgency level is low.

Experimenting with system changes
Assumption When experimenting with the model, system changes that led to waiting time reductions where found, based on which interesting interventions could be identified later in the research. While experimenting with the optimization of shift lengths and capacity, specific assumptions were made on when capacity changes should take place.

Limitation It is important to realize that the results from running these experiments can be different if the conditions for capacity changes are changed and set to different constraints. It could be the case that changing the optimization settings create much better or worse outcomes.

Quantitative model
Limitation A last limitation (but also strength) of the model is that it is an entirely quantitative model. It cannot be known whether qualitative interventions such as campaigns will really have a demand reducing effect that leads to waiting time reductions: the model can only prove whether those demand reducing effects lead to good system outcomes for the out-hours-department, direct effects from qualitative interventions cannot be measured with this model. The identified interventions that are more focused on the behavior of people within the healthcare system therefore have to be implemented in real life, perhaps first on a small scale, to see whether the desired system changes to reduce waiting times are really reached.

8.2 IMPLICATIONS OF RESULTS

In this Section, the value of the results is discussed in light of the transferability of the results. Next, the implications for practical business use of the results and the implications of the results for literature are discussed.

8.2.1 Transferability of results

The system data analysis and the identification of system changes that lead to a reduction of waiting times and better performance on norms is performed for two out-of-hours departments. It was concluded that many of the demand and service time patterns are similar for the two departments: they follow the same trends and have roughly the same shape: demand peaks often happen in the late morning for weekend days and in the early evening for weekdays. The results of the research have also been presented to a third out-of-hours department, whose data was not used for this research. Many of the factors that were found to impact demand and service times in the two researched data-sets are recognized by this department as well. Additionally, because many of the model conditions that lead to the reduction of waiting times as identified in Section 6.6 also have a beneficial effect when running them with scenarios and data from the second data-set, it can be concluded that if scenarios and distributions based on data of an out-of-hours department are known, the model and its results can be transferred to an out-of-hours department with a similar telephone triage system.

8.2.2 Practical implications

In this section, the business value of the research and its results for out-of-hours departments is presented, in which feedback on the feasibility of the results from three out-of-hours departments in the Netherlands is taken into account, collected after oral presentations of the results to those departments. The possible interventions that lead to statistically significant lower waiting times in out-of-hours departments without having to keep increasing triagist capacity are divided into three categories: quick win interventions, long term interventions and interventions that are practically hard to implement for out-of-hours departments or might not have the intended effect on waiting times. The first two categories refer to the difference in the length of the implementation paths as they require different amounts of work from the out-of-hours departments itself and other healthcare stakeholders.

SHORT TERM INTERVENTIONS: QUICK WIN

The following short term interventions are recommended to start with to reduce waiting times at out-of-hours departments. For each solution, the implementation path is discussed and when applicable specific numbers that should be considered when implementing are mentioned.

- Regular use of the simulation model. To implement this, the departments indicate that the model and its outcomes need to become connected to their dashboard applications or that someone at the department gets instructions for direct use of the model itself. Besides this, the department needs to re-identify demand and service time scenarios yearly in order to keep the scenarios that the model uses up to date.

Using the model like this enables departments to predict system performance with the resources at hand.

- Make it possible to let shifts overlap by letting them start every hour of every two hours, based on how crowded it is, in stead of the standard times every 4-8 hours that the shifts start now. This is especially efficient on the busier days like Fridays and Saturdays. The departments indicate that this is something that is doable in the schedule: the only thing changing for triagists is the fact that the start times of their shifts change, in fact, one of the departments already has many start times for shifts and this yields better response to waiting times and demand. This intervention results in a reduction of peak waiting times by at least 10%, which can increase to percentages around 50% on Saturdays if a shift takes around 4 hours.
- Shift a few patients from peak demand to less busy periods of the day, especially on the most busy day of the week: Saturday. A shift of 1 patient per hour in the weekend from peak demand in the morning to the afternoon already greatly reduces waiting times. The departments are eager to implement this because the technology at hand allows them to easily persuade people to call back later: audio fragments are played to patients waiting in the queue, and when demand is high the department can play an extra audio fragment indicating that it is very busy right now and that if the problem is less urgent it is better to call back at certain hours of the day in which it tends to be less busy. The impact of this solution can be directly measured. The departments should carefully choose the hours of the day in which they want to 'move' patients because moving too many patients to another part of the day only moves the peak and does not flatten it. The model can be used to simulate how waiting times are impacted by different hours and amounts of shifted patients. The departments however indicate that it is key that the patients call back themselves, so that this task does not also lay on the shoulders of the triagists, as that resulted in stressful situations in the past. This intervention results in a waiting time reduction of at least 10-20% on the busy Saturdays when 1 patient per hour is shifted from peak demand hours to lower demand hours, possibly increasing to at least 50% for higher amounts of shifted patients. For the weekdays, a waiting time reduction can be seen when shifting a patient (up to 50% for weekdays, around 10% for Fridays), but its effect is less than for weekend days. This could be due to the fact that demand is not that low in the hours that the patients are shifted to on weekdays, but shifting to a later moment leads to shifting patients to the night, something that is not desirable.
- Reduce service times by retrieving personal information in the queue. In this research, it is found that small reductions in service times lead to big waiting time reductions, but small increases lead to a big increase in waiting times as well. As service times have been increasing over the years, it is key to try to reduce them again or at least avoid an extra increase. This can be achieved when personal information is automatically asked for and retrieved in the queue, in stead of by the triagists. This is a solution that is also quite easily implementable by the departments as the technology exists. Already one of the departments that the results were presented to is starting a pilot by asking social security numbers in the queue, because this number often takes much time of a triagist as many people do not have it ready when calling the department. It can be easily measured if this intervention results in a service time reduction by comparing service time data before and after implementation. An addition to the telephone data that tells your whether or not a call was a video call makes this comparison even more complete. If already a small service time reduction is reached, waiting times are reduced: it was found that a 10% service time reduction can reduce waiting times by at least 50% for most hours of the day in all scenarios.

LONG TERM INTERVENTIONS

Lastly, there are a few interventions possible for out-of-hours departments that might not have direct effects, but have proven to have demand reducing effects over time in other healthcare systems like emergency departments, and demand reduction quantitatively proves to reduce waiting times and increase norm performance in this research: it was found that a 10% demand reduction can reduce waiting times by at least 50% for most hours of the day in all scenarios. These long term solutions require more preparation and cooperation between stakeholders and are recommended to start implementing after the quick win solutions.

- Implement separate lines for home care and nursing homes that cause busyness especially in weekends. The departments indicate that they do not have a clear image of what percentage of the calls are of this kind, but they know that it accounts for quite a portion of the calls when it is already busy. In cooperation with nursing homes and home care, a separate staff member specialized in this type of care and with knowledge of the patient population can then be assigned to handle these types of calls in a certain time frame of the day.
- Increase the accessibility and understanding of regular general practitioners. This is a solution to reduce waiting times that the out-of-hours department agree with strongly, as well as literature on out-of-hours departments. Implementing this solution is however seen as not that easy and straight-forward as there are so many regular general practitioners and they all have different opening hours and rules. To make it an easily accessible solution, the department can start with a small amount of general practitioners (1-5) that they cooperate with. This means coordinating the opening hours of the general practitioner department in such a way that people do not miss that window and are forced to turn to out-of-hours care, but it also means targeting all patients with this information and a guide on how to make appointments and which healthcare department to call when. As it was found that educative measures often do not have the intended effect on health literacy of patients if they are stand-alone, this intervention has a bigger chance of working as it combines an attempt to improve accessibility with education on the healthcare system. The effect of this solution can be measured by comparing the frequency that the patients of these general practitioner departments contact the out-of-hours department before and after these measures are implemented. If it proves to have a reducing effect on patients calling the out-of-hours department, and an increasing effect on people calling their own general practitioner, the solution can be extended to more general practitioners that the out-of-hours department serves.
- Monitor patients who regularly contact out-of-hours departments. Some patients have quite a big record of calling the departments, of which data is available. Out-of-hours departments can cooperate with the regular general practitioners of these patients to monitor and check in with those patients regularly to reduce their use of the out-of-hours departments.
- Implement a small financial (dis)incentive for out-of-hours care. People tend to choose the cheaper option, and are found to choose their regular general practitioner over out-of-hours care if the former is free and the latter is not.
- Implement working from home for triagists in the near future. This could enable short and spontaneous shifts that reduce waiting time peaks on busier days like Friday and in the afternoons of the weekend days.

NOT PRACTICALLY IMPLEMENTABLE

Some of the identified system changes leading to a reduction in waiting times were found to be hard to implement or likely do not have the intended effect on waiting times, after consultation with three out-of-hours departments:

- Implementation of shift lengths of 1 or 2 hours on the busier weekend days or Fridays as long as working from home is not possible in the current infrastructure of out-of-hours departments. Very short shift lengths are hard to sell to triagists according to the departments themselves, as also indicated in Boyle et al. (2012). It creates irregular work schedules and salaries and travel times only for short shifts. This will not improve work environments for triagists contrary to what this research intends to do.
- Use of a website or app for self-triaging of patients before contacting the out-of-hours department, similar to "Moet ik naar de dokter" already implemented in the Netherlands (Verzantvoort et al., 2018): a website where people can check whether or not they should call the out-of-hours department for their problem. The departments indicate that due to the conservatism of this website, they fell like it leads almost all people to the department anyways to avoid taking risks, while in reality many of them do not need to call the department. The intended effect of the website will in that case not be achieved. This notion made by the departments needs to be researched, mentioned in Section 9.5.

8.2.3 Literature implications

In this section, the implications of the results in light of the literature review are discussed. Some of the research results add value to the literature on out-of-hours care or on (healthcare) simulation research, and it is discussed whether findings are contradicting or corresponding to the literature.

The first implication for literature is the identification of the service time and demand scenarios for out-of-hours care. It was found that (temporal) factors such as year, season, weekpart, weekday, patient urgency holidays and days after holidays have an effect on demand and service times in an out-of-hours department, following the results of Hamrock et al. (2013); Marcilio et al. (2013); McCarthy et al. (2008) where it was found that these factors have an impact on the demand for emergency departments. However, the applicability of these factors on service times, besides the urgency levels, was not yet identified in research, as well as the applicability of these factors at all on out-of-hours departments. Next to this implication, it was found that weather factors also influence demand and service times at out-of-hours departments.

The second implication for literature is the notion that a simulation model is built that uses identified scenarios and empirical distributions derived from these scenarios as input. The notion that varying demand and service times are identified is not that unique, many papers within different sectors use complicated models and data analysis to find contributing factors for arrival rates (Hill & Böse, 2017; Xu et al., 2013), but directly using the outcomes of those analyses as input into a simulation model is not seen before for a healthcare system. Often in simulation models of healthcare systems, no extensive prior data analysis is executed or no data is available at all, leading to average values of arrival rates and service times being used that do not take into account variability of demand for healthcare (Abo-Hamad & Arisha, 2013; Keshtkar et al., 2015). The out-of-hours model built in this research is therefore closer to reality than these types of simulation models and can be seen as an interesting addition to healthcare system simulation research as it is completely data-driven.

A third implication for literature is the use of system recommendations made in out-of-hours literature as basis for experiments in the model. In the literature, many recommendations are made to reduce demand for care at these departments, like financial incentives Keizer et al. (2016), using an app Verzantvoort et al. (2018) or using education campaigns to improve health literacy (Jansen et al., 2018; Keizer et al., 2016). From these recommendations from literature, financial incentives and education campaigns are in this research proposed as possible means to reach demand reduction. It remains unknown what the potential effects on demand of these interventions are in an out-of-hours department, but in Anantharaman (2008); Van den Heede and Van de Voorde (2016) it was found that these interventions have demand reducing effects on emergency departments, and as the simulation model quantitatively proved that small demand reductions lead to very significant waiting time reductions this interventions prove very promising. Another recommendation from literature, the use of an app or website such as 'Moet ik naar de dokter?' in Verzantvoort et al. (2018) was considered for the reduction of low urgency calls, which proves promising to reduce waiting times. However, when consulting three out-of-hours departments, they mention that the website is very conservative and that it sends a large part of patients to the out-of-hours department anyway and therefore does not appear to reduce demand. Therefore, the results of the pilot app and website in the paper might be true, but should be complemented with the fact that these apps and websites are conservative because it is medical problems that they deal with, and therefore might not help in demand reduction which was the intended effect. This empirical notion of the out-of-hours departments is therefore an interesting subject for further research, discussed in Section 9.5.

Lastly, it was found in literature reviews that discuss interventions with the goal to reduce waiting times in emergency departments, that the use of a telephone triage system or improvement of the out-of-hours care availability can lead to a reduction of the waiting times and high demand at emergency departments (Van den Heede & Van de Voorde, 2016). It is very interesting to see that the telephone triage at an out-of-hours care department itself has been used as an intervention to *reduce* waiting times at the emergency department, but which in reality deals with waiting times itself and needs its own separate interventions to reduce this. Furthermore, (Van den Heede & Van de Voorde, 2016) also mentions that proof that telephone triage systems can reduce the use of emergency departments is lacking and in Huntley et al. (2013) it is found that increasing the accessibility of out-of-hours services does not reduce the use of emergency departments. This implies that the notion in literature that telephone triage and out-of-hours care can reduce the use of emergency departments is outdated and will only put more demand on the already overcrowded out-of-hours systems.

8.3 CONCLUSION

In this chapter, the main assumptions and limitations of the research were discussed. The found limitations mainly focus on the fact that some of the data in the model might be simplified due to aggregation of some scenarios and due to the model only including incoming patient calls that are longer than 1 or 2 minutes respectively for the normal line and for the emergency line. Also, no other types of calls in the out-of-hours department are included into the model. When drawing conclusions from the research, it is important to be aware of these assumptions and limitations as they might have effects on the implications especially for practical use. In this research, it is already found that waiting times need to be reduced only when taking into account patient calls, and therefore the implications would definitely count if demand was even higher with other types of calls. Next, the quick win, long term and not practically implementable interventions were discussed in the practical section, after consultation with three out-of-hours departments. Lastly, the scientific implications of the research were discussed, mainly focusing on the system data analysis and its identification of data-driven scenarios that are used in combination with a simulation model in healthcare system research, and on quantitative proof that can be added to the recommendations for out-of-hours care that are already made in literature.

9 CONCLUSION

In this final chapter, the conclusions of the research are presented by answering the main research question and the research subquestions. Next, the scientific societal contribution are presented, after which recommendations for further research are made.

9.1 MAIN RESEARCH QUESTION

First, the main research question as defined in Chapter 1 is answered, which was as follows:

How can waiting times at the telephone triage of out-of-hours general practitioner departments in the Netherlands be reduced?

This question addressed the practical and scientific lack of knowledge into how the telephone triage at an out-of-hours department is affected by external and internal factors and how interventions could address the problem of high demand and high waiting times that is experienced in theory and in practice. The main research question is answered by identifying the factors that influence variables within the out-of-hours system, after which ways of reducing waiting times are identified.

In out-of-hours departments in the Netherlands, high demand for care and waiting times that are longer than the designated norms is experienced, leading to undesirable side-effects such as high pressure on staff and inadequate medical response to patients in potential need of urgent medical attention. These waiting times are partly caused by external factors that impact the demand for healthcare and the length of service times within the system, which were identified in this thesis by extensive data analysis of the system. These external factors were found to be temporal factors such as the season, whether it is a holiday or not, the day of the week or the hour of the day, weather factors and the urgency of a patient. The identification of these factors led to multiple scenarios grouped on these factors in which demand and service times are different. Waiting times are however not only caused by demand and service times, but also by other internal system factors such as triagist capacity and scheduling, handling times after calls and the efficiency of triagists at the department. Identifying how the demand and service time scenarios interplay with internal system variables reveals how waiting times can emerge in the out-of-hours department. Identification of this interplay can be achieved with a simulation model, which in its turn can be used to answer the research question by finding out the main ways in which waiting times can be reduced. To come to this answer, the out-of-hours department was modelled in a discrete event simulation model using real-life data from two telephone triage systems. The use of the model for the current situation in the system revealed that the amount of standard triagists that the department uses is in many scenarios not enough to meet the waiting time norms and therefore that other interventions to reduce waiting times within the system are necessary. These interventions were identified by experimenting with system changes in the model and all lead to a reduction of waiting times in the simulated out-of-hours departments. To identify the feasibility of the interventions, they were discussed with three out-of-hours departments and divided in three: quick win interventions, that can be easily and quickly implemented by out-of-hours departments and whose results can be measured directly, long term interventions that focus more on behavior change, require more cooperation between stakeholders and whose results have to be measured over a longer time frame and lastly the interventions that the departments indicate will likely not result in a waiting time reduction or are practically hard to implement.

Quick win interventions

1. A shift of peak demand to less busy periods of the day. This can be implemented by adding extra information to the audio fragments that are played to patients in the queue about peak hours and when to call back. A shift of 1 patient per hour in the peak hours of the busiest day of the week, a Saturday, to the less busy hours reduces waiting times by at least 10-20%. This increases to 50% when shifting 4 patients per hour. For the weekdays, a slightly lower waiting time reduction can be seen when shifting 1 patient: up to 50% for weekdays, around 10% for Fridays. The simulation model needs to be consulted to find the optimum amount of patients to shift, as shifting too many patients only moves the waiting time peak to another part of the day.
2. More overlap in triagist shifts by starting them every hour or every few hours, based on how crowded it is, especially on the busy Fridays and Saturdays. With this flexibility, sudden demand increases that were not foreseen can be handled with staff capacity at hand. Optimizing the schedule hourly on these days reduces peak waiting time of at least 10% which can increase to percentages around 50% on Saturdays if a shift takes around 4 hours.
3. Automatic retrieval of personal patient information in the queue in stead of by a triagist. This reduces the time a call takes, which greatly impacts the waiting times: a 10% call time reduction reduces the waiting times by at least 50% in most hours of the day in all scenarios. The call times have been increasing over the past years, so it is key to try to reduce them again or at least avoid an extra increase.

Long term interventions

4. An increase of accessibility and understanding of the primary healthcare system. By working together with regular general practitioners *("huisartsen")* to coordinate accessible opening hours and by informing patients on making appointments and when to use what service, less of the burden lays on out-of-hours departments. Combining accessibility increase with education on the healthcare system will have a better impact, as stand-alone educative measures often prove less effective for better general use of healthcare systems than intended.
5. Implementation of a small financial (dis)incentive for out-of-hours care. People are found to choose the regular general practitioner over out-of-hours care if the former is free and the latter is not.
6. Implementation of separate telephone lines for home care and nursing homes to reduce the amount and length of these types of calls in weekends.

7. Monitoring of patients who regularly contact out-of-hours departments. Cooperate with the regular general practitioners of these patients and plan frequent check-ins to reduce their use of out-of-hours care.
8. Implementation of working from home for triagists in the near future. This could enable short and spontaneous shifts that reduce waiting time peaks on busier days like Friday and in the afternoons of the weekend days.

The interventions that the departments found to be practically hard to implement or that they found to be unlikely to reduce waiting times are self-triaging by means of a website or app and implementing very short shift lengths of 1 hour within the current infrastructure where working from home is not possible.

If out-of-hours departments actively use the simulation model and keep demand and service time scenarios up to date to track how waiting times are emerging, combined with the implementation of the quick win interventions and by starting with the long term interventions, waiting times can be reduced. This leads to much better work conditions for triagists where long queues and high pressure are not the norm anymore and to better service to patients that might be in need of urgent medical care.

9.2 RESEARCH SUBQUESTIONS

The research subquestions are answered in this section. They go deeper into partial conclusions after answering the main research question. The subquestions followed the modelling cycle: from conceptualization of the model and its processes in the first two subquestions towards model formalization, implementation and model use in the next two subquestions.

1. What concepts and performance indicators are needed to accurately model telephone triage at an out-of-hours department?
In the literature review it was concluded that a discrete event simulation modelling approach is a good choice for accurately modelling a telephone triage system. This subquestion addressed the need for understanding of the system before being able to accurately implement a model that displays the system. The following important system concepts, data and performance indicators were identified to be necessary for a good model implementation. Some of them are (stochastic) internal variables of the system and were quantified in the data-analysis of subquestion 2. Others have fixed values or are assumed to be fixed in the state of the out-of-hours system as it is now.

VARIABLE CONCEPTS

- Inter arrival times between calling patients and emergency patients
- Service times of patients
- Composition of the subprocesses over demand for healthcare
- Composition of urgency levels over demand for healthcare

FIXED CONCEPTS

- Delays before a call
- Handling times after a call
- Triagist schedule and capacity

- Threshold for performance norms for waiting times

These identified concepts were based on the necessary components of a discrete event simulation model as identified in Chapter 2, combined with concepts that are measured in the real life data of out-of-hours departments. The performance indicators that were identified based on these concepts are the waiting times itself, the performance on the norms and the occupancy of triagists.

2. What variables effect the demand and service time behavior of the out-of-hours telephone triage system?
The identified variable concepts in subquestion 1 , the first list displayed just above, were derived from real life data from the telephone triage of out-of-hours departments. Because of the availability of this data, it was possible to analyze the impact of temporal, weather and urgency factors on system performance and to compare the demand and service times between data grouped on those factors. This led to the identification of the factors that have an impact on demand for care and on the service times within the system, from which scenarios could be derived to be used as input into the model to account for the variability of demand and service times. The identified factors that impact healthcare demand and service times at out-of-hours departments were:

- Season
- Weekend or weekday
- Day of the week
- Hour of the day
- Temperature
- Holidays and days after holidays
- Only for service times: subprocess
- Only for service times: urgency levels

Each of these factors prove to have an effect on demand and service-times in one or more of the identified scenarios, for both of the data-sets of out-of-hours departments that were analyzed. It can be concluded that the behavior of the system can be explained by temporal, weather, subprocess and urgency factors and that simulation models of out-of-hours systems should take this variability into account and avoid aggregation over these factors as the resulting system performance might not be corresponding to real life performance.

3. What system changes can reduce waiting times in the out-of-hours telephone triage system compared to the current situation?
Various experiments were run with the implemented simulation model to identify which system changes can lead to a reduction of waiting times and therefore better performance on the waiting time norms, a better work environment for triagists and better service to patients. For all system changes, it was tested whether the waiting times statistically significantly decrease compared to normal system performance to come to valid conclusions.

Firstly, it can be concluded that optimization of triagist capacity under different shift lengths has a reducing effect on waiting times when it is possible to have shift lengths of one hour on the busier days of the week, such as weekends and Fridays, or when it is possible to start triagist shifts hourly in stead of just a few times a day especially on Fridays and Saturdays. The former can reduce afternoon waiting time peaks on Fridays and weekend days by 50%, the latter can reduce total waiting times over the day by at least 10%, which can increase to percentages around 50% on Saturdays if a shift takes around 4 hours.

Secondly, it was found that a reduction of low or no urgency calls can reduce the amount of calls that have to be handled by triagists and therefore reduce waiting times within the system. A low urgency call reduction of 10% or 20% already yields at least a 20% reduction of waiting times in all scenarios, climbing up to a reduction of at least 30% when the low urgency call reduction is also 30%

Thirdly, it was found that small increases in demand or service times have a disruptive effect on the system that the out-of-hours departments would not be able to handle, as waiting times increase to over a few hours when this happens. However, small reductions of demand and service times also yield beneficial results: a decrease of 10% leads to a reduction of at least 50% of waiting times for most hours of the day in all scenarios.

Lastly, it was found that shifting a few patients from peak demand of the busiest day of the week, a Saturday, towards a less busy period of the day already reduces waiting time peaks when only a very few amount of patients per hour are shifted: a shift of 1 patient reduces waiting times by at least 10-20%. This increases to at least 50% for higher amounts of shifted patients. For the weekdays, a slightly lower waiting time reduction can be seen when shifting 1 patient: up to 50% for weekdays, around 10% for Fridays. This lower effect could be due to the fact that demand is not that low in the hours that the patients are shifted to, but shifting to a later moment leads to shifting patients to the night, something that is not desirable.

4. What possible interventions can lead to the system changes (identified in subquestion 3) of the out-of-hours telephone triage system in which waiting times are reduced?
The identified interventions that can potentially lead to the system changes that were identified in subquestion 3 and that lead to a reduction of waiting times can be divided into three categories: quick win interventions that out-of-hours departments can implement themselves and whose effectivity can be measured almost directly, long term interventions often focused more on the behavioral change of people for which cooperation with other stakeholders is necessary and measurements need to take place over a longer period of time, and interventions that are not practically implementable or are not likely to have a reducing effect on waiting times based on a consultation with three out-of-hours departments.

In Table 9.1, all system changes that lead to a waiting time reduction, identified in subquestion 3, are displayed with the interventions that could lead to this system change. The category of intervention is also displayed.

A more elaborate explanation of these interventions can be read in Section 8.2.2.

9.3 SCIENTIFIC CONTRIBUTION

The performed research contributes to science in various ways, the first of which is the most straightforward one: no simulation research has been performed yet in the out-of-hours care field and this research filled that gap by creating a discrete event simulation

Table 9.1: System changes and interventions that can make them happen

System change	Intervention	Category
Reduction of shift lengths	Flexible schedule policy	Not practically implementable
Reduction of shift lengths	Possibility for triagists to work from home	Long term intervention
Optimized start times of shifts	Flexible schedule policy	Quick win
Optimized start times of shifts	Possibility for triagists to work from home	Long term intervention
Shifting of demand	Extra audio fragment in peak demand	Quick win
Reduction of (low urgency) demand	Financial (dis)incentive	Long term intervention
	Increased accessibility and knowledge of primary healthcare system	Long term intervention
	Monitoring of frequent users of out-of-hours care	Long term intervention
	Separate telephone lines for nursing homes and home care	Long term intervention
	Use of website or app for self-triaging	Not practically implementable (too conservative)
Service time reduction	Personal patient information retrieval in queue	Quick win

model that can be used to find out how waiting times can be reduced in these healthcare systems. Different sub-parts of the research can be seen as contributions to literature, some in the direction of simulation research and some in the direction of healthcare and out-of-hours care research.

The first scientific contribution to out-of-hours care research is that the research concludes that temporal factors such as seasonality, part of the week, day of the week and hour of the day have an impact on demand for healthcare at out-of-hours departments and on the service-times within the departments, based on data analysis and on the interpretation of this analysis by people working in the field. In literature, the effect of these factors was found for emergency departments and other types of healthcare systems, but was not yet found for out-of-hours departments. Next to temporal and urgency factors, it was also found that the temperature of a day impacts the demand and service times within an out-of-hours department. Where in out-of-hours literature the demand for care at these departments is only researched in terms of demographic indicators that explain the use of the departments by certain groups of people, this research adds these external weather, urgency and temporal factors to the demographic indicators, leading to a more broad scientific insight into the use and inner processes of telephone triage at out-of-hours departments.

A next scientific contribution is the data-driven variability of demand and service times that was identified and then used as input and internal variables of a discrete event simulation model. This combination is new in healthcare system simulation literature, where this variability has not been taken into account before, as often no historical data on system performance is available, or simply averages of arrival and service times are taken, reducing the practical use of the model for daily predictions, for capacity optimization or for other system interventions such as demand shifting (Abo-Hamad & Arisha, 2013; Keshtkar et al., 2015).

A last scientific contribution is the use of system recommendations made in out-of-hours literature as basis for experiments in the model. This was done especially for the low urgency demand reduction experiments, for which the idea was derived from Verzantvoort et al. (2018) where the website and app "Moet ik naar de dokter" was implemented for self-triaging of patients before contacting the out-of-hours department to reduce the amount of calls. If 65% of the people using this website or app accept its advise, as mentioned in the paper, very promising results when it comes to the reduction of low or no urgency calls in the system are possible. However, when consulting three out-of-hours departments, they mention that the website is very conservative and that it sends a large part of patients to the out-of-hours department anyway and therefore does not appear to reduce demand. Therefore, the results of the pilot app and website in the paper might be true, but should be complemented with the fact that these apps and websites are conservative because it is medical problems that they deal with, and therefore might not help in demand reduction which was the intended effect. This empirical notion of the out-of-hours departments is therefore an interesting subject for further research, discussed in Section 9.5.

Next, the use of financial incentives and educational campaigns was mentioned in Jansen et al. (2018); Keizer et al. (2016) as a potential means to reduce demand, and has proven to have a reducing effect on demand in other healthcare systems Anantharaman (2008). Combining this with the fact that very small demand reductions in the simulation model lead to great performance improvement on the norms and on waiting times gives a very firm updated recommendation for financial incentives and education that is now also based in quantitative simulation research.

A more extensive literature implication section is found in Section 8.2.3.

9.4 SOCIETAL CONTRIBUTION

The societal and practical contribution of this research is twofold: one the one hand, it is based on the conclusions from the system data analysis where factors that impact demand and service times within the system are identified, and on the other hand it is based in the recommendations for interventions that can be made, categorized in quick win interventions, long term interventions and not practically implementable interventions or interventions that the departments indicate do not have a decreasing effect on waiting times.

The first societal contribution is therefore the system data analysis. This gives the department insights into what factors have to be taken into account when scheduling triagists or when predicting demand and service times. With the results of this analysis, they know what seasons, holidays or specific days of the week tend to have higher waiting times than other days. Even without any system changes, the knowledge that a very busy day is coming beforehand can give them time to respond and make a schedule accordingly.

The second societal contribution are the recommendations for interventions, as listed in the answer of the main research question and in Section 8.2.2. The implementation of the quick win interventions can already reduce waiting times on a short term, and combining them with some of the long term interventions will lead to much better work conditions for triagists where long queues and high pressure are not the norm anymore and to better service to patients that might be in need of urgent medical care.

These contributions are of a practical nature for out-of-hours departments, but when the interventions are implemented they can lead to an improvement of the functioning of the primary healthcare system as a whole, by spreading out patients accurately over the healthcare facilities that they should go to with their problem. This is not only beneficial for the departments and its staff, but also leads to better service to patients and better knowledge of patients on how to manoeuvre within the healthcare system.

A more extensive elaboration on the practical implications of this research can be read in Section 8.2.2.

9.5 RECOMMENDATIONS FOR FUTURE RESEARCH

The simulation model in this research covers the telephone triage system within an out-of-hours department. The simulation model provided insights in to what system changes might be desirable to reduce waiting times at these types of departments. However, the model does not cover calls between healthcare departments that the triagists also have to take and it does not cover the part of the out-of-hours department that comes after the call with a triagist is finished. A main recommendation for further research would be to extend the model with other types of calls than patient calls and with the part of the system that comes after the telephone triage. The simulation model would then be extended with the appointment system for people who have to be seen by a doctor at the department, who have to be visited at their homes by a doctor or for whom an ambulance has to be dispatched and also with waiting times at the department and doctor and room capacity at the department. This would give a more complete image of what the effects are of system changes for waiting times all over the system.

Another recommendation that extends the current research is to analyze whether the type of medical problem of a patient has an influence on demand and service times is a recommendation for further research. Now, the urgency level of a patient is taken into account, but it might be interesting to see what types of problems occur on different days of the week and times of the day and how they effect demand and service times. This can be performed if a data-set that includes the specific medical category that a problem falls into per call is made available. This can then be connected to the data analysis on demand and service times in this research.

Next to this main recommendation, some other recommendations for further research can be made. One of those recommendations is to perform research on the preferences of the public when it comes to out-of-hours and primary care to identify what drives their choices for using primary healthcare services. A good method to approach this would be a stated-choice-experiment, where people can make a choice between two ways of contacting the primary services where certain factors are altered: for example the time of reaching out to the service, the type of complaint you have, the amount of time you have to wait or a financial incentive. This can lead to identification of why and how people use primary care services, and could lead to the identification of trade-offs between for example waiting times and money or the severity of the problems and waiting times or money to be able to adequately implement financial incentives and education and healthcare campaigns that were recommended in this research.

A next recommendation is based on the notion from the three consulted out-of-hours departments that the "Should I see a doctor?" or "Moet ik naar de dokter?" website and app is very conservative and that it sends a large part of patients to the out-of-hours department anyway and therefore does not appear to reduce demand. The real impact of the website and app has to be researched in such a way that it is clear what percentage of people that call the department after having used the website or app had an urgency level that was actually low and did not need to call the department. This can be done by asking patients on the phone if they first used the website or app before calling. This can later be connected to their final urgency level, to see the differences in urgency levels between the group of people that didn't use the website or app and the people who did. If the urgency levels of people who used the website or app are significantly lower, the notion that a large group of patients with low urgency levels is still directed to the out-of-hours department by the website or app is true. This can serve as a reason to either stop the website or app, or to make it more conservative if that is medically possible.

A next recommendation involves research that explores the use of the data-driven simulation model for other types of healthcare systems and its corresponding data. It is interesting to find out if it can be slightly altered to also be used in for example hospitals or emergency departments, which are also often modelled using a discrete event simulation approach, and whether experiments can be run with system changes that lead to good performance within other types of healthcare systems.

REFERENCES

Abo-Hamad, W., & Arisha, A. (2013). Simulation-based framework to improve patient experience in an emergency department. *European Journal of Operational Research*, *224*(1), 154–166. doi: 10.1016/j.ejor.2012.07.028

Akkerman, R., & Knip, M. (2004). Reallocation of beds to reduce waiting time for cardiac surgery. *Health Care Management Science*, *7*(2), 119–126. doi: 10.1023/B:HCMS.0000020651.02085.49

Anantharaman, V. (2008). Impact of health care system interventions on emergency department utilization and overcrowding in Singapore. *International Journal of Emergency Medicine*, *1*(1), 11–20. doi: 10.1007/s12245-008-0004-8

Bloemhoff, A., Schoon, Y., Smulders, K., Akkermans, R., Vloet, L. C., Van Den Berg, K., & Berben, S. A. (2020). Older persons are frailer after an emergency care visit to the out-of-hours general practitioner cooperative in the Netherlands: A cross-sectional descriptive TOPICS-MDS study. *BMC Family Practice*, *21*(1). doi: 10.1186/s12875-020-01220-y

Boyle, A., Beniuk, K., Higginson, I., & Atkinson, P. (2012). Emergency Department Crowding: Time for Interventions and Policy Evaluations. *Emergency Medicine International*, *2012*, 1–8. doi: 10.1155/2012/838610

Brasseur, E., Servotte, J. C., Donneau, A. F., Stipulante, S., D'Orio, V., & Ghuysen, A. (2019). Triage for out-of-hours primary care calls: a reliability study of a new French-language algorithm, the SALOMON rule. *Scandinavian Journal of Primary Health Care*, *37*(2), 227–232. doi: 10.1080/02813432.2019.1608057

Caro, J. J., & Möller, J. (2016). Advantages and disadvantages of discrete-event simulation for health economic analyses. *Expert Review of Pharmacoeconomics and Outcomes Research*, *16*(3), 327–329. doi: 10.1586/14737167.2016.1165608

Caro, J. J., Möller, J., Getsios, D., Coudeville, L., El-Hadi, W., Chevat, C., ... Caro, I. (2007). Invasive meningococcal disease epidemiology and control measures: A framework for evaluation. *BMC Public Health*, *7*(1), 130. doi: 10.1186/1471-2458-7-130

Caro, J. J., Ward, A., Deniz, H. B., O'Brien, J. A., & Ehreth, J. L. (2007). Cost-benefit analysis of preventing sudden cardiac deaths with an implantable cardioverter defibrillator versus amiodarone. *Value in Health*, *10*(1), 13–22. doi: 10.1111/j.1524-4733.2006.00140.x

Chen, G., Govindan, K., Yang, Z. Z., Choi, T. M., & Jiang, L. (2013). Terminal appointment system design by non-stationary M(t)/E k/c(t) queueing model and genetic algorithm. *International Journal of Production Economics*, *146*(2), 694–703. doi: 10.1016/j.ijpe.2013.09.001

Choon, O. H., Dali, Z., Beng, P. T., & Magdalene, C. P. Y. (2016). Uncovering Effective Process Improvement Strategies in an Emergency Department Using Discrete Event Simulation. *Operational Research for Emergency Planning in Healthcare: Volume 1*, 139–163. doi: 10.1057/9781137535696_7

Cook, D. A., & Skinner, J. M. (2005). How to perform credible verification, validation, and accreditation for modeling and simulation. *CrossTalk. The Journal of Defense Software Engineering*. doi: 10.1.1.640.7396

Demir, E., Gunal, M. M., & Southern, D. (2017). Demand and capacity modelling for acute services using discrete event simulation. *Health Systems*, *6*(1), 33–40. doi: 10.1057/hs.2016.1

Dhamodharan, A., & Proano, R. A. (2012). Determining the optimal vaccine vial size in developing countries: A Monte Carlo simulation approach. *Health Care Management Science*, *15*(3), 188–196. doi: 10.1007/s10729-012-9200-4

Dransfield, B. (2021). *Kruskal-Wallis ANOVA- Principles.* Retrieved from https://influentialpoints.com/Training/kruskal-wallis_anova-principles-properties-assumptions.htm

EFSA PPR Panel. (2014). Scientific Opinion on good modelling practice in the context of mechanistic effect models for risk assessment of plant protection products. *EFSA Journal*, *12*(3). doi: 10.2903/j.efsa.2014.3589

e Oliveira, B. R., de Vasconcelos, J. A., Almeida, J. F., & Pinto, L. R. (2020). A Simulation-Optimisation approach for hospital beds allocation. *International Journal of Medical Informatics*, *141*, 104174. doi: 10.1016/j.ijmedinf.2020.104174

Erkelens, D. C., Rutten, F. H., Wouters, L. T., Dolmans, L. S., de Groot, E., Damoiseaux, R. A., & Zwart, D. L. (2020). Accuracy of telephone triage in patients suspected of transient ischaemic attack or stroke: a cross-sectional study. *BMC Family Practice*, *21*(1). doi: 10.1186/s12875-020-01334-3

Erkelens, D. C., van Charldorp, T. C., Vinck, V. V., Wouters, L. T., Damoiseaux, R. A., Rutten, F. H., ... de Groot, E. (2021). Interactional implications of either/or-questions during telephone triage of callers with chest discomfort in out-of-hours primary care: A conversation analysis. *Patient Education and Counseling*, *104*(2), 308–314. doi: 10.1016/j.pec.2020.07.011

Erkelens, D. C., Zwart, D. L., Van Der Meer, G. H., Wouters, L. T., De Groot, E., Damoiseaux, R. A., ... Rutten, F. H. (2020). Is the time of calling helpful for differentiating transient ischaemic attack and stroke from mimics in primary care out-of-hours services? A cross-sectional study. *BMJ Open*, *10*(12). doi: 10.1136/bmjopen-2020-041408

Exalto, L. G., van Doorn, S., Erkelens, D. C. A., Smit, K., Rutten, F. H., Kappelle, L. J., & Zwart, D. L. M. (2021). Call Characteristics of Patients Suspected of Transient Ischemic Attack (TIA) or Stroke During Out-of-Hours Service: A Comparison Between Men and Women. *Frontiers in Neurology*, *12*. doi: 10.3389/fneur.2021.669090

Ghanes, K., Jouini, O., Diakogiannis, A., Wargon, M., Jemai, Z., Hellmann, R., ... Koole, G. (2015). Simulation-based optimization of staffing levels in an emergency department. *SIMULATION*, *91*(10), 942–953. doi: 10.1177/0037549715606808

Giesen, M. J., Keizer, E., Van De Pol, J., Knoben, J., Wensing, M., & Giesen, P. (2017). The impact of demand management strategies on parents' decision-making for out-of-hours primary care: Findings from a survey in the Netherlands. *BMJ Open*, *7*(5). doi: 10.1136/bmjopen-2016-014605

Giesen, P., Mokkink, H., Hensing, M., van den Bosch, W., & Grol, R. (2008). Rude or aggressive patient behaviour during out-of-hours GP care: Challenges in communication with patients. *Patient Education and Counseling*, *73*(2), 205–208. doi: 10.1016/j.pec.2008.04.009

Gorunescu, F., McClean, S. I., & Millard, P. H. (2002). Using a queueing model to help plan bed allocation in a department of geriatric medicine. *Health Care Management Science*, *5*(4), 307–312. doi: 10.1023/A:1020342509099

Green, L. (2006). Queueing analysis in healthcare. In *International series in operations research and management science* (Vol. 91, pp. 282–307). Springer, Boston, MA. doi: 10.1007/978-0-387-33636-7_10

Gualandi, S., & Toscani, G. (2018). Call center service times are lognormal: A Fokker-Planck description. *Mathematical Models and Methods in Applied Sciences*, *28*(8), 1513–1527. doi: 10.1142/S0218202518500410

Gul, M., & Celik, E. (2020). An exhaustive review and analysis on applications of statistical forecasting in hospital emergency departments. *Health Systems*, *9*(4), 263–284. doi: 10.1080/20476965.2018.1547348

Günal, M. M., & Pidd, M. (2010). Discrete event simulation for performance modelling in health care: A review of the literature. *Journal of Simulation*, *4*(1), 42–51. doi: 10.1057/jos.2009.25

Gupta, D. (2013). Queueing models for healthcare operations. In *International series in operations research and management science* (Vol. 184, pp. 19–44). Springer, New York, NY. doi: 10.1007/978-1-4614-5885-2_2

Hamrock, E., Paige, K., Parks, J., Scheulen, J., & Levin, S. (2013). Discrete event simulation for healthcare organizations: A tool for decision making. *Journal of Healthcare Management*, *58*(2), 110–124. doi: 10.1097/00115514-201303000-00007

Haraldseide, L. M., Sortland, L. S., Hunskaar, S., & Morken, T. (2020). Contact characteristics and factors associated with the degree of urgency among older people in emergency primary health care: A cross-sectional study. *BMC Health Services Research*, *20*(1). doi: 10.1186/s12913-020-05219-0

He, J., Hou, X.-Y., Toloo, S., Patrick, J. R., & Gerald, G. F. (2011). Demand for hospital emergency departments: a conceptual understanding. *World Journal of Emergency Medicine*, *2*(4), 253. doi: 10.5847/wjem.j.1920-8642.2011.04.002

Heutmekers, M., Naaldenberg, J., Verheggen, S. A., Assendelft, W. J., Van Schrojenstein Lantman-De Valk, H. M., Tobi, H., & Leusink, G. L. (2017). Does risk and urgency of requested out-of-hours general practitioners care differ for people with intellectual disabilities in residential settings compared with the general population in the Netherlands? A cross-sectional routine data-based study. *BMJ Open*, *7*(11). doi: 10.1136/bmjopen-2017-019222

Hill, A., & Böse, J. W. (2017). A decision support system for improved resource planning and truck routing at logistic nodes. *Information Technology and Management*, *18*(3), 241–251. doi: 10.1007/s10799-016-0267-3

Höpken, W., Eberle, T., Fuchs, M., & Lexhagen, M. (2021). Improving Tourist Arrival Prediction: A Big Data and Artificial Neural Network Approach. *Journal of Travel Research*, *60*(5), 998–1017. doi: 10.1177/0047287520921244

Hossain, N. U. I., Debusk, H., Hasan, M., Jaradat, R., & Khasawneh, M. T. (2017). Reducing patient waiting time in an outpatient clinic: A discrete event simulation (DES) based approach. In *67th annual conference and expo of the institute of industrial engineers 2017* (pp. 241–246). Retrieved from https://www.researchgate.net/publication/317730692

Huibers, L., Giesen, P., Smits, M., Mokkink, H., Grol, R., & Wensing, M. (2012). Nurse telephone triage in Dutch out-of-hours primary care: The relation between history taking and urgency estimation. *European Journal of Emergency Medicine*, *19*(5), 309–315. doi: 10.1097/MEJ.0b013e32834d3e67

Huibers, L., Keizer, E., Carlsen, A. H., Moth, G., Smits, M., Senn, O., & Christensen, M. B. (2018). Help-seeking behaviour outside office hours in Denmark, the Netherlands and Switzerland: A questionnaire study exploring responses to hypothetical cases. *BMJ Open*, *8*(10). doi: 10.1136/bmjopen-2017-019295

Huibers, L., Moth, G., Andersen, M., Van Grunsven, P., Giesen, P., Christensen, M. B., & Olesen, F. (2014). Consumption in out-of-hours health care: Danes double Dutch? *Scandinavian Journal of Primary Health Care*, *32*(1), 44–50. doi: 10.3109/02813432.2014.898974

Huibers, L., Philips, H., Giesen, P., Remmen, R., Christensen, M. B., & Bondevik, G. T. (2014). EurOOHnet-the European research network for out-of-hours primary health care. *European Journal of General Practice*, *20*(3), 229–232. doi: 10.3109/13814788.2013.846320

Huibers, L. A., Moth, G., Bondevik, G. T., Kersnik, J., Huber, C. A., Christensen, M. B., ... Wensing, M. (2011). Diagnostic scope in out-of-hours primary care services in eight European countries: An observational study. *BMC Family Practice*, *12*. doi: 10.1186/1471-2296-12-30

Huntley, A., Lasserson, D., Whye, L., Morris, R., Checkland, K., England, H., ... Purdy, S. (2013). Which features of primary care affect unscheduled secondary care use? A systematic review. *BMJ Open*. doi: 10.1136/bmjopen-2013-004746

Jahn, B., Theurl, E., Siebert, U., & Pfeiffer, K. P. (2010). Tutorial in medical decision modeling incorporating waiting lines and queues using discrete event simulation. *Value in Health*, *13*(4), 501–506. doi: 10.1111/j.1524-4733.2010.00707.x

Jansen, T., Hek, K., Schellevis, F. G., Kunst, A. E., & Verheij, R. A. (2020). Socioeconomic inequalities in out-of-hours primary care use: An electronic health records linkage study. *European Journal of Public Health*, *30*(6), 1049–1055. doi: 10.1093/eurpub/ckaa116

Jansen, T., Hek, K., Schellevis, F. G., Kunst, A. E., & Verheij, R. A. (2021, 6). Income-related differences in out-of-hours primary care telephone triage using national registration data. *Emergency Medicine Journal*, *38*(6), 460–466. doi: 10.1136/emermed-2020-209649

Jansen, T., Rademakers, J., Waverijn, G., Verheij, R., Osborne, R., & Heijmans, M. (2018). The role of health literacy in explaining the association between educational attainment and the use of out-of-hours primary care services in chronically ill people: A survey study. *BMC Health Services Research*, *18*(1). doi: 10.1186/S12913-018-3197-4

Jansen, T., Zwaanswijk, M., Hek, K., & De Bakker, D. (2015). To what extent does sociodemographic composition of the neighbourhood explain regional differences in demand of primary out-of-hours care: A multilevel study. *BMC Family Practice*, *16*(1). doi: 10.1186/s12875-015-0275-0

Jun, J. B., Jacobson, S. H., & Swisher, J. R. (1999). Application of discrete-event simulation in health care clinics: A survey. *Journal of the Operational Research Society*, *50*(2), 109–123. doi: 10.1057/palgrave.jors.2600669

Keizer, E., Bakker, P., Giesen, P., Wensing, M., Atsma, F., Smits, M., & Van Den Muijsenbergh, M. (2017). Migrants' motives and expectations for contacting out-of-hours primary care: A survey study. *BMC Family Practice*, *18*(1). doi: 10.1186/s12875-017-0664-7

Keizer, E., Maassen, I., Smits, M., Wensing, M., & Giesen, P. (2016). Reducing the use of out-of-hours primary care services: A survey among Dutch general practitioners. *European Journal of General Practice*, *22*(3), 189–195. doi: 10.1080/13814788.2016.1178718

Keizer, E., Senn, O., Christensen, M. B., & Huibers, L. (2021). Use of acute care services by adults with a migrant background: a secondary analysis of a EurOOHnet survey. *BMC Family Practice*, *22*(1). doi: 10.1186/s12875-021-01460-6

Keizer, E., Smits, M., Peters, Y., Huibers, L., Giesen, P., & Wensing, M. (2015). Contacts with out-of-hours primary care for nonurgent problems: Patients' beliefs or deficiencies in healthcare? Knowledge, attitudes, behaviors, education, and communication. *BMC Medical Research Methodology*, *15*(1). doi: 10.1186/s12875-015-0376-9

Keshtkar, L., Salimifard, K., & Faghih, N. (2015). A simulation optimization approach for resource allocation in an emergency department. *QScience Connect*, *2015*(1), 8. doi: 10.5339/connect.2015.8

Kim, Y. B., Lee, H., & Choi, H. G. (2005). *A discrete event simulation study for incoming call centers of a telecommunication service company* (Vol. 3397). Springer Verlag. doi: 10.1007/978-3-540-30583-5_42

Knight, V. A., Williams, J. E., & Reynolds, I. (2012). Modelling patient choice in healthcare systems: Development and application of a discrete event simulation with agent-based decision making. *Journal of Simulation*, *6*(2), 92–102. doi: 10.1057/jos.2011.21

Kwakkel, J. H. (2017). The Exploratory Modeling Workbench: An open source toolkit for exploratory modeling, scenario discovery, and (multi-objective) robust decision making. *Environmental Modelling and Software*, *96*, 239–250. doi: 10.1016/j.envsoft.2017.06.054

Leutgeb, R., Engeser, P., Berger, S., Szecsenyi, J., & Laux, G. (2017). Out of hours care in Germany - High utilization by adult patients with minor ailments? *BMC Family Practice*, *18*(1). doi: 10.1186/s12875-017-0609-1

Levit, L., Smith, A., Benz, E., & Ferrell, B. (2010). Ensuring Quality Cancer Care through the Oncology Workforce. *Journal of Oncology Practice*, *6*(1), 7–11. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2805353/ doi: 10.1200/JOP.091067

Lin, C. C., Wu, C. C., Chen, C. D., & Chen, K. F. (2019). Could we employ the queueing theory to improve efficiency during future mass causality incidents? *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, *27*(1). doi: 10.1186/s13049-019-0620-8

Marcilio, I., Hajat, S., & Gouveia, N. (2013). Forecasting daily emergency department visits using calendar variables and ambient temperature readings. *Academic Emergency Medicine*, *20*(8), 769–777. doi: 10.1111/acem.12182

McCarthy, M. L., Zeger, S. L., Ding, R., Aronsky, D., Hoot, N. R., & Kelen, G. D. (2008). The challenge of predicting demand for emergency department services. *Academic Emergency Medicine*, *15*(4), 337–346. doi: 10.1111/j.1553-2712.2008.00083.x

Mielczarek, B. (2016). Review of modelling approaches for healthcare simulation. *Operations Research and Decisions*, *1*(1), 55–72. doi: 10.5277/ord160104

Mohammed, M. A., Clements, G., Edwards, E., & Lester, H. (2012). Factors which influence the length of an out-of-hours telephone consultation in primary care: A retrospective database study. *BMC Health Services Research*, *12*(1). doi: 10.1186/1472-6963-12-430

Montalto, M., Dunt, D. R., Day, S. E., & Kelaher, M. A. (2010). Testing the safety of after-hours telephone triage: Patient simulations with validated scenarios. *Australasian Emergency Nursing Journal*, *13*(1-2), 7–16. doi: 10.1016/j.aenj.2009.11.003

Mukherjee, A. K. (1991). A Simulation Model for Management of Operations in the Pharmacy of a Hospital. *Simulation*, *56*(2), 91–103. doi: 10.1177/003754979105600205

Nørøxe, K. B., Huibers, L., Moth, G., & Vedsted, P. (2017). Medical appropriateness of adult calls to Danish out-of-hours primary care: A questionnaire-based survey. *BMC Family Practice*, *18*(1). doi: 10.1186/s12875-017-0617-1

PBL Netherlands Environmental Assessment Agency. (2013). Het PBL-normenkader voor modellen. *Planbureau voor de Leefomgeving*. Retrieved from https://www.pbl.nl/publicaties/het-pbl-normenkader-voor-modellen

Pinto, L. R., De Campos, F. C. C., Perpetuo, I. H. O., & Ribeiro, Y. C. N. M. B. (2015). Analisys of hospital bed capacity via queuing theory and simulation. In *Proceedings - winter simulation conference* (Vol. 2015-Janua, pp. 1281–1292). Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/WSC.2014.7019984

Razali, M., N, & Wah. Y.B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, 2(1), 21–33. Retrieved from https://www.researchgate.net/publication/267205556_Power_Comparisons_of_Shapiro-Wilk_Kolmogorov-Smirnov_Lilliefors_and_Anderson-Darling_Tests

Salmon, A., Rachuba, S., Briscoe, S., & Pitt, M. (2018). A structured literature review of simulation modelling applied to Emergency Departments: Current patterns and emerging trends. *Operations Research for Health Care*, *19*, 1–13. doi: 10.1016/j.orhc.2018.01.001

Sargent, R. G. (2011). Verification and validation of simulation models. *Proceedings - Winter Simulation Conference*, *37*(2), 166–183. doi: 10.1109/WSC.2010.5679166

Smith, S., & Carragher, L. (2021). Prioritization and management of calls from older people to GP out-of-hours services. *International Journal for Quality in Health Care*, *33*(1). doi: 10.1093/intqhc/mzab021

Smits, M., Keizer, E., Huibers, L., & Giesen, P. (2014). GPs' experiences with out-of-hours GP cooperatives: A survey study from the Netherlands. *European Journal of General Practice*, *20*(3), 196–201. doi: 10.3109/13814788.2013.839652

Smits, M., Peters, Y., Broers, S., Keizer, E., Wensing, M., & Giesen, P. (2015). Association between general practice characteristics and use of out-of-hours GP cooperatives. *BMC Family Practice*, *16*(1). doi: 10.1186/s12875-015-0266-1

Steeman, L., Uijen, M., Plat, E., Huibers, L., Smits, M., & Giesen, P. (2021). Out-of-hours primary care in 26 European countries: An overview of organizational models. *Family Practice*, *37*(6), 744–750. doi: 10.1093/FAMPRA/CMAA064

Tikkanen, R., Osborn, R., Mossialos, E., Djordjevic, A., & Wharton, G. (2020). *International Profiles of Health Care Systems 2020*. Retrieved from https://www.researchgate.net/publication/347011106

Tinker, A. (2002). The social implications of an ageing population. *Mechanisms of Ageing and Development*, *123*(7), 729–735. doi: 10.1016/S0047-6374(01)00418-3

Tiwari, S. K., Gupta, V. K., & Joshi, T. N. (2016). M/M/S Queueing Theory Model to Solve Waiting Line and to Minimize Estimated Total Cost. *International Journal of Science and Research (IJSR)*, *5*(5), 1901–1904. doi: 10.21275/v5i5.20051603

Van den Heede, K., & Van de Voorde, C. (2016). Interventions to reduce emergency department utilisation: A review of reviews. *Health Policy*, *120*(12), 1337–1349. doi: 10.1016/j.healthpol.2016.10.002

van der Ham, R. (2018). salabim: discrete event simulation and animation in Python. *Journal of Open Source Software*, *3*(27), 767. doi: 10.21105/joss.00767

Van Der Meer, M. G., Appelman, Y., Rutten, K. H., Van Der Graaf, Y., Nathoe, H. M., Doevendans, P. A., . . . Rutten, F. H. (2019). Are there gender disparities in symptom presentation or triage of patients with chest discomfort at primary care out-of-hours services? An observational study. *BMJ Open*, *9*(11). doi: 10.1136/bmjopen-2019-031613

Van Uden, C. J., Ament, A. J., Hobma, S. O., Zwietering, P. J., & Crebolder, H. F. (2005). Patient satisfaction with out-of-hours primary care in the Netherlands. *BMC Health Services Research*, *5*. doi: 10.1186/1472-6963-5-6

Verzantvoort, N. C., Teunis, T., Verheij, T. J., & Van Der Velden, A. W. (2018). *Self-triage for acute primary care via a smartphone application: Practical, safe and efficient?* (Vol. 13) (No. 6). Public Library of Science. doi: 10.1371/journal.pone.0199284

Walker, N. J., Van Woerden, H. C., Kiparoglou, V., & Yang, Y. (2016). Identifying seasonal and temporal trends in the pressures experienced by hospitals related to unscheduled care. *BMC Health Services Research*, *16*(1), 1–10. doi: 10.1186/s12913-016-1555-7

Welch, S. J., Jones, S. S., & Allen, T. (2007). Timeliness and efficiency: Mapping the 24-hour emergency department cycle to improve patient flow. *Joint Commission Journal on Quality and Patient Safety*, *33*(5), 247–255. doi: 10.1016/S1553-7250(07)33029-8

World Health Organization. (2018). *Ageing and health.* Retrieved from https://www.who.int/news-room/fact-sheets/detail/ageing-and-health

Wouters, L. T., Rutten, F. H., Erkelens, D. C., De Groot, E., Damoiseaux, R. A., & Zwart, D. L. (2020). Accuracy of telephone triage in primary care patients with chest discomfort: a cross-sectional study. *Open Heart*, *7*(2), e001376. doi: 10.1136/openhrt-2020-001376

Wouters, L. T., Zwart, D. L., Erkelens, D. C., Huijsmans, M., Hoes, A. W., Damoiseaux, R. A., ... de Groot, E. (2020). Tinkering and overruling the computer decision support system: Working strategies of telephone triage nurses who assess the urgency of callers suspected of having an acute cardiac event. *Journal of Clinical Nursing*, *29*(7-8), 1175–1186. doi: 10.1111/jocn.15168

Xu, M., Wong, T. C., & Chin, K. S. (2013). Modeling daily patient arrivals at Emergency Department and quantifying the relative importance of contributing variables using artificial neural network. *Decision Support Systems*, *54*(3), 1488–1498. doi: 10.1016/j.dss.2012.12.019

Zeltyn, S., Marmor, Y. N., Mandelbaum, A., Carmeli, B., Greenshpan, O., Mesika, Y., ... Basis, F. (2011). Simulation-based models of emergency departments: Operational, tactical, and strategic staffing. *ACM Transactions on Modeling and Computer Simulation*, *21*(4), 1–25. doi: 10.1145/2000494.2000497

Zhang, C., Grandits, T., Härenstam, K. P., Hauge, J. B., & Meijer, S. (2018). A systematic literature review of simulation models for non-technical skill training in healthcare logistics. *Advances in Simulation*, *3*(1). doi: 10.1186/s41077-018-0072-7

Zhu, Z., Hen, B. H., & Teow, K. L. (2012). Estimating ICU bed capacity using discrete event simulation. *International Journal of Health Care Quality Assurance*, *25*(2), 134–144. doi: 10.1108/09526861211198290

Zwaanswijk, M., Nielen, M. M., Hek, K., & Verheij, R. A. (2015). Factors associated with variation in urgency of primary out-of-hours contacts in the Netherlands: A cross-sectional study. *BMJ Open*, *5*(10). doi: 10.1136/BMJOPEN-2015-008421

A ANALYSIS OF SECOND DATASET

In this appendix, a summary of the analysis of a second data-set from an out-of-hours department is reported. This department is located in another region than the first data-set, analyzed in Chapter 5. The analysis was performed to see whether the found demand and service time patterns are similar for the two out-of-hours departments.

A.1 SYSTEM DATA ANALYSIS

First, the different demand scenarios were identified by grouping the data on temporal and weather factors. Afterwards, the same was performed for service time scenarios. The Kruskal-Wallis test was used, similar to the analysis in Chapter 5, to test whether groups can be aggregated or not. First, the figures that were created when analyzing the data are displayed. The end results of the analysis can be seen in the scenarios in Tables A.1 and A.2.

A.1.1 Demand

First, the demand scenarios were identified by performing the Kruskal-Wallis test on the data that was grouped on years, seasons, week or weekday, specific day of the week, holiday and day after holiday and weather. The results of plotting those different groups are visible in Figures A.1 up and till A.7. The identified demand scenarios based on these plots and statistical testing between the groups can be found in Table A.1.

Figure A.1: Comparison of hourly arrivals between years

A.1.2 Service times

Next, the demand scenarios were identified by performing the Kruskal-Wallis test on the data that was grouped on years, seasons, week or weekday, specific day of the week, holiday, day after holiday, weather and urgency level. The results of plotting those different

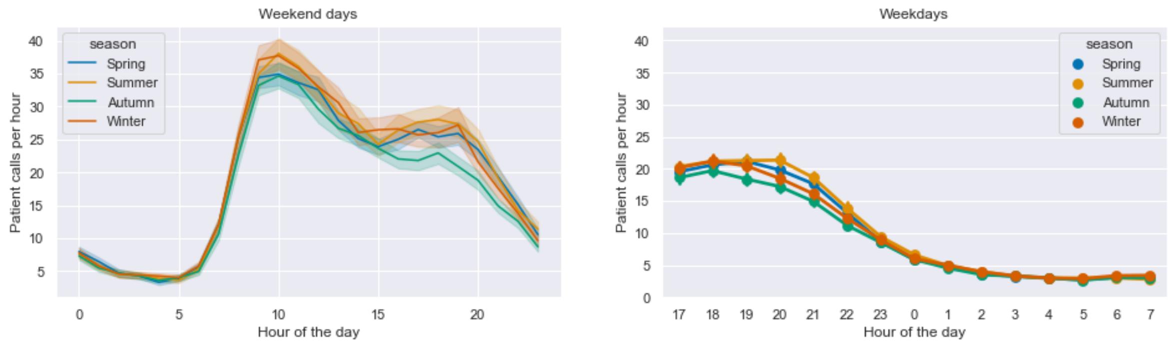

Figure A.2: Comparison of hourly arrivals between seasons

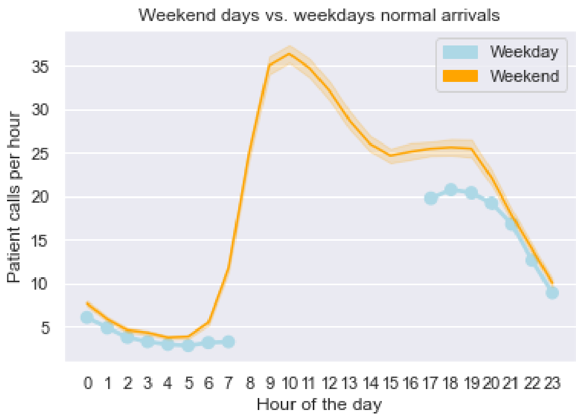

Figure A.3: Comparison of hourly arrivals between weekend and weekdays

groups are visible in Figures A.8 up and till A.12. The identified service time scenarios based on these plots and statistical testing between the groups can be found in Table A.2.

It can be concluded that, next to great similarities in patterns for demand and service times, many of the factors that influence behavior in the first data-set, also influence behavior of demand and service times in the second data-set.

A.2 AGGREGATE HOURS FOR SERVICE TIME

For the first data-set, it was found in 5 that the following hours can be aggregated:

1. Hours 7, 8, 9, 10, 11
2. Hours 17, 18, 19, 20, 21, 22
3. All other hours, so 23 - 6 and 12 - 16

The same analysis for the second data-set gives similar results for a few of the scenarios. For most of the scenarios, all hours can be aggregated. For simplicity, when running the model with the second data-set, the same aggregate hour groups are used as for the first data-set.

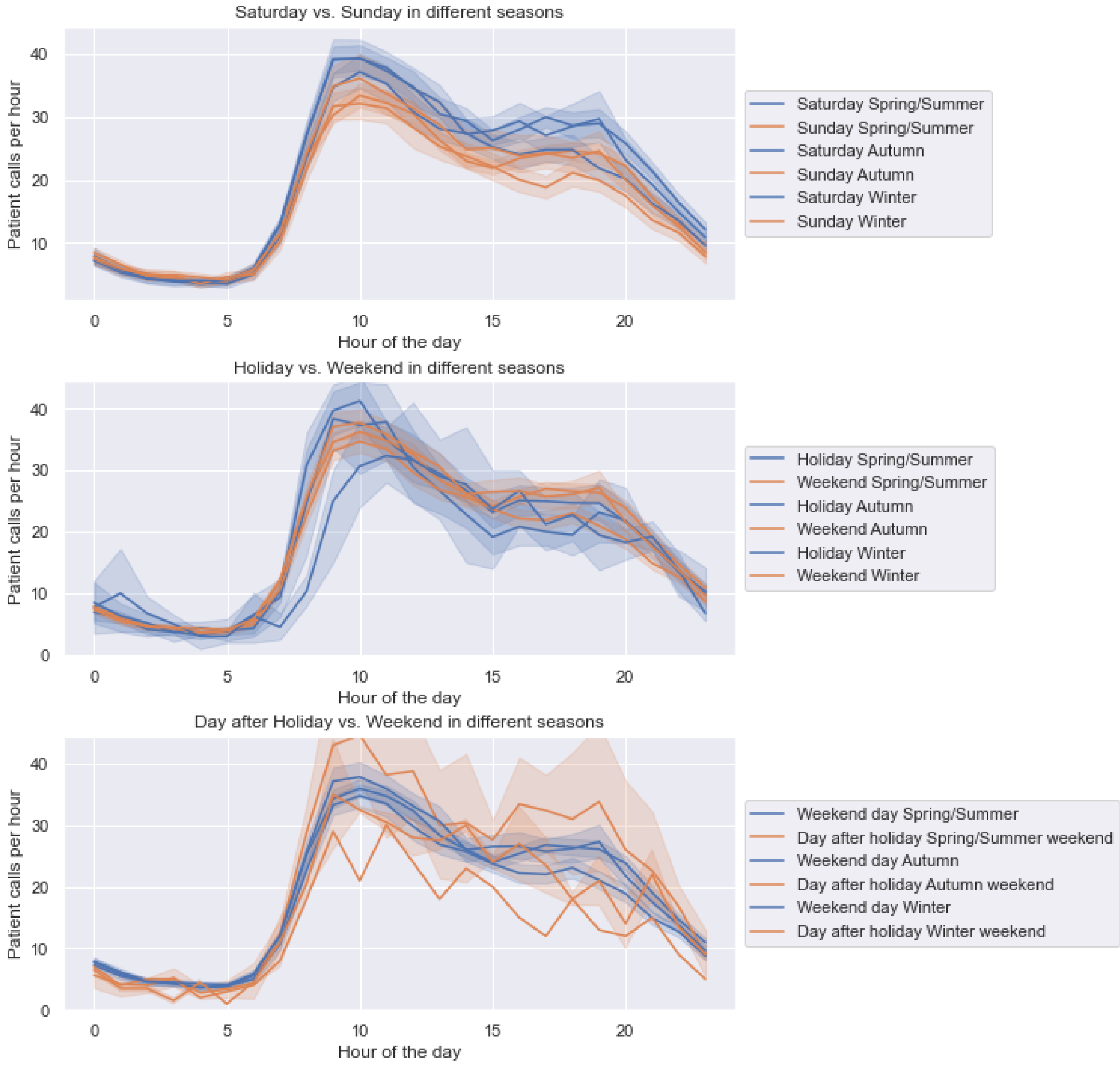

Figure A.4: Comparison of hourly arrivals between weekend days and holidays

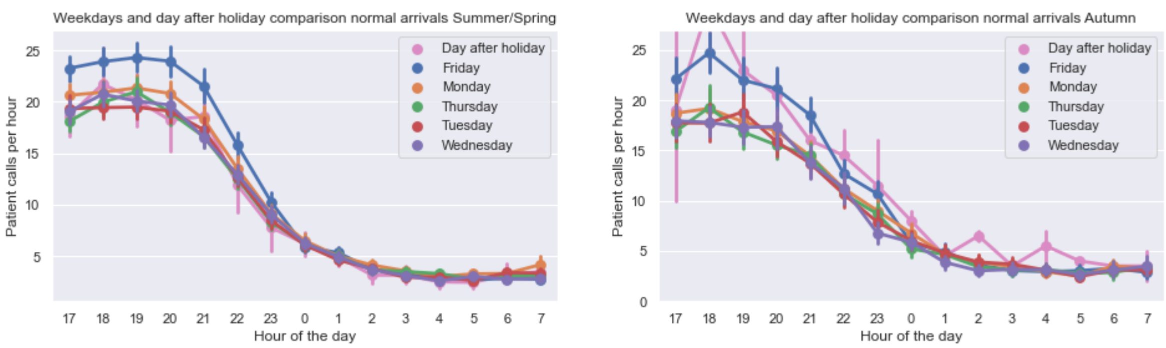

Figure A.5: Comparison of hourly arrivals between weekdays

A.3 IDENTIFIED SCENARIOS

In Tables A.1 and A.2, the identified demand and service time scenarios from performing the system data analysis on the second department data are displayed. The scenarios can all be subdivided into urgency levels 1 and 2 together, and urgency levels 3, 4 and 5 together.

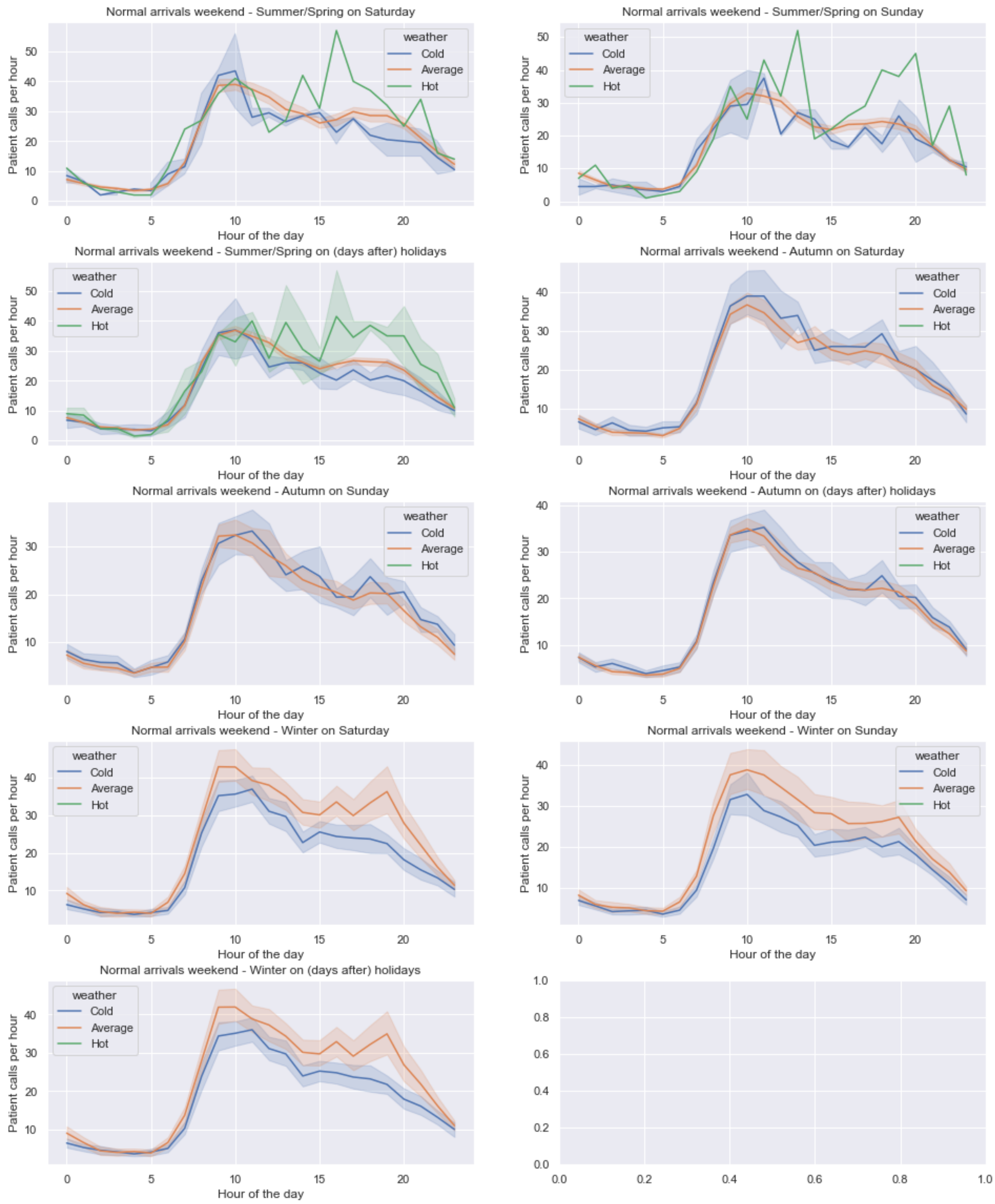

Figure A.6: Comparison of hourly arrivals between weather categories weekend days

A.4 WAITING TIME VALIDATION

The standard triagist schedule is visible in Table A.3. In Figures A.13 and A.14 it can be seen that the model simulates the waiting times correctly (again, standard deviations are displayed as confidence intervals are almost none existent because of the great amount of data) with a slight alteration of the schedule: sometimes when it is busy, some of the capacity of triagists that is actually not on the phone is used or some triagists might be doing something else. This is indicated in the tables. In the mornings, In the night, similar to the first data-set, 2 triagists are needed.

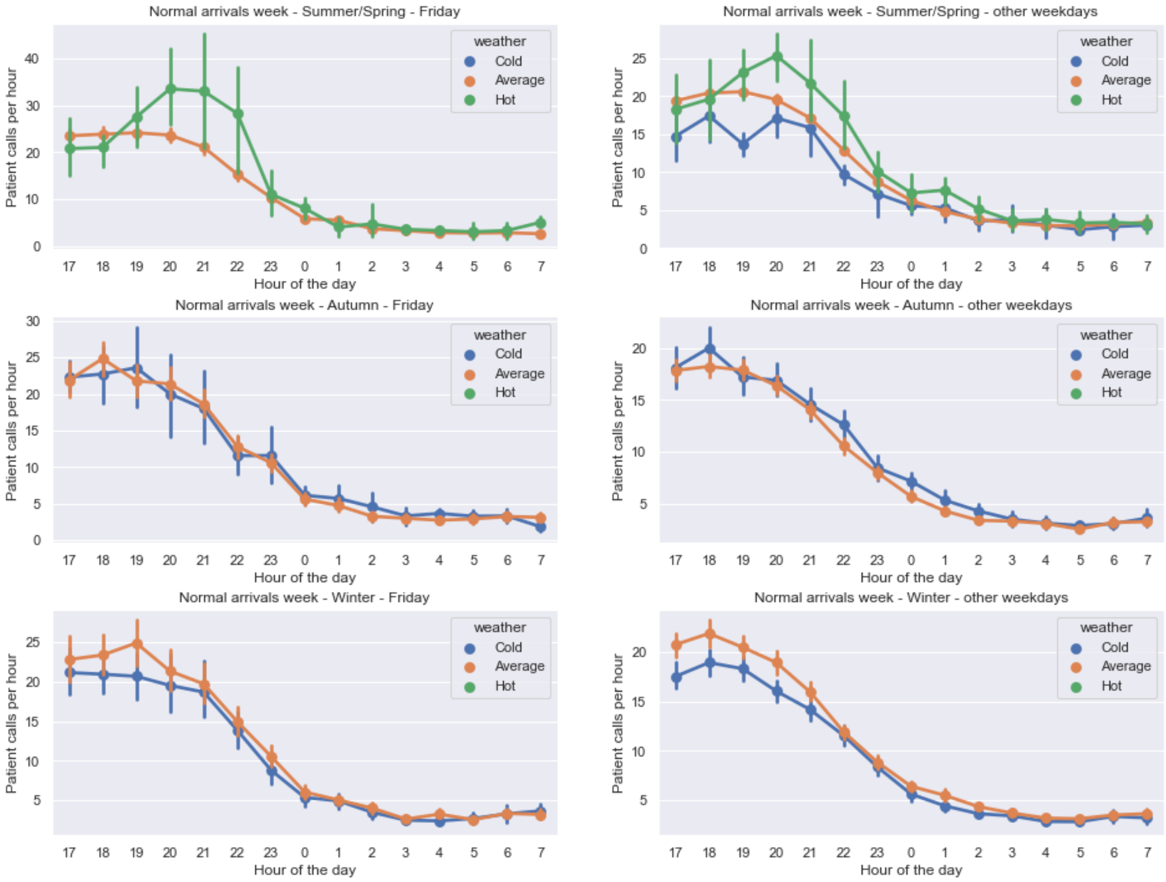

Figure A.7: Comparison of hourly arrivals between weather categories weekdays

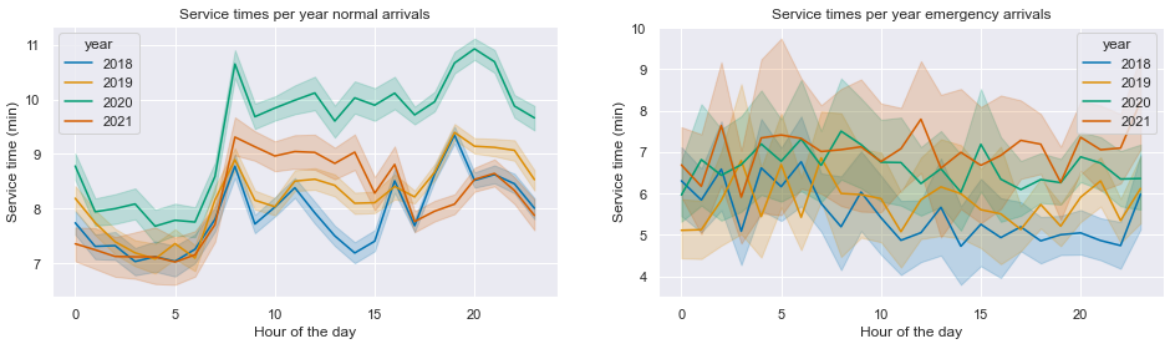

Figure A.8: Comparison of service times between years

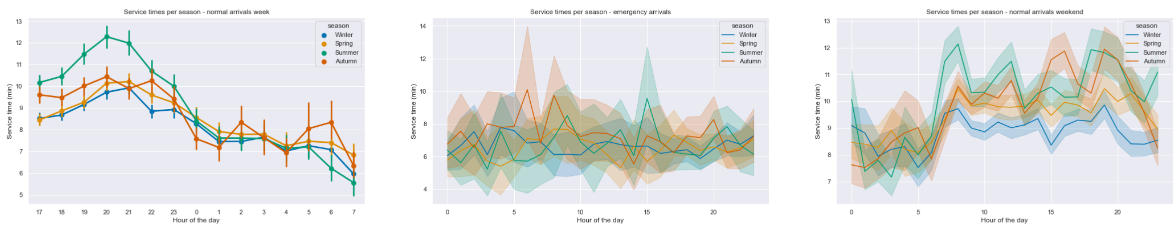

Figure A.9: Comparison of service times between seasons

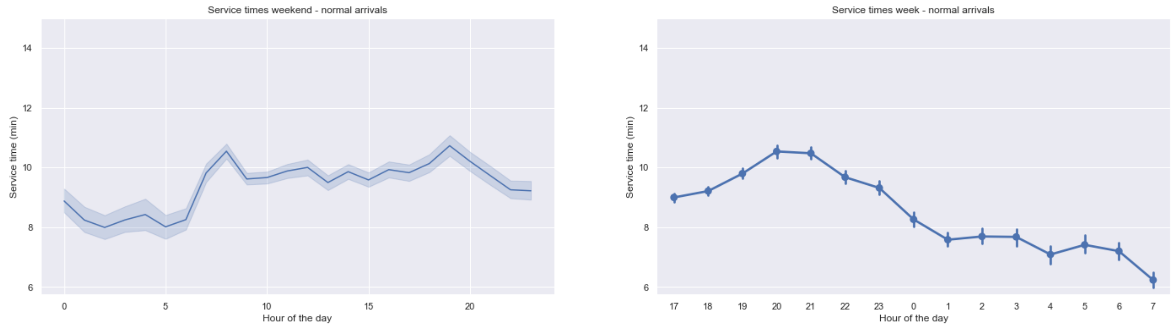

Figure A.10: Comparison of service times between weekend and weekdays

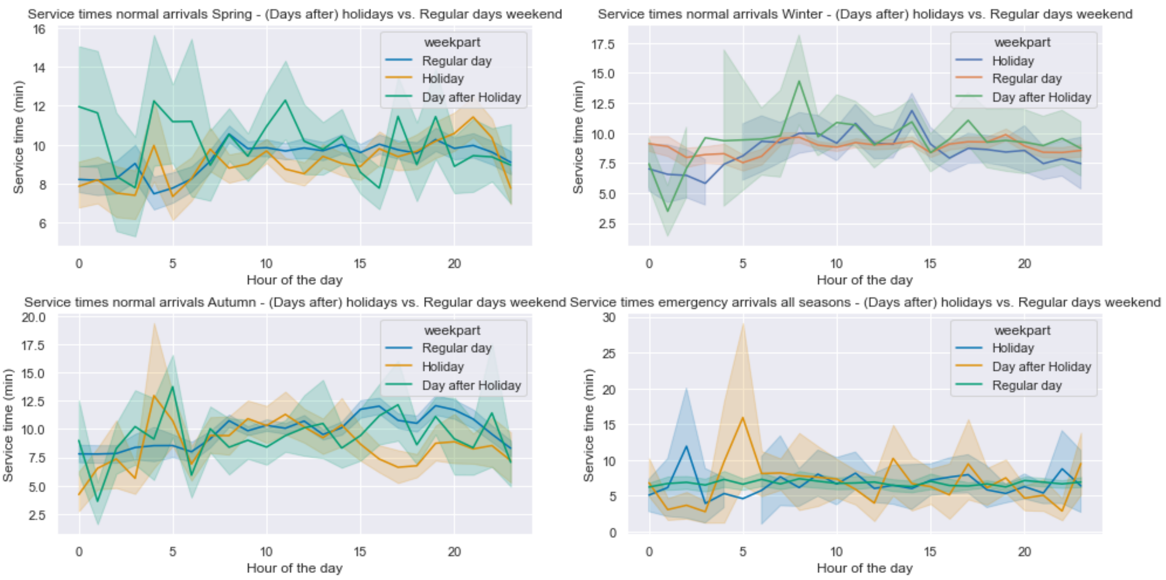

Figure A.11: Comparison of service times between weekend days and holidays

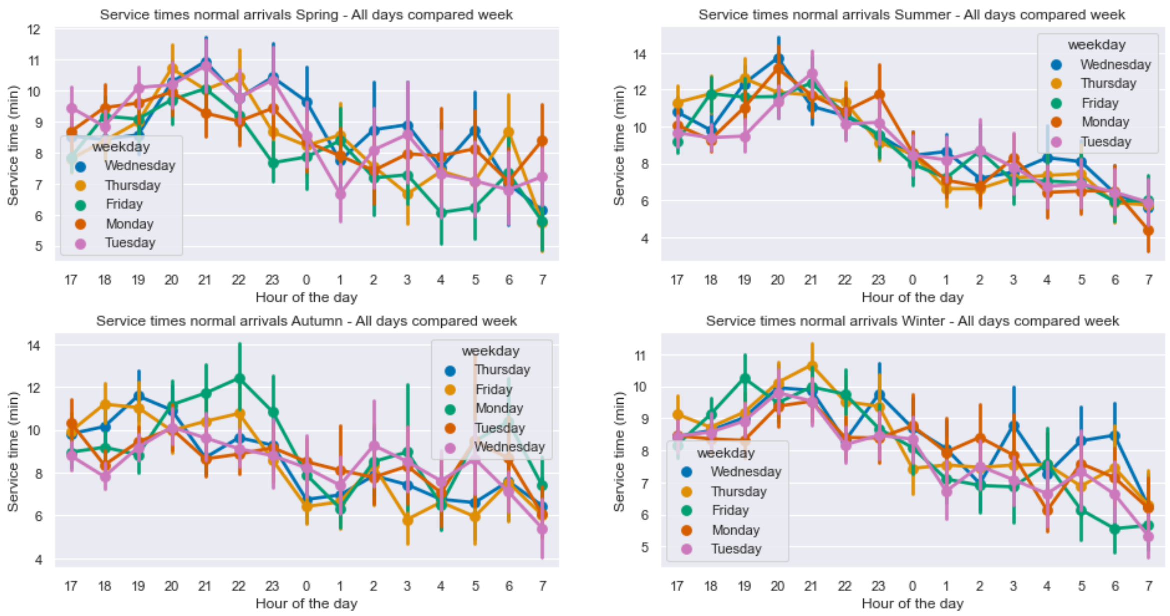

Figure A.12: Comparison of service times between weekdays

Table A.1: Identified demand scenarios second department

Line	**Weekpart**	**Season**	**Day**	**Weather**
Normal	Weekend	Spring/Summer	Saturday	All
Normal	Weekend	Spring/Summer	Sunday	All
Normal	Weekend	Spring/Summer	(Days after) holidays	All
Normal	Weekend	Autumn	Saturday	Cold or Average
Normal	Weekend	Autumn	Sunday	All
Normal	Weekend	Autumn	(Days after) holidays	All
Normal	Weekend	Winter	Saturday	Cold or Average
Normal	Weekend	Winter	Sunday	Cold or Average
Normal	Weekend	Winter	(Days after) holidays	Cold or Average
Normal	Week	Spring/Summer	Friday	Average or Hot
Normal	Week	Spring/Summr	All other weekdays, DAH	Cold, Average or Hot
Normal	Week	Autumn	Friday	All
Normal	Week	Autumn	All other weekdays, DAH	Cold or Average
Normal	Week	Winter	Friday	All
Normal	Week	Winter	All other weekdays, DAH	Cold or Average
Emergency	Weekend	All seasons	All days	All
Emergency	Week	All seasons	All days	All

Table A.2: Identified service time scenarios second department

Line	**Weekpart**	**Season**	**Day**	**Weather**
Normal	Weekend	Spring	All	Average or Hot
Normal	Weekend	Summer	Saturday	Average or Hot
Normal	Weekend	Summers	Sunday	Average or Hot
Normal	Weekend	Autumn	Saturday	Cold or Average
Normal	Weekend	Autumn	Sunday and DAH	All
Normal	Weekend	Autumn	Holiday	All
Normal	Weekend	Winter	All	Cold or Average
Normal	Week	Spring	All	Cold, Average or Hot
Normal	Week	Summer	All	Average or Hot
Normal	Week	Autumn	All	Cold or Average
Normal	Week	Winter	All	Cold or Average
Emergency	All	All	All	All

Table A.3: Standard triagist schedule second department

Day	Shift	Phone capacity
Weekday	17-18: 3	3
	18-21: 6	4
	21-23: 5	3
	23-0: 3	1
Friday	17-18: 5	5
	18-19: 5	3
	19-21: 6	4
	21-23: 5	3
	23-0: 2	0
Saturday	7-8: 2	2
	8-9: 5	3
	9-14: 9	7
	14-15: 7	5 (in model, one triagist less is used here)
	15-16: 7	5
	16-17: 5	3 (in model, two triagists more are used here)
	17-20: 6	4 (in model, one triagist more is used here)
	20-21: 7	5
	21-22: 6	4
	22-23: 5	3
	23-0: 2	0
Sunday	7-8: 2	2
	8-9: 5	3
	9-12: 8	6
	12-14: 9	7
	14-15: 7	5
	15-16: 7	5
	16-17: 6	4
	17-19: 7	5
	19-20: 6	4
	20-21: 7	5
	21-22: 6	4
	22-23: 5	3
	23-0: 4	2
Night	Always 2	2

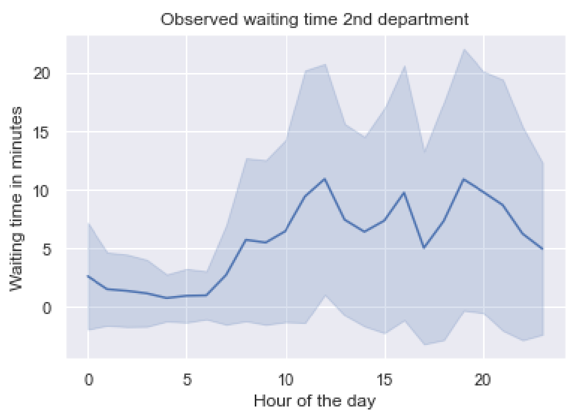

Figure A.13: Waiting time observed on a Saturday second department

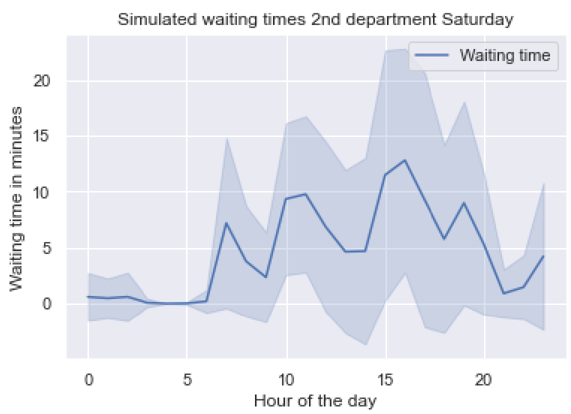

Figure A.14: Waiting time simulated on a Saturday second department

B MODEL CONDITIONS FOR PERFORMANCE IMPROVEMENT - ADDITIONS

In this Appendix, the results of running the different explored model conditions for performance improvement in Chapter 6 in other scenarios are displayed. Each of the different options is tested on a busy week scenario, a Friday, a less busy week scenario: all other days and on a less busy weekend day: a Sunday. The options for model performance improvement are also tested on a second data-set from another out-of-hours department, to establish whether interesting system changes also work for out-of-hours departments in general. For all tested system changes on the scenarios, a statistical test (Kruskal-Wallis and posthoc-Dunn, see Section 5.3) is performed to see whether it reduces waiting times compared to normal system performance. The results of these tests are briefly mentioned for each scenario.

B.1 VARIANT 1

In this section, the results of using shift length and capacity optimization in other scenarios are displayed, from which in general similar conclusions can be drawn: for variant 1, lower shift lengths lead to better model performance on waiting times and norms than longer shift length, for example visible in Figures B.3 and B.4. For variant 2, the model performance is much more consistent throughout shift lengths for most of the scenarios, similar to when its run in the Saturday scenario, visible in for example Figure B.16 where the variant is run for the second department.

B.1.1 Friday scenario

In Figures B.1 and B.2, the results of running the first variant of shift and capacity optimization on a Friday are displayed. When statistically testing whether short shift lengths have a significant reducing impact on waiting times compared to normal model performance, it was found that for all hours of the day, a shift length of 1 hour reduces waiting times compared to normal model performance.

B.1.2 Sunday scenario

In Figures B.3 and B.4, the results of running the first variant of shift and capacity optimization on a Sunday are displayed. When statistically testing whether short shift lengths have a significant reducing impact on waiting times compared to normal model performance, it was found that a shift length of 1 hour reduces waiting times in the afternoon compared to normal model performance.

Figure B.1: Length of stay in queue for different shift lengths - Friday

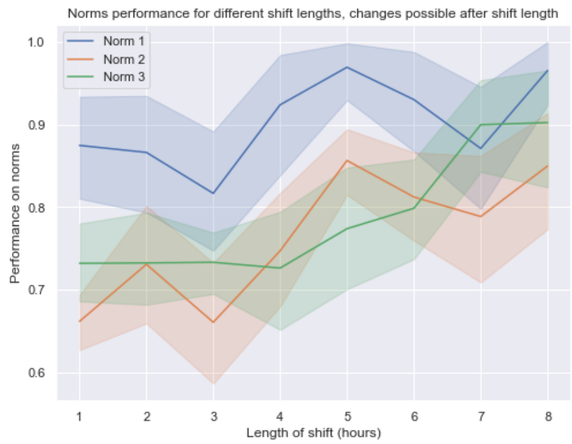

Figure B.2: Norm performance for different shift lengths - Friday

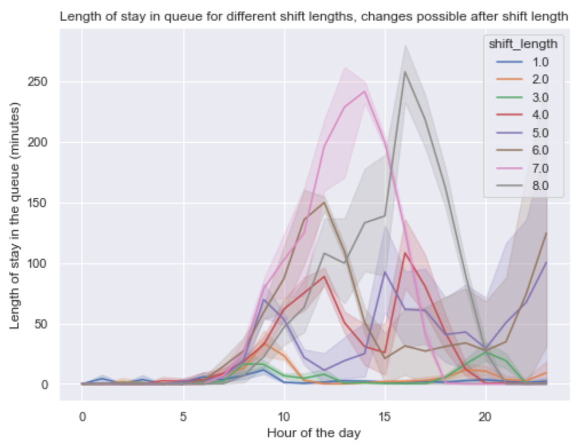

Figure B.3: Length of stay in queue for different shift lengths - Sunday

Figure B.4: Norm performance for different shift lengths - Sunday

B.1.3 Weekday scenario

In Figures B.5 and B.6, the results of running the first variant of shift and capacity optimization on a weekday are displayed. When statistically testing whether short shift lengths have a significant reducing impact on waiting times compared to normal model performance, it was found that for all hours of the day, a shift length of 1 hour performs the same or slightly worse in the first hours of the day. This can be explained by the way the optimization works, and the fact that during weekdays it is already not that busy and optimizing therefore does not have the intended effect.

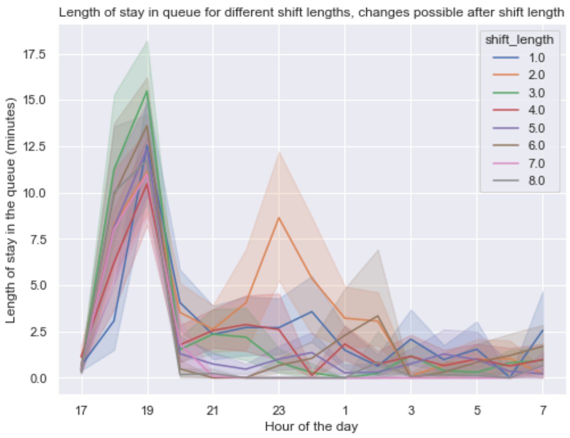

Figure B.5: Length of stay in queue for different shift lengths - Weekday

Figure B.6: Norm performance for different shift lengths - Weekday

B.1.4 Second data-set scenario

In Figures B.7 and B.8, the results of running the first variant of shift and capacity optimization on a Saturday for the second department are displayed. When statistically testing whether short shift lengths have a significant reducing impact on waiting times compared to normal model performance, it was found that a shift length of 1 hour reduces waiting times to zero for almost all hours of the day, except around 10 o'clock, where the peak stays roughly the same.

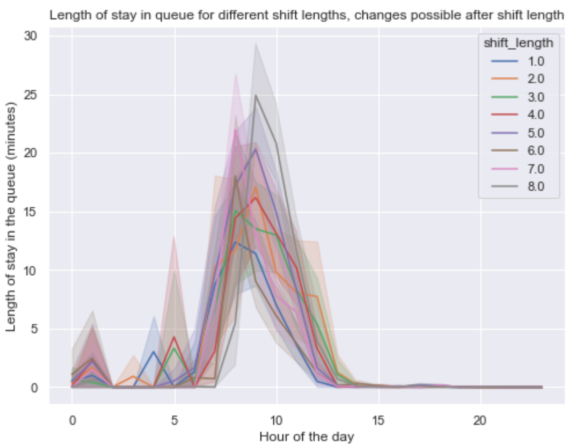

Figure B.7: Length of stay in queue for different shift lengths - second department Saturday

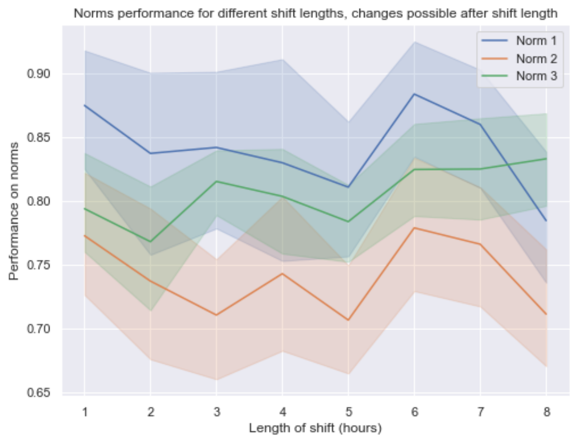

Figure B.8: Norm performance for different shift lengths - second department Saturday

B.2 VARIANT 2

The same has been done for variant 2.

B.2.1 Friday scenario

In Figures B.9 and B.10, the results of running the second variant of shift and capacity optimization on a Friday are displayed. When statistically testing whether the shift lengths have a significant reducing impact on waiting times compared to normal model performance, it was found that all shift lengths perform differently to normal model performance for hours 17 till 7, which are all hours that the department is open. For most shift lengths except 2 and 3 hours, the waiting times of the optimization are better than normal model performance.

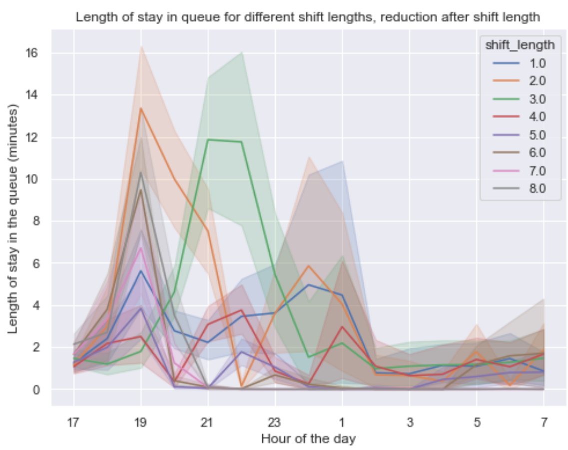

Figure B.9: Length of stay in queue for different shift lengths - Friday

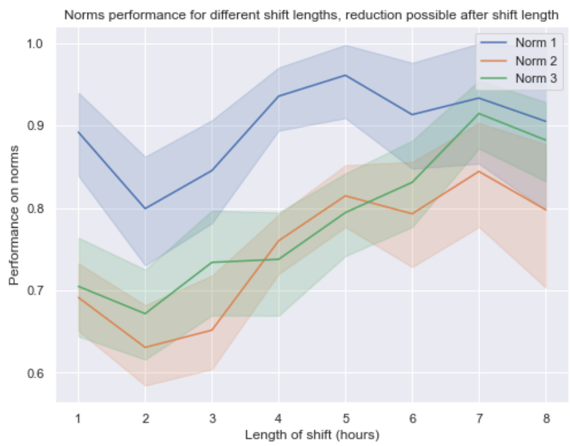

Figure B.10: Norm performance for different shift lengths - Friday

B.2.2 Sunday scenario

In Figures B.11 and B.12, the results of running the second variant of shift and capacity optimization on a Sunday are displayed. When statistically testing whether the shift lengths have a significant reducing impact on waiting times compared to normal model performance, it was found that all shift lengths perform differently to normal model performance for hours 0 up and till 22 with sometimes an exception in the mornings when waiting times are 0. The first peak of the day is often slightly higher than normal model performance, but the second peak of the day is reduced: total waiting time stays approximately the same even though the performance hourly is statistically different. It should however also be noted here that the way the optimization is run greatly impacts performance of the waiting times. Other thresholds for increasing capacity could yield very different results.

B.2.3 Weekday scenario

In Figures B.13 and B.14, the results of running the second variant of shift and capacity optimization on a Weekday are displayed. When statistically testing whether the shift lengths have a significant reducing impact on waiting times compared to normal model performance, it was found that all shift lengths perform differently to normal model performance for almost all hours. As the waiting times during the week are not that high, a

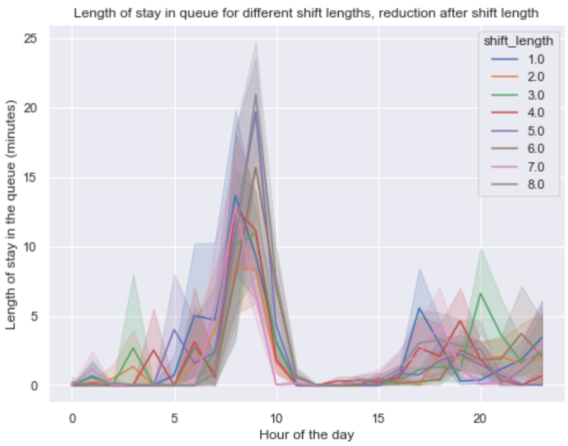

Figure B.11: Length of stay in queue for different shift lengths - Sunday

Figure B.12: Norm performance for different shift lengths - Sunday

reduction compared to normal performance is not achieved for weekdays with this type of optimization.

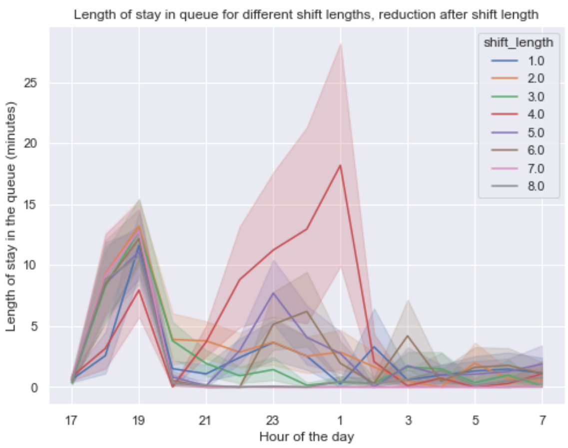

Figure B.13: Length of stay in queue for different shift lengths - Weekday

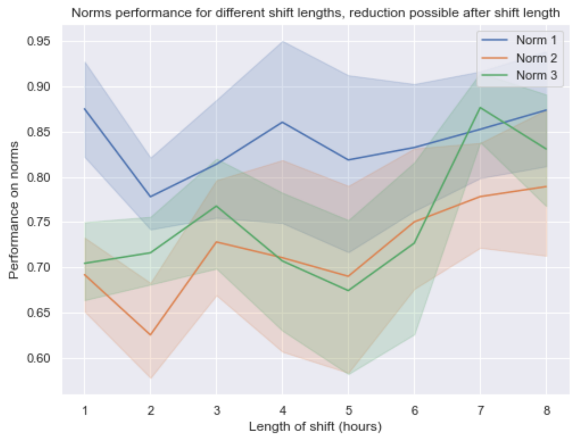

Figure B.14: Norm performance for different shift lengths - Weekday

SECOND DATA-SET SCENARIO

In Figures B.15 and B.16, the results of running the second variant of shift and capacity optimization on a Saturday for the second department are displayed. When statistically testing whether the shift lengths have a significant reducing impact on waiting times compared to normal model performance, it was found that all shift lengths perform differently to normal model performance for hours 8 up and till 22. The one peak of the day is often slightly higher than normal model performance, but all other waiting time throughout the day is reduced.

B.3 DECREASE IN LOW URGENCY CALLS

In this section, the results of decreasing low urgency calls in other scenarios than the busy Saturday scenario in the main text are displayed. In these scenarios, low urgency call reduction has similar effects as in the Saturday scenario displayed in Chapter 6. For example when looking at Figure B.17 it can be seen that low urgency call reductions can

Figure B.15: Length of stay in queue for different shift lengths - second department Saturday

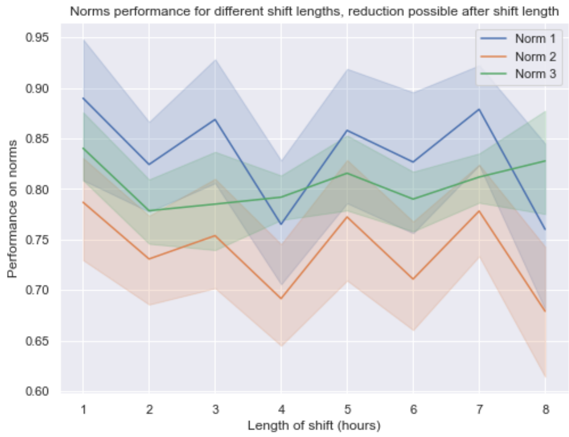

Figure B.16: Norm performance for different shift lengths - second department Saturday

reduce waiting times already starting at a reduction of 10%. Also, when implementing this reduction with the data and scenarios of the second data-set as analyzed in Appendix A, similar outcomes can be seen and similar conclusions can be drawn, visible in Figure B.20. This indicates that a low urgency call reduction works in all scenarios and in other out-of-hours departments with similar systems.

B.3.1 Friday scenario

In Figure B.17, the results of implementing different levels of low urgency call reduction on a Friday are displayed. When statistically testing whether a 10% reduction has a statistically significant reducing impact on waiting times compared to normal model performance, it was found that for 17 up and till 21 excluding 18 where you can see that the peak is almost similar, the reduction performs significantly better on waiting times than the normal waiting times. These are the peak hours of the evening and it can therefore be concluded that a small reduction has a big reducing effect already of around 50% in total.

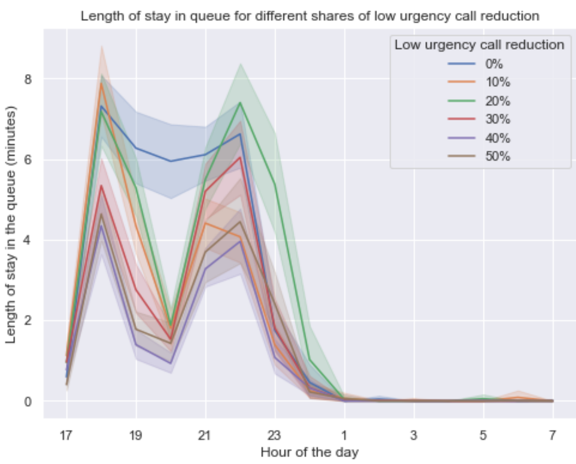

Figure B.17: Waiting time with different levels of low urgency call reduction - Friday

B.3.2 Sunday scenario

In Figure B.18, the results of implementing different levels of low urgency call reduction on a Sunday are displayed. It can be concluded that a 10% reduction of low urgency calls statistically significantly reduces waiting times on a Sunday in the second peak of the day from hour 16 till 22 by at least 20%.

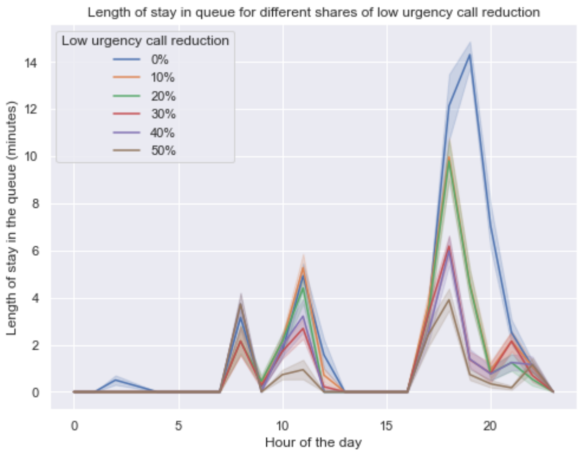

Figure B.18: Waiting time with different levels of low urgency call reduction - Sunday

WEEKDAY SCENARIO

In Figure B.19, the results of implementing different levels of low urgency call reduction on a Weekday are displayed. When statistically testing whether a 20% reduction (the 10% does not have a big impact here, waiting times are already small in this scenario) has a statistically significant reducing impact on waiting times compared to normal model performance, it was found that for 19, 20 and 22 the low urgency demand reduction gives a waiting time reduction compared to normal model performance. It can be concluded that for weekdays where it is less busy, at least a 20% reduction is necessary for waiting time reduction.

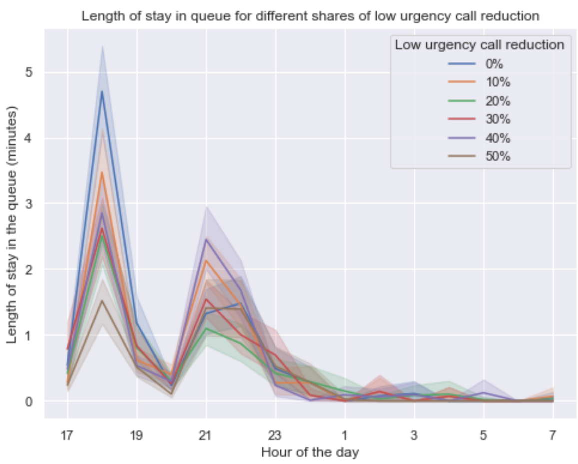

Figure B.19: Waiting time with different levels of low urgency call reduction - Weekday

SECOND DEPARTMENT SCENARIO

In Figure B.20, the results of implementing different levels of low urgency call reduction on a Weekday are displayed. When statistically testing whether any of the reductions lead to a statistically significant waiting time reduction, a 20% reduction is found to be significant for a waiting time reduction of 20% in the afternoons.

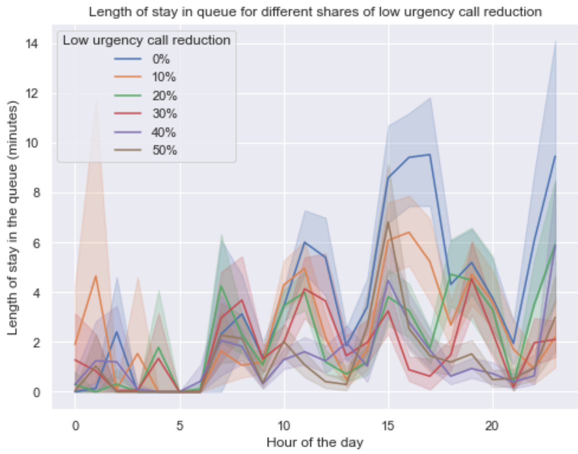

Figure B.20: Waiting time with different levels of low urgency call reduction - second department

B.3.3 Changes in demand

The effect of demand increases and reductions in other scenarios was also tested. It can be concluded that demand increases and reductions yields beneficial outcomes on waiting time and demand for all scenarios. All statistical tests yield at least a statistically significant waiting time reduction of 40%, often 50%, already when a demand decrease of 10% takes place.

FRIDAY SCENARIO

In Figures B.21 and B.22, the results of different values of demand reduction on a Friday are displayed.

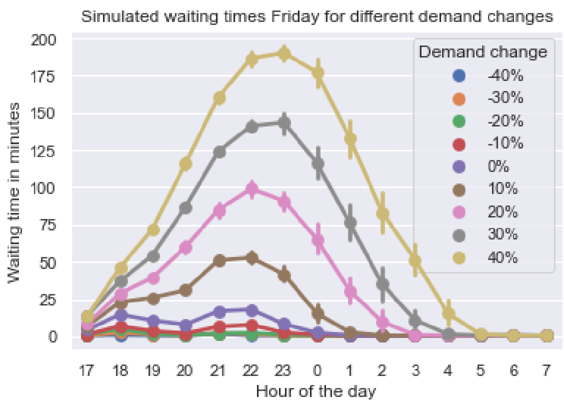

Figure B.21: Waiting time with different levels of demand increases and reductions - Friday

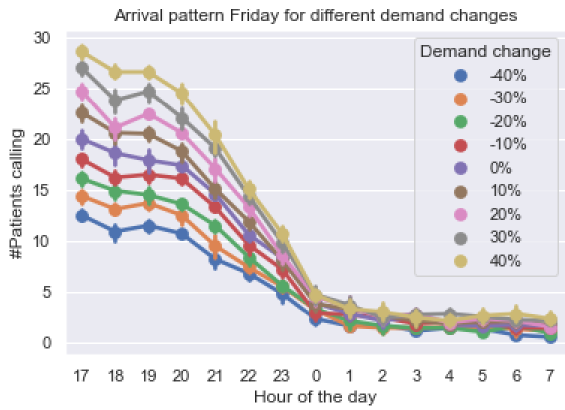

Figure B.22: Demand with different levels of demand increases and reductions - Friday

SUNDAY SCENARIO

In Figures B.23 and B.24, the results of different values of demand reduction on a Sunday are displayed.

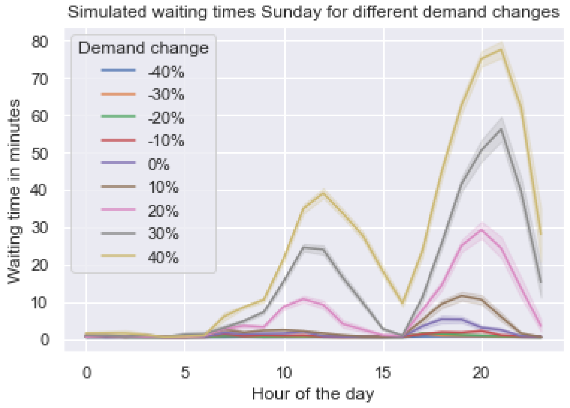

Figure B.23: Waiting time with different levels of demand increases and reductions - Sunday

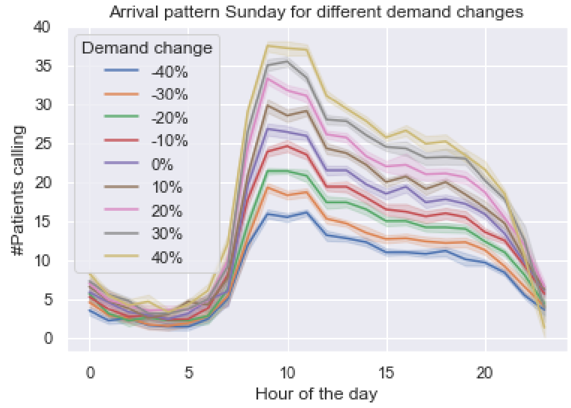

Figure B.24: Demand with different levels of demand increases and reductions - Sunday

WEEKDAY SCENARIO

In Figures B.25 and B.26, the results of different values of demand reduction on a Weekday are displayed.

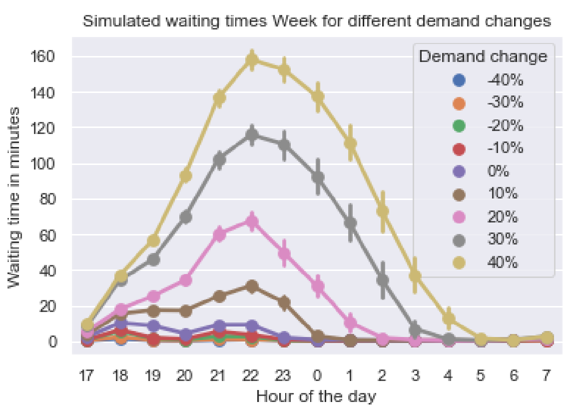

Figure B.25: Waiting time with different levels of demand increases and reductions - Weekday

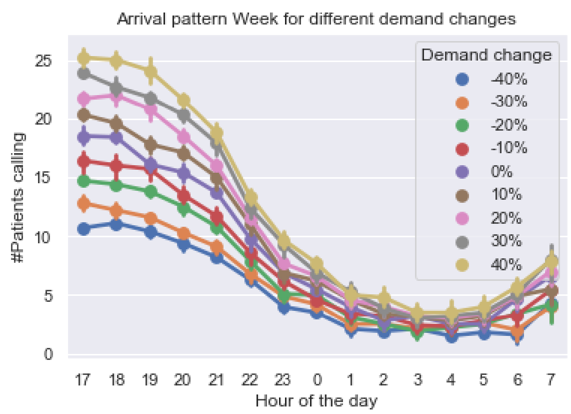

Figure B.26: Demand with different levels of demand increases and reductions - Weekday

SECOND DEPARTMENT SCENARIO

In Figures B.27 and B.28, the results of different values of demand reduction on a Saturday for the second department are displayed.

B.3.4 Demand shifting

When using demand shifting in a weekday scenario, demand from hours 17, 18 and 19PM is shifted to 20, 21 and 22PM: these are hours in which people might accept a call at a later moment and are also the most crowded hours of a workday. It is concluded that also for the less busy scenarios, demand shifting can help to reduce waiting times and workload.

FRIDAY SCENARIO

In Figures B.29 and B.30, the results of different values of demand shifting on a Friday are displayed. When statistically testing whether any of the demand shifts lead to a statistically

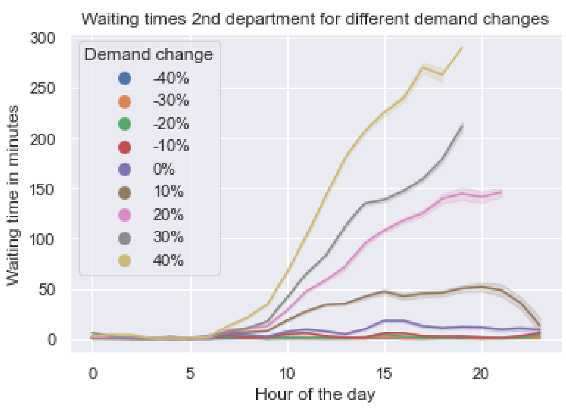

Figure B.27: Waiting time with different levels of demand increases and reductions - second department Saturday

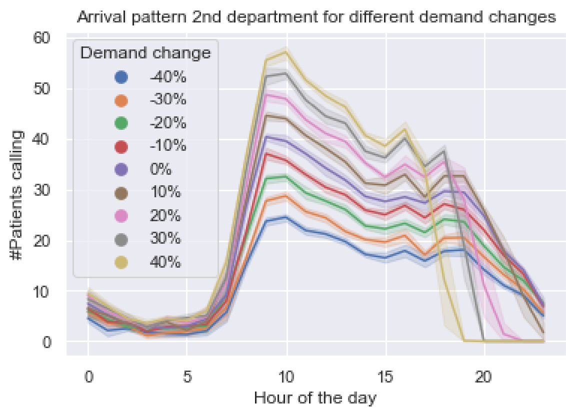

Figure B.28: Demand with different levels of demand increases and reductions - second department Saturday

significant waiting time reduction, a shift of 1 patient on Fridays significantly helps to (very) slightly reduce waiting times especially around hour 18 and 19, but as on weekdays the shift can only be a few hours later as otherwise the calls are shifted to the night, the impact of shifting is not that big.

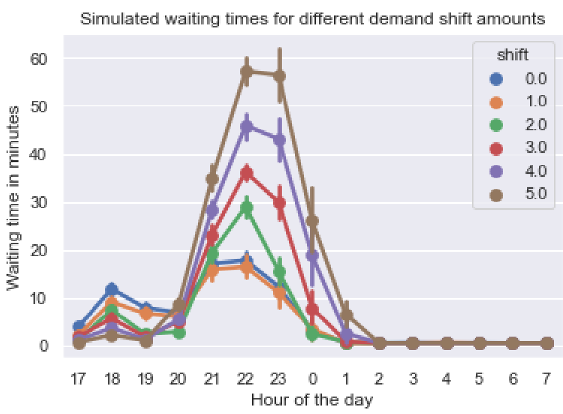

Figure B.29: Length of stay in queue with different amount of demand shifting Friday

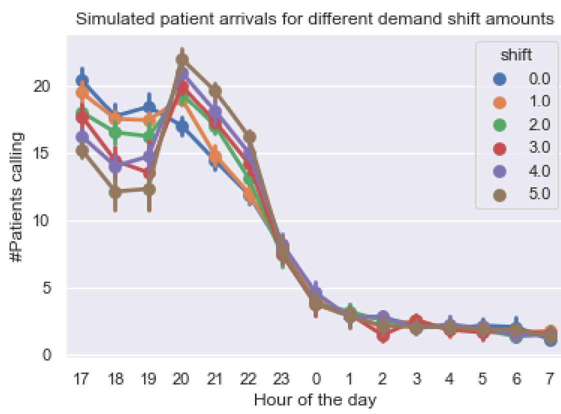

Figure B.30: Demand with different amount of demand shifting Friday

SUNDAY SCENARIO

In Figures B.31 and B.32, the results of different values of demand shifting on a Sunday are displayed. When statistically testing whether any of the demand shifts lead to a statistically significant waiting time reduction, a shift of 1 patient on Sundays significantly helps to slightly reduce waiting times, but as the waiting times are already not that high, shifting a patient does in general not have a big impact.

WEEKDAY SCENARIO

In Figures B.33 and B.34, the results of different values of demand shifting on a Weekday are displayed. When statistically testing whether any of the demand shifts lead to a statistically significant waiting time reduction, a shift of 1 patient on Weekdays significantly helps to slightly reduce waiting times especially around hour 18 and 19, but as on weekdays the shift can only be a few hours later as otherwise the calls are shifted to the night, the impact of shifting is not that big. It is however bigger than on a Friday, this could be related to the

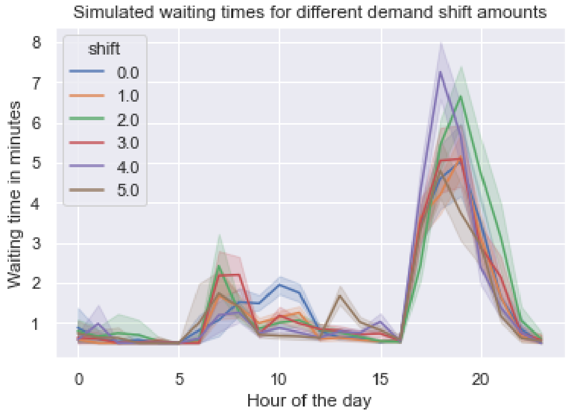

Figure B.31: Length of stay in queue with different amount of demand shifting Sunday

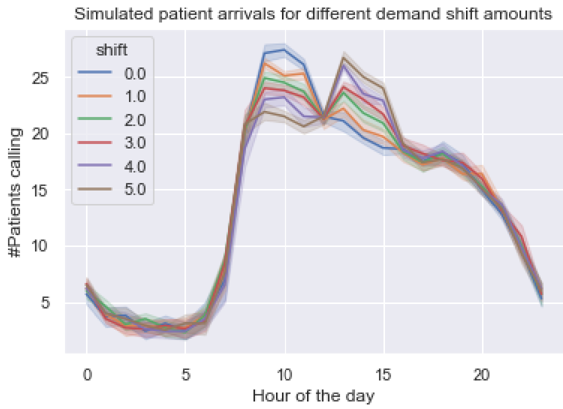

Figure B.32: Demand with different amount of demand shifting Sunday

fact that it is less busy on weekdays, so a larger proportion of demand is shifted away with a shift of 1 compared to Friday.

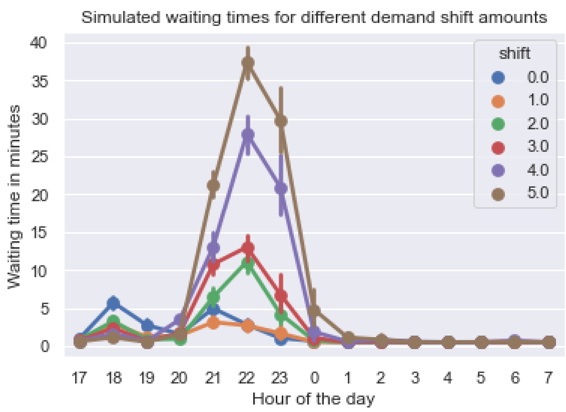

Figure B.33: Length of stay in queue with different amount of demand shifting Weekday

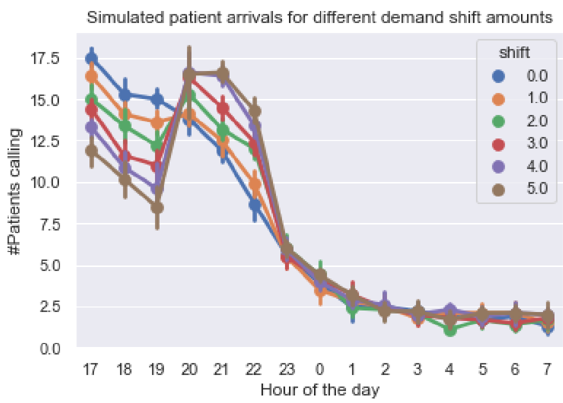

Figure B.34: Demand with different amount of demand shifting Sunday

SECOND DEPARTMENT SCENARIO

In Figures B.35 and B.36, the results of different values of demand shifting on a Saturday for the second department are displayed. When statistically testing whether any of the demand shifts lead to a statistically significant waiting time reduction, it is found that a shift of 1 or 2 significantly helps to slightly reduce waiting times by around 10-30% over the day, visible in the figure when comparing the blue with the orange and the green line.

B.3.5 Changes in service times

Lastly, the effects of service time reduction and increases within other model scenarios were tested. It can be concluded that increasing and decreasing service times yield the same outcomes for different scenarios: disruptive effects for increases, but beneficial effects for decreases. All statistical tests yield at least a statistically significant waiting time reduction of 40%, often 50%, already when a service time decrease of 10% takes place.

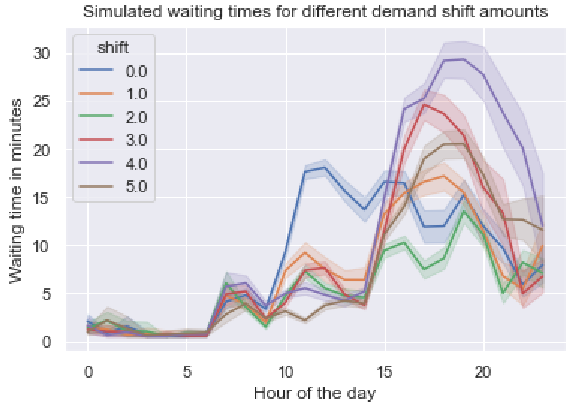

Figure B.35: Length of stay in queue with different amount of demand shifting second department Saturday

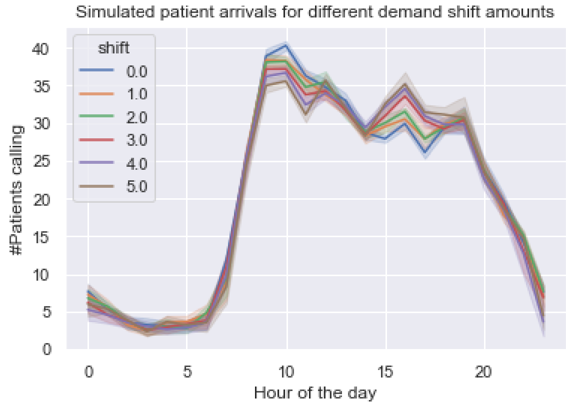

Figure B.36: Demand with different amount of demand shifting second department Saturday

FRIDAY SCENARIO

In Figures B.37 and B.38, the results of different values of service times changes on a Friday are displayed.

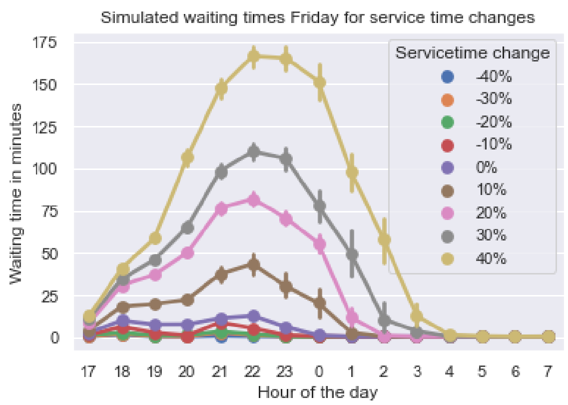

Figure B.37: Waiting time with different levels of demand increases and reductions - Friday

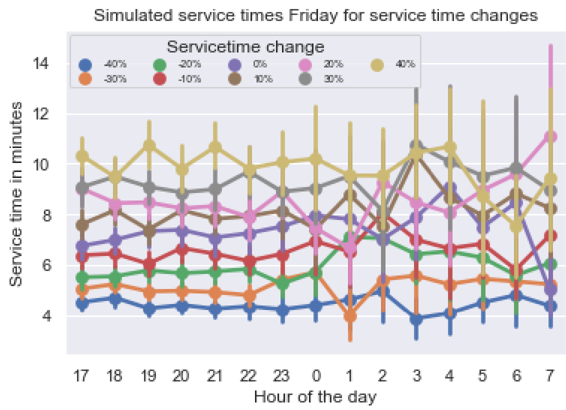

Figure B.38: Demand with different levels of demand increases and reductions - Friday

SUNDAY SCENARIO

In Figures B.39 and B.40, the results of different values of service times changes on a Sunday are displayed.

WEEKDAY SCENARIO

In Figures B.41 and B.42, the results of different values of service times changes on a Weekday are displayed.

SECOND DEPARTMENT SCENARIO

In Figures B.43 and B.44, the results of different values of service times changes on a Saturday for the second department are displayed.

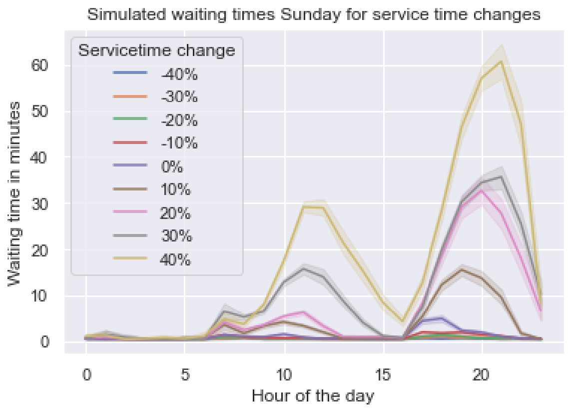

Figure B.39: Waiting time with different levels of demand increases and reductions - Sunday

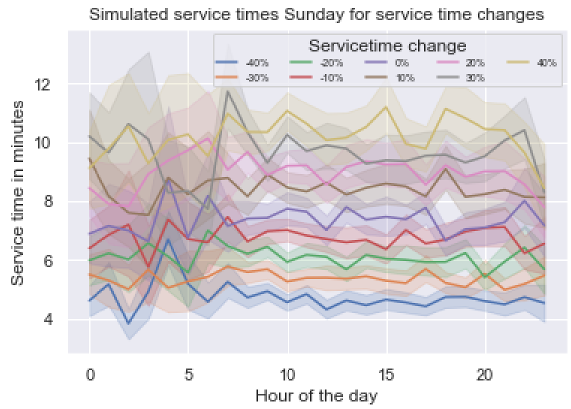

Figure B.40: Demand with different levels of demand increases and reductions - Sunday

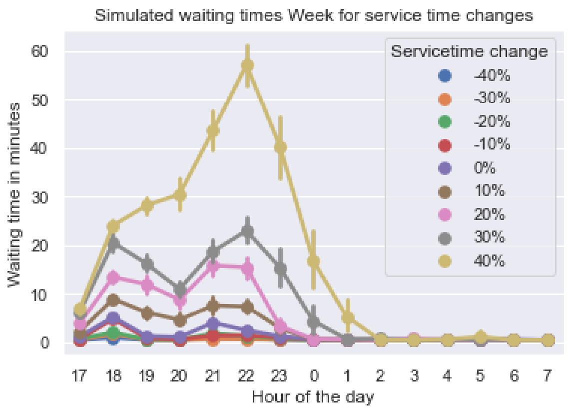

Figure B.41: Waiting time with different levels of demand increases and reductions - Weekday

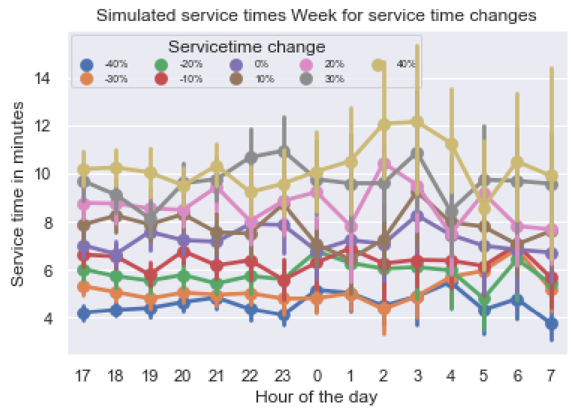

Figure B.42: Demand with different levels of demand increases and reductions - Weekday

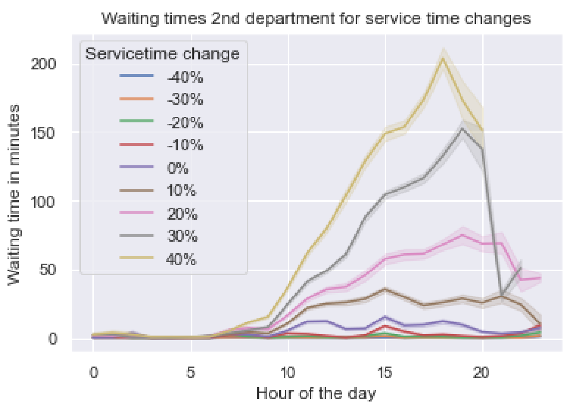

Figure B.43: Waiting time with different levels of demand increases and reductions - second department Saturday

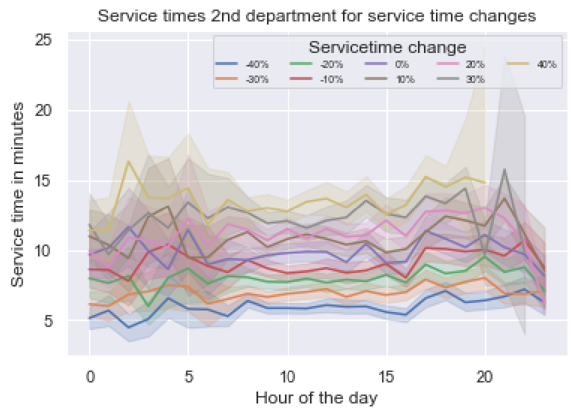

Figure B.44: Demand with different levels of demand increases and reductions - second department Saturday