UMC Utrecht
Wilhelmina Kinderziekenhuis

UNIVERSITY OF TWENTE.

Minimizing travel time in a Neonatal Care Network

Author:
R. Buter

Supervisors:
prof. dr. ir. E.W. Hans
dr. ir. A.G. Maan - Leeftink
dr. W.B. de Vries
dr. M.O. Blanken

November 18, 2019

MASTER THESIS

General information

Author
Robin Buter
r.buter@alumnus.utwente.nl

Educational institution
University of Twente
Faculty of Behavioural Management and Social Sciences
Department of Industrial Engineering and Business Information Systems
Center for Healthcare Operations Improvement and Research

Educational program
MSc Industrial Engineering and Management
Track: Production and Logistics Management
Orientation: Operations Management in Healthcare

Supervisors
prof. dr. ir. E.W. Hans
University of Twente
Center for Healthcare Operations Improvement and Research

dr. ir. A.G Maan-Leeftink
University of Twente
Center for Healthcare Operations Improvement and Research

dr. W.B. de Vries
Universitair Medisch Centrum Utrecht, location Wilhelmina Kinderziekenhuis
Department of Neonatology

dr. M.O. Blanken
Universitair Medisch Centrum Utrecht, location Wilhelmina Kinderziekenhuis
Department of Neonatology

Preface

Before you lies my thesis that has been written to complete my master's degree in Industrial Engineering and Management. Time has gone by fast and my time as a student has come to an end after 5 years. I would like to thank several people for helping me write and finish this thesis.

First of all, I would like to thank Willem and Maarten. I was motivated by your enthusiasm and interest in the operations management side of healthcare. I want to thank you for finding time in your schedule for discussing things and helping me write this thesis. I would like to thank Erwin and Gréanne for being my supervisors. Discussions with you provided me with new ideas and insights. I also want to thank you for the feedback you've given me, which allowed me to improve my thesis.

I would also like to thank all NICUs for sharing their operational capacity. In addition, I want to thank the people working at Perined for providing us with a dataset of the number of births.

I hope you enjoy reading my thesis.

Robin Buter,
Enschede, November 2019

Management summary

Background During a pregnancy it is possible that a child is prematurely born (<37 weeks). In such a situation, the child might need intensive care and monitoring in a hospital, depending on his condition at birth, such as for example his gestational age and weight. Currently in the Netherlands, clinical guidelines mandate to start active treatment *only* for newborns that are 24 weeks or older (de Laat et al. (2010)). Parents are heavily involved in both the treatment decisions and the care for their child.

If a newborn's condition is severe, he might require more complex care or surgery than general hospitals can provide. This means that the patient will be transferred to one of the nine Neonatal Intensive Cares Units (NICUs), mainly located in academic hospitals. This research project takes place at one particular NICU, namely the one of the Wilhelmina Children's hospital (WKZ) in Utrecht.

Every general hospital is assigned to one primary NICU. During this transport to a NICU, the newborn is accompanied by a doctor and a nurse. The primary NICU is responsible for providing the transportation for hospitals in their own region.

Neonatology care is not only complex medically, but also logistically. If a NICU is fully occupied and a new request for a bed comes from a NICUs own region, then that NICU department is responsible for finding a new place to admit this patient at another NICU. Transferring a patient is not only stressful for the family, but is also expensive and time consuming for the NICU staff.

Goal and methods Our research goal is to minimize the Neonatal Intensive Care travel and transport time by optimizing the assignment of general hospitals to NICUs.

Using data from Perined on the number of births at each hospital, we estimated the expected NICU demand of each hospital. In addition we gathered all travel times between hospital and NICUs. And finally, we obtained and visualized the current assigned of hospitals to NICUs.

We formulated an Integer Linear Programming (ILP) model for both an uncapacitated and capacitated scenario. Transfers of patients are not included in these two models.

To include transfers, we modeled the NICUs as a network of queues $M|M|c|c$, meaning there is no waiting room. Each NICU has their own Poisson arrival process for patients of their region. In case such a patient must be rejected because the NICU is fully occupied, the patient is admitted at another NICU. This new NICU is found using a predefined prioritization matrix.

We used two different methods to analyze this network of queues. The first method we used is the Continuous-Time Markov Chain (CTMC). CTMC is unable to solve large instances in reasonable time, so therefore we introduced Discrete Event Simulation (DES) as a second method.

In the CTMC the prioritization matrix is used for transferring arriving patients to other NICUs, if their primary NICU was fully occupied. We used the steady state distribution π and the PASTA property to construct an admission table which we could use to calculate the total travel time.

The number of the states in the CTMC formulation increase exponentially with the included number of NICUs. We can find a lower bound of the number of admission at each NICU using the steady state distribution of individual $M|M|c|c$ queues. This reduces the total state space, but still only allowed us to evaluate networks of at most five NICUs in reasonable time.

We introduced Discrete Event Simulation as the second method to solve larger instances. This model, after a sufficient run time, approximates the result obtained from a CTMC. DES also enabled us to use hospital-to-NICU prioritization.

When designing an optimization heuristic, we had to take several points into account. The search space is extremely large (4.1×10^{70} unique solutions) and contains many bad solutions. Evaluating one solution is slow and takes at least 20 seconds (only 180/hour).

We introduced an optimization heuristic consisting of three steps. First, we decrease the search space by disallowing the combination of certain hospitals and NICUs. In the second step we use the metaheuristic Reduced Variable Neighbourhood Search (RVNS) to quickly find a good quality solution. In the third step, after applying RVNS, we used steepest descent with a 1-move neighborhood search until no improvements can be found.

Results We used the best assignment resulting from the deterministic model without capacity restrictions as the absolute lower bound of the travel time. The difference in travel time between this lower bound and the current solution is undesirable travel time and indicates how much improvement is possible.

Using a stochastic model, we found for the current situation a travel time of 73.47 minutes on average per patient and 690 transfers per year. By applying our optimization heuristic, we found a travel time reduction of approximately 4.6 minutes on average per patient compared the to current situation. Compared to the lower bound, this is a reduction of 25.8% of the undesirable transport time. In addition, the number of transfers are reduced by 15.7%.

By reallocating all 163 beds optimally among the NICUs, we found an average travel time of approximately 65.08 minutes per patient. Compared to the current situation, we found a reduction of 47% in undesirable travel time and a decrease of 35.5% in the number of transfers.

It is clear that calculating the required capacity using deterministic demand (133 beds) leads to a severe underestimation of the travel time and number of transfers. Furthermore, it seems that increasing capacity moderately from 163 to 170 beds, can still result in a significant and efficient reduction in transfers. Moving towards and beyond 190 beds results in high diminishing returns.

Conclusions and recommendations The current assignment can be evaluated and improved using operations research methods. Providing higher quality data for input parameters will increase confidence in the values of the performance indicator (travel time). If the total state space allows for it, the CTMC method is preferred because there is no uncertainty in the mean value of the chosen performance indicator. Furthermore, DES scales well with number of NICUs, while CTMC scales well with number of patients (arrival rate).

We recommend the Neonatal Care Network to evaluate the current assignment of hospitals to NICUs, for example once a year. Doing this will result in a better match of current available capacity and demand, and keep transfers of patients to a minimum. In addition, setting performance targets such as the number of transfers within the network should be based on total capacity in the network. We provided a reference points for setting realistic targets.

Further research might be on the topic of where to locate NICUs and/or specialty care. In addition, more research on online operational decisions regarding transferring patients, depending on the state of the network, might prove useful.

Abbreviations and definitions

WKZ	Wilhelmina Kinderziekenhuis (Wilhelmina Children's Hospital)
NIC	Neonatal Intensive Care
NICU	Neonatal Intensive Care Unit
HC	High Care
MC	Medium Care
General hospital	A non-academic hospital
Patient[1]	A newborn that requires treatment in a NICU. Patient might also refer to a pregnant woman that must be admitted to a maternity ward, depending on the context.
Neonatal Intensive Care Network	We define this as the network of all nine[2] NICUs.
Neonatal Care Network	We define this as the network of the NICUs and the general hospitals, with regard to providing care for very ill neonates.
Perinatal center	Birth centre with a NICU

In this report, a newborn is referred to as "he", for sake of simplicity and consistency. When referring to a NICU, we use the name of the hospital and name of the the city interchangeably (e.g."Utrecht" and "WKZ").

[1]Patient/client are interchangeable. We choose to use patient in this report.
[2]Following the recent merge of the hospitals AMC and VUmc in Amsterdam, we merge their NICUs as well.

Contents

Chapter 1

Problem introduction

The care for prematurely born children in the Netherlands is of high quality. Many medical professionals work day and night to provide the complex care these still fragile children deserve. However, capacity problems put the neonatal intensive care under pressure. They experience that too many newborns and pregnant women must be transferred to another birth centre, which is undesirable for all parties involved.

This chapter analyzes this problem and introduces our approach to solving it. In Section 1.1 we identify several core problems we could try to solve, of which we choose one. Section 1.2 formulates and explains our problem solving approach.

1.1 Problem formulation

We start by giving background information that is required to understand the context of the problem in Section 1.1.1. After that, we analyze the problem in detail and motivate our chosen core problem to solve in Section 1.1.2.

1.1.1 Background information

During a pregnancy it is possible that a child is prematurely born (<37 weeks). In such a situation, the child might need intensive care and monitoring in a hospital, depending on his condition at birth, such as for example his gestational age and weight. Currently in the Netherlands, clinical guidelines mandate to start active treatment *only* for newborns that are 24 weeks or older (de Laat et al. (2010)). Parents are heavily involved in both the treatment decisions and the care for their child.

If a newborn's condition is severe, he might require more complex care or surgery than general hospitals can provide. This means that the patient will be transferred to one of the nine Neonatal Intensive Cares Units (NICUs), mainly located in academic hospitals. Every general hospital is assigned to one primary NICU. During this transport to a NICU, the newborn is accompanied by a doctor and a nurse. The primary NICU is responsible for providing the transportation and the staff.

If there is an indication that the unborn child might need a NICU bed, it is preferable that the pregnant woman is admitted to a birth centre with a NICU (perinatal centre), just before she gives birth. Then no transport is necessary and the newborn can be treated immediately.

If there is no available bed at the NICU for a new admission at the time of a request from their own region, this NICU is responsible for finding a place in one of the other NICUs. However, there is no overview of the free capacity in the network, which makes finding a new place difficult. If another NICU is contacted and denies the request, it will count as an additional rejection. This means that one patient can be rejected multiple times, before he is admitted somewhere. The initially assigned NICU is still responsible for the transport of the newborn.

This research project takes place at one particular NICU, namely the one of the Wilhelmina Children's hospital (WKZ) in Utrecht. Their neonatalogy department provides Intensive Care (IC), High Care (HC), and Medium Care (MC). A total of 24 beds are available for IC, divided into three units of eight beds, on the same floor. Of those 24 beds, 20 were operational at the time of writing due to staff shortages.

A few other research projects took place at the WKZ before. Most notably, Oude Weernink (2018) provides us with a recent analysis of the logistical processes at the department. Moreover, at the time of writing, WKZ is developing a model to predict whether a pregnant woman will give birth within a certain time period from now. This will hopefully support the decision making process for NICU admissions.

1.1.2 Problem description

For years the Neonatology department of the WKZ has been struggling with employing sufficient personnel (see Hoek (2015), Otten (2017), Oude Weernink (2018)). The type of care that is provided at a NICU requires highly qualified nurses and doctors. The resulting problem the department faces is that ***too many requests for a NICU bed must be rejected and too many pregnant women are transferred to other birth centres***.

Rejecting and transferring patients have several consequences. Transferring a pregnant woman to another hospital means that time is wasted on preparing for a birth that eventually takes place somewhere else, with each transfer potentially resulting in a loss of information. This loss of information might mean that time is spent on procedures or checks that had already been performed in the last hospital. And of course, the transfer itself is also a discomfort to the pregnant woman and her family.

Rejecting a request for a NICU bed might cause a delay in the start of a treatment of the newborn, which impacts the quality of care. In addition, if the newborn is admitted to another NICU, then that location is likely farther away from the parents' house than the original NICU. Transferring and rejecting a patient leads to even more time spent on transportation, which could be spend more productively.

Since finding additional qualified personnel while adhering to the same department budget is not possible, available capacity must be used more efficiently. In addition, quality of labor and care must not deteriorate, but preferably improve. For this purpose, five core problems are identified using root cause analysis by interviewing the problem owners.

We will discuss each core problem in the following paragraphs. Appendix A visualizes the effects and underlying causes of the action problem.

1. At the moment, there is no overview of national and regional capacity in the Neonatal Intensive Care Network. This makes finding a new place for a rejected request for an admission difficult. It results in hectic phone calls to other (somewhat arbitrary) NICUs, asking if they have a free bed. If they do not, then it counts as another rejection. Because all NICUs must make decisions based on limited information, the capacity of the network is not used optimally.

2. The admission policy of WKZ is currently insufficiently based on (objective) acuity of the patients. Hoek (2015) has already developed a neonatal acuity measurement model for the WKZ. However, more research is required to validate this model. By basing the admission policy on acuity, the *true* capacity of the department can be utilized, while improving quality of labor and care.

3. The NICUs receive a reimbursement for every day a bed is occupied by a patient. In theory, this would mean that financially the best decision would be to keep the patient as long as possible. In reality, this fortunately does not occur and the patient is transferred back to the general hospital as soon as safety and logistics allows it. Contradictory, while providing emergency care, the NICUs are not compensated for providing accessibility of that care. This system puts financial pressure on the NICUs.

4. Demand for neonatal care is inherently highly variable, which makes planning for the department difficult. However, some of the variability might be predictable, but it is unknown to what extent. At the moment, a project team of the WKZ is already investigating whether prediction of demand for a NICU bed from the hospital's own population can be improved.

5. As mentioned before, every general hospital is assigned to one NICU. However, this assignment is historically determined and might not be optimal with regards to travel time and (current) available capacity at each NICU. Time spent on transportation should be minimized, since it is an expensive resource, requiring both ambulance and NICU personnel.

Because of the the time constraints of this project, we can only choose one core problem to solve. Problem (1) and (3) require full support, cooperation, and involvement of all nine NICUs, which is not feasible for this project at the moment. As mentioned before, problem (4) is already investigated at WKZ. Of problem (2) and (5), we think solving problem (5) will have more impact than problem (2).

Therefore, the focus of the project is on the NIC-transport between all general hospitals and the nine perinatal centres. Special attention will be paid to analyzing transfers within a network, available capacity at the NICUs, stochasticity of demand from the general hospitals, and travel time.

Using the framework of Hans et al. (2012), this problem could be interpreted as tactical resource capacity planning. The assignment of general hospitals to NICUs is historically determined. No methodology for planning and control is applied, meaning there is no (periodic) evaluation of the match between *current* capacity and demand at the locations.

1.2 Problem approach

In this section we formulate our problem solving approach. In Section 1.2.1, we start by defining the research goal and its research objectives. To achieve this research goal, we formulated several research questions, and a plan of approach to answering them in Section 1.2.1. Afterwards, we discuss the scope in Section 1.2.3.

1.2.1 Research goal and objectives

Our research goal is to minimize the Neonatal Intensive Care travel and transport time by optimizing the assignment of general hospitals to NICUs, and to provide an independent perspective from outside of the Neonatal Care Network.

Our research objectives are therefore:

1. to analyze the current situation in the Neonatal Care Network;
2. to develop and test mathematical models to determine the optimal catchment areas of the NICUs for different scenarios;
3. and to make a recommendation for improvement.

Achieving our research goal will hopefully lead to less time spent on transportation and better match between demand and capacity at each NICU. As a result, fewer requests for admissions will be rejected, meaning that the quality of care will improve and that the newborn will be treated closer to its parents' home.

1.2.2 Research questions and methodology

We would have to answer the following main research question to achieve our research goal:

"Which assignments of general hospitals to Neonatal Intensive Cares lead to minimized transportation time?"

We will answer this question by modeling the key characteristics of the Neonatal Intensive Care Network using operations research methods. We will use mathematical optimization techniques to find a (close-to) optimal assignment. We will model the NICUs as a network of nine interconnected $M|M|c|c$ queues, in which in case of rejection a new NICU must be found for this patient. In particular, we will investigate how we can analyze transfers between NICUs, and how demand and capacity allocations in this network affect the total travel time. A thorough understanding of the problem context and the used methods is required. To guide this process and to systematically answer the main research question, the main question is split into multiple smaller sub-questions.

1. ***What models or methods are commonly used to determine or evaluate catchment areas of health care facilities, taking travel time into account?***
   Question *1* is answered in Chapter 2 by means of a literature review. The initial set of articles is constructed using broadly defined search terms on Scopus, which is afterwards filtered on relevance. Additional articles are included by backward reference searching.

2. ***How can uncertainty in demand of health care services be included in mathematical programming?***
   Question *2* is also answered in Chapter 2 by means of a literature review. Uncertainty of demand is an important characteristic of neonatal intensive care, which should somehow be incorporated into the models to find a robust solution.

3. ***What is the current situation in the Neonatal Care Network?***

   (a) *How is Neonatal Intensive Care organized in the WKZ?*

   (b) *How is the Neonatal Intensive Care Network organized?*

   (c) *How is the Neonatal Care Network organized?*

   (d) *How is accessibility to Neonatal Intensive Care related to travel time from the parents' house?*

   These questions are answered in Chapter 3. First, we take the perspective of one particular NICU, namely the WKZ. The WKZ provides us with expert opinions and opportunities to visit the NICU itself. Second, we broaden our perspective from one NICU, to the network of all NICUs. Third, we include the general hospitals as well, and look at the whole Neonatal Care Network. And finally, we want to know how accessibility to NIC is related to travel time, from the parents' perspective, using methods found in literature. Many publicly available resources can be used to help us get insight into the nationwide situation, such as statistics from the CBS or Perined, and annual reports of hospitals.

Analyzing the current situation allows us to obtain a deeper understanding of the problem context. This helps identifying core characteristics of the problem that should be modeled, and helps identifying which simplifications and assumptions can, or must, be made. In addition, it allows us to gather input data for the model.

4. ***How can we analyze a network of*** $M|M|c|c$ ***queues in which rejected patients must be relocated?***
   This question is answered in Chapter 4. Results from Chapter 2 and 3 will be used here as input for the model. In addition, opinions from experts will be critical in validating the model and the chosen approach for analyzing it.

5. ***What are the effects of different assignments of general hospitals to Neonatal Intensive Cares?***

   (a) *What is the optimal assignment?*

   (b) *Where should we increase capacity and what would its effect be?*

   (c) *What is the optimal assignment, given that it is allowed to change the allocation of nationwide capacity at the NICUs?*

   These questions are answered in Chapter 5. The results of the models and scenarios formulated in Chapter 4 are analyzed and discussed. First we want to know what the optimal assignment would be. In addition, we want to investigate the impact of adding capacity at the optimal location and decreased transportation time. And lastly, we want to know what is the best we (hypothetically) could do with the available nationwide capacity, with regards to transportation time.

In Chapter 6 we conclude our research and answer the main research question. In addition to that, we will give our recommendations to the Neonatal Care Network, discuss the limitations of this research, and give suggestions for further research.

While these sub-questions are answered chronologically in this report, the research process, however, is not be linear. Figure 1.1 shows the main research activity in each chapter and how these activities are related. Since formulating, verifying, and validating a model is iterative, it might require taking a step backwards in the chain. For example, different data might be required for a model formulation than initially obtained.

1.2.3 Scope

This project focuses on all nine NICUs in the Netherlands. Any MC or HC departments of the perinatal centres are excluded. Capacity and logistical processes of the general hospitals are not taken into account. The physical locations of the NICUs are assumed as fixed. In addition, the islands Texel, Terschelling, Vlieland, Schiermonnikoog, and Ameland are excluded from analyses and models.

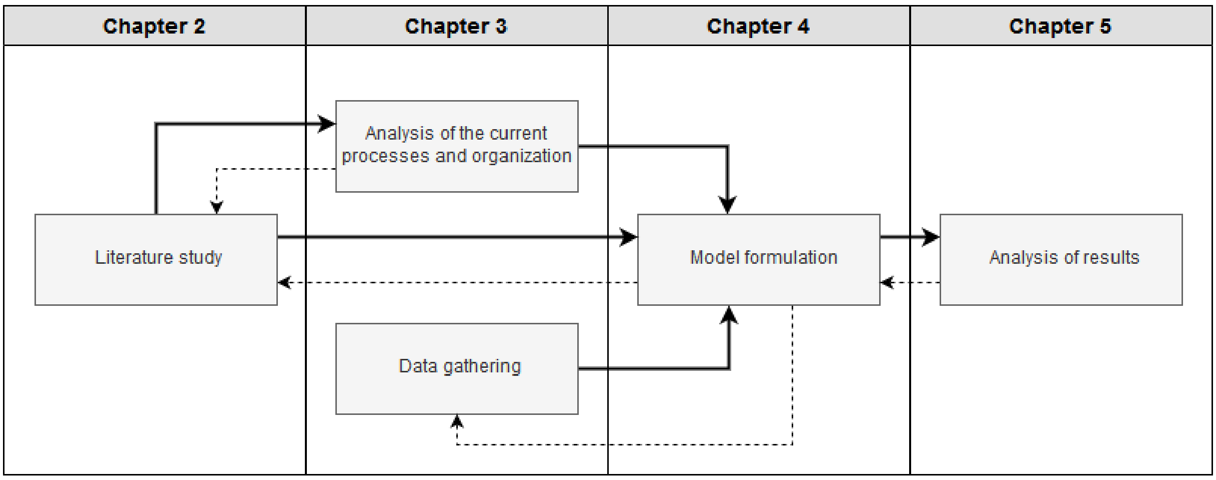

Figure 1.1: The main activities of each chapter and their relationship

Chapter 2

Literature review

In this chapter, the first and second research questions are answered by means of literature search. The question *"What models or methods are commonly used to determine or evaluate catchment areas of health care facilities, taking travel time into account?"* is answered in Section 2.1. In Section 2.2 we discuss the Generalized Assignment Problem, and some of its applications in health care. The question *"How can uncertainty in demand of health care services be included in mathematical programming?"* is answered in Section 2.3.

2.1 Catchment areas

In literature, catchment areas and spatial accessibility of health care are often intertwined. Spatial accessibility is commonly defined as a combination of availability (volume) of care, and accessibility (distance) to care (e.g. Guagliardo (2004); Delamater et al. (2019)). Floating Catchment Area (FCA) is a family of methods that can be used to measure spatial accessibility to health care, and are easy to interpret.

From a broad literature search on models and methods with regard to catchment areas of health care facilities, we conclude that most of the relevant literature concerns spatial accessibility to health care and use a (new) variant of FCA as a metric thereof. The details on the approach of this literature search can be found in Appendix B.1.

In Sections 2.1.1, 2.1.2, and 2.1.3, we discuss three main developments in FCA methods. In Section 2.1.4 we mention some, *but not all*, variants or applications of FCA methods. We discuss other, non-FCA, methods we found in Section 2.1.5. And finally, we discuss the relevance of the found literature to our research, in Section 2.1.6.

2.1.1 Two-Step Floating Catchment Area

Initially, a method called the Two-Step Floating Catchment Area (2SFCA), of which Radke and Mu (2000) laid the groundwork, is further developed and popularized by Luo and Wang (2003). As the name suggests, this method is performed in two steps.

The first step Luo and Wang (2003) formulated is to calculate the ratio of supply (physicians) to demand (population), called R_j. For each location at which physicians work (j), the capacity S_j is divided by the total population that can travel to this location within a certain time threshold d_0, say 30 minutes.

$$R_j = \frac{S_j}{\sum\limits_{k \in \{d(k,j) \leq d_0\}} P_k}$$

The second step identifies for every demand location i, all supply locations j that can be reached within the time threshold used in step 1. The accessibility score of a demand location is then the sum of the supply-to-demand ratios (calculated in step 1) of those identified supply locations.

$$A_i = \sum_{j \in \{d(i,j) \leq d_0\}} R_j$$

2.1.2 Enhanced Two-Step Floating Catchment Area

For 2SFCA, the assumptions that patients will not travel further than the distance threshold and that all patients within this area have equal access to care, are often critiqued (e.g., Luo and Wang (2003); Luo and Qi (2009); Wan et al. (2012); Ma et al. (2018)). In addition, it seems that accessibility is overestimated in areas where multiple health facilities overlap (Luo and Qi (2009)).

Therefore, Luo and Qi (2009) developed the Enhanced Two-Step Floating Catchment Area method (E2SFCA), in which multiple travel time zones are defined and weighted differently. Shi et al. (2012) proposed a Gaussian function for obtaining a set of weights, making the weights form a bell shaped curve. Luo and Qi (2009) argued that using discretized weights might be preferred to using a continuous distance decay function, because people are indifferent to a small difference in travel time when driving to a health care facility. To give an example of a continuous distance decay function, Guagliardo (2004) used a kernel density function to model this effect. Just like 2SFCA, E2SFCA is a more intuitive variant of the gravity model, of which it is originated from.

The steps of E2SFCA are similar to those of 2SFCA. However, in the first step the time threshold d_0 is divided into R smaller intervals. D_r is the rth travel time zone of the catchment area for some supply location j. Then, the weighted ratio of physicians to population can be calculated, using weight W_r for the rth travel time zone.

$$R_j = \frac{S_j}{\sum\limits_{r=1}^{R} \sum\limits_{k \in \{d(k,j) \in D_r\}} P_k W_r}$$

In the second step, A_i is calculated for every demand location i using the same travel time zones and weights as in step 1.

$$A_i = \sum_{r=1}^{R} \sum_{j \in \{d(i,j) \in D_r\}} R_j W_r$$

2.1.3 Three-Step Floating Catchment Area

Another variation, which builds on E2SFCA, is called the Three-Step Floating Catchment Area method (3SFCA), proposed by Wan et al. (2012). The authors introduced competition between service providers into the model, resulting in less overestimation of demand when multiple facilities overlap in catchment area.

In the first step, for every demand location i, the total catchment area limited by d_0 is divided into R intervals which are given weights W_r. A selection criteria G_{ij} is calculated between demand location i, and supply location j that is within the catchment area. T_{ij} and T_{ik} are the weights for supply locations j and k. To illustrate, in case demand location i can reach two supply locations within the time threshold d_0, the supply location that is closer will be given a larger share of the demand, unless both supply locations are in the same travel time zone D_r .

$$G_{ij} = \frac{T_{ij}}{\sum_{k \in \{d(i,j) \leq d_0\}} T_{ik}}$$

In the second step, for every supply location j, the total catchment area bounded by d_0 is divided into R intervals. The weighted physician to population ratio R_j can be calculated using weights W_r.

$$R_j = \frac{S_j}{\sum_{r=1}^{R} \sum_{k \in \{d(k,j) \in D_r\}} P_k W_r G_{kj}}$$

And in the third step, the spatial accessibility index of demand location i is computed.

$$A_i = \sum_{r=1}^{R} \sum_{j \in \{d(i,j) \in D_r\}} R_j W_r G_{ij}$$

2.1.4 Some variations and applications of FCA

FCA methods can be used for more than only identifying regions that lack accessibility to health care. For example, Delamater et al. (2019) compared multiple FCA metrics for predicting the destination hospital for hospitalizations originating from a certain ZIP code. Calovi and Seghieri (2018) used 2SFCA to evaluate three different interventions for reorganizing outpatient care by comparing accessibility to care.

Other authors proposed modifications to existing FCA methods. For example, Ma et al. (2018) improved 3SFCA by more accurately predicting the demand for health care in a region by distinguishing different age groups (e.g. the elderly require more care). In addition, real-time travel information is used for comparing accessibility during several time periods, showing that accessibility is dynamic. Kim et al. (2018) proposed Seoul Enhanced 2-Step Floating Catchment Area, which reduces the overestimation of accessibility from E2SFCA for regions with high population and hospital density. Cheng et al. (2016) identified sub-districts in Shenzhen that lack access to high level hospitals. They used a Kernel Density Two-Step Floating Catchment Area method, which is equivalent to E2SFCA with a continuous impedance function for travel time.

2.1.5 Other methods and models

Geographical Information Systems can be used to analyze networks of supply and demand, and their geographical relationship. For example, Murad (2007) used a GIS application for exploring the location of hospital demand. Service areas of 15 minutes travel time were visualized for hospitals. Schuurman et al. (2006) developed a GIS tool to model geographical catchment areas of rural hospitals, based on travel time.

Other methods require a large dataset of for example hospitalizations. Xiong et al. (2018) used population and hospitalization data to calculate hospitalization probabilities, which are used to determine the sphere of influence of the top hospitals in Shanghai. Gilmour (2010) used K-means clustering to allocate local authority districts to the catchment area of a certain hospital, based on a multivariate dataset. Klauss et al. (2005) performed a patient origin study in Switzerland by using small area analysis (SAA). Regions were assigned to its most frequent hospital provider region. Afterwards, hospital service area were obtained.

King et al. (2019) formulated a location-allocation model to determine the optimal allocation of general hospitals to current pediatric intensive care retrieval teams, minimizing travel time. However, capacity and availability of those retrieval teams were not taken into account, and the focus lies on evaluating how much of the demand is accessible within a certain amount of time.

2.1.6 Discussion

Using FCA methods to measure accessibility to health care seems to become increasingly popular. Searching for *"Floating Catchment Area"* on Scopus shows that the majority of these articles have been published since 2015 and are still growing in numbers in 2019, despite the fact that one of the initial popular works is from 2003. However, only a handful of those paper show the application of FCA to specifically emergency medical services (see Xia et al. (2019), Shin and Lee (2018), Rocha et al. (2017), and Tansley et al. (2016)). To the best of our knowledge, it has also not been applied to a situation in the Netherlands.

To measure the difference in accessibility to Neonatal Intensive Care, either E2SFCA or 3SFCA seem most suitable to start with. Coordinates and population statistics of municipalities are publicly available. Travel times between NICUs and coordinate centroids of municipalities can be gathered using open source software. Other methods and variations of FCA are either too specific or require hospitalization data.

At first sight, the location-allocation model of King et al. (2019) seems unsuitable to our project since evaluating locations of NICUs is out of scope. However, the set of possible locations can be reduced to the current ones. In addition, Ross and Soland (1977) has shown that many of the important location-allocation models can be rewritten as Generalized Assignment Problems.

2.2 The Generalized Assignment Problem

The Generalized Assignment Problem (GAP) has been researched since the 1970s and can be formulated as assigning tasks to agents, such that

- each task is assigned to exactly one agent;
- the required resources of the assigned tasks to an agent do not exceed the agent's capacity;
- and the total cost (profit) of all assignments is minimized (maximized).

This means that multiple tasks can be assigned to one agent. If the number of tasks and agents are equal, then the problem is reduced to the *assignment problem*. The GAP can be formulated as an ILP as follows (see any paper on this topic, e.g. Fisher et al. (1986), Nauss (2003), etc):

Consider the case that n tasks must to be assigned to m agents, assuming $n \geq m$. Define $c_{i,j}$ as the cost of assigning task j to agent i, $d_{i,j}$ as the resources required for task j if performed by agent i, and b_i as the capacity of agent i.

$$
\begin{aligned}
\text{Decision variable} \quad & X_{i,j} = \begin{cases} 1 & \text{if agent i performs task j} \\ 0 & \text{otherwise} \end{cases} \\
\text{minimize} \quad & \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j} X_{i,j} \\
\text{subject to:} \quad & \sum_{j=1}^{n} d_{i,j} X_{i,j} \leq b_i \quad i = 1, \ldots, m \\
& \sum_{i=1}^{m} X_{i,j} = 1 \qquad j = 1, \ldots, n \\
& X_{i,j} = \{0, 1\} \qquad \forall i, j
\end{aligned}
$$

For a comprehensive overview of applications of GAP, we refer to Öncan (2007). The author also discusses eleven modifications to the base formulation. In particular, the following variants might be interesting for our case:

- Bottleneck GAP
- Stochastic GAP

An advantage of formulating the problem as a GAP is that the objective function and constraints can easily be modified or extended for different scenarios, e.g. to include uncertainty of parameters. In addition, the formulation and results are intuitive and easy to interpret for this problem. Furthermore, it is possible to find exact solutions. However, the GAP is NP-hard (Fisher et al. (1986)), which means finding an exact solution for this problem size is not guaranteed within reasonable time. Therefore, meta-heuristics might have to be used for finding an acceptable solution within time limits.

2.3 Demand uncertainty

A literature search is performed to investigate how uncertainty in parameters can be included in a mathematical programming formulation, for a health care application. Table 2.1, at the end of this section, summarizes the results. Details on the approach of this literature search are found in Appendix B.2.

We categorize and discuss the literature based on the method that is used to incorporate stochasticity in models. In Section 2.3.1 we discuss the application of fuzzy variables. Scenario based formulations and robust formulations are mentioned in Sections 2.3.1 and 2.3.3, respectively. We discuss other methods in Section 2.3.4.

2.3.1 Fuzzy variables

Fuzzy sets are first introduced by Zadeh (1965). Fuzzy variables are imprecise and vague. For example, one might find the weather 'hot', and someone else might say it is just 'warm'. Sadatasl et al. (2017) proposed a model for facility location and network design. A back up facility is assigned to each facility for demand that could not be fulfilled. Demand is considered uncertain and is therefore included as triangular fuzzy numbers. Ahmadvand and Pishvaee (2018) formulated a Credibility-based Fuzzy Common Weights Data Envelopment Analysis method to match available kidneys for transplantation to patients. Fuzzy variables were used to incorporate uncertainty in input variables such as transportation time and laboratory measurements.

2.3.2 Scenarios

If the probability distribution of a parameter is known, a set of scenarios may be generated from that distribution and can be incorporated into a model. For example, Vieira et al. (2018) allocated radiation therapy technologists to operations by means of stochastic Mixed Integer Linear Programming. Several scenarios of patient arrivals were generated from the Poisson distribution. Results, using real data, show that less capacity will be required and more patients will be treated within waiting time limits.

Liu et al. (2015) proposed a stochastic planning model for medical resources order and shipment scheduling, in which scenarios were generated according to a probability distribution. Cardoso et al. (2015) incorporated health gains of long term care into a location-allocation model. A scenario tree was constructed using empirical distributions and represented combinations of stochastic parameters. Koppka et al. (2018) precomputed probabilities of the operating room finishing on time at the end of the day, for each combination of patients assigned to an OR and the OR capacity. Afterwards, scenarios of daily arrivals are weighted according to the case mix of the hospital.

Stochastic formulations can be transformed into a deterministic formulation by means of sample average approximation (SAA), if probability distributions of parameters are known. Wang et al. (2014) applied SAA in operating theater allocation, Daldoul et al. (2018) in allocating resources in an emergency department, and Bagheri et al. (2016) in developing a nurse schedule.

2.3.3 Robust formulation

A robust formulation can be used to make a trade-off between conservativeness of the solution and the objective value. Tang and Wang (2015) proposed a model for allocating OR capacity to subspecialties, and to decide how much capacity to reserve for emergencies. Demand is assumed to be uniform, based on historical lower and upper bounds. Conservativeness of the model can be adjusted by setting a limit on the total demand in a scenario. The worst-case revenue loss is minimized. Karamyar et al. (2018) formulated a bi-objective model that minimizes the total cost of locating facilities, and minimizes the completion time of demand. The costs are uncertain with an ambiguous distribution. The authors proposed an algorithm in which the problem is divided into two parts, and solved by using Simulated Annealing and Benders decomposition, sequentially. Zarrinpoor et al. (2018) suggested a location-allocation model of an hierarchical hospital network. Several disruptive scenarios were formulated. The model is solved by using Benders decomposition.

2.3.4 Other

Vidyarthi and Jayaswal (2014) modeled a system as a network of independent M/G/1 queues. An integer program was formulated, in which waiting time was penalized by a cost. The formula for waiting time was linearized at the expense of additional variables and constraints. Carello and Lanzarone (2014) uses a cardinality-constrained approach for guaranteeing the solution can deal with the worst scenarios. An advantage of this approach is that no probability distributions or scenarios have to be assumed. Given uncertain service times, Wang et al. (2017) proposed a chance-constrained model for surgery planning. These constraints limit the probability of overtime and are independent of a type of distribution.

Authors	Problem	Objective	Method	Uncertain parameters	Solving approach
Carello and Lanzarone (2014)	A	Cost	Cardinality constraints	Demand	CPLEX
Vidyarthi and Jayaswal (2014)	LA	Cost	Queuing theory	Demand (Poisson) Service time (general)	CPLEX
Wang et al. (2014)	RA	Cost	Scenarios	Service time (lognormal) Interarrival time (exp)	CPLEX, SAA
Cardoso et al. (2015)	LA	Cost, health gains	Scenarios	Care requirement (emperical) LOS (emperical)	CPLEX
Liu et al. (2015)	S	Cost	Scenarios	Demand	CPLEX
Tang and Wang (2015)	RA	Worst case revenue loss	RF	Demand (uniform)	CPLEX, IAA
Bagheri et al. (2016)	S	Cost	Scenarios	Demand (discrete uniform) LOS (discrete uniform)	SAA
Sadatasl et al. (2017)	FL	Cost	Fuzzy variables	Demand	CPLEX
Wang et al. (2017)	RA	Cost	DRCC	Service time	CPLEX
Ahmadvand and Pishvaee (2018)	RA	Deviation of efficiency	Fuzzy variables	Transport time	Unspecified
Daldoul et al. (2018)	RA	Waiting time	Scenarios	Patient arrival (Poisson) Service time (normal, exp)	CPLEX, SAA
Karamyar et al. (2018)	LA, S	Cost, completion time	RF	Cost	CPLEX, Benders, SA
Koppka et al. (2018)	RA	Overtime or cancellations	Scenarios	Patient arrival (multiple)	Gurobi, DES
Vieira et al. (2018)	RA	Timely treated patients	Scenarios	Patient arrival (Poisson)	CPLEX
Zarrinpoor et al. (2018)	HLA	Cost	Scenarios, RF	Capacity Reliability Demand Referral rate Geographical accessibility	CPLEX, Benders

Table 2.1: A summary of papers found in literature. Probability distributions of parameters are mentioned, if applicable. A: Assignment; (H)LA: (Hierarchical) Location-allocation; RA: Resource allocation; S: Scheduling; FL: Facility location; SAA: Sample average approximation; DRCC: Distributionally robust chance constraint; RF: Robust formulation; IAA: implementer-adversary algorithm; Benders: Benders decomposition; SA: Simulated annealing; DES: Discrete event simulation.

2.4 Conclusion

The first research question *"What models or methods are commonly used to determine or evaluate catchment areas of health care facilities, taking travel time into account?"* is answered in this chapter.

Catchment areas and spatial accessibility of health care are frequently related to each other in literature. Spatial accessibility can be defined as a combination of availability (volume) of care, and accessibility (distance) to care. Floating Catchment Area (FCA) is a family of methods that can be used to measure spatial accessibility to health care, and are easy to interpret.

Three main developments in FCA methods are identified. First, Two-Step Floating Catchment Area (2SFCA) was developed. In the first step the supply-to-demand ratio is calculated for every supply location. Only the demand that can be reached with in a certain time threshold is taken into account. The second step calculates the accessibility score of every demand location by taking the sum of the score calculated in the first step of all supply locations that can be reached within the time threshold. Enhanced Two-Step Floating Catchment Area (E2SFCA) defines multiple travel time zones with decreasing weights. And finally, Three-Step Floating Catchment Area (3SFCA) introduces competition between supply locations. If more than one supply location can be reached from a certain demand location, a fraction of the demand is assigned to each supply location, proportional to its distance from the demand location.

The second research question *"How can uncertainty in demand of health care services be included in mathematical programming?"* is answered in this chapter.

There are multiple methods to include uncertainty of parameters in mathematical programming approaches. Scenarios can either be sampled from probability distributions of stochastic parameters, or formulated using for example expert opinions. Using a robust formulation, a trade-off between conservativeness and the objective value of a model can be made accordingly. Chance-constrained models can be used for dealing with risks or probabilities, such as limiting the probability of overtime. Queuing networks and expressions can be utilized in mathematical programming as well.

Chapter 3

Current situation

In this chapter, the third research question is answered: *What is the current situation in the Neonatal Care Network?*

To develop a valid model for assigning general hospitals to Neonatal Intensive Cares, we must first analyze the current situation. This includes obtaining a deeper understanding of the problem context, identifying characteristics of the problem that should be modeled, and gathering input data for the model.

This chapter is structured as follows. We start by describing the relevant key processes of one NICU (WKZ) in Section 3.1. After knowing how one individual NICU operates, we investigate how the network of Intensive Care is organized, in Section 3.2. And finally, we analyze the complete Neonatal Care Network in Section 3.3, which also includes the general hospitals in addition to the nine NICUs. Figure 3.1 helps visualize the scope of Section 3.1 to 3.3.

Figure 3.1: The three maps, from left to right, present the scope of Section 3.1, 3.2, and 3.3, respectively. Administrative boundaries: GADM (2012)

3.1 WKZ

3.1.1 Types of care

The neonatology department of the WKZ provides three types of care: Intensive Care (IC), High care (HC), and Medium Care (MC). The focus of this project is on the seriously ill newborns that are admitted to the IC. General hospitals might provide Medium and High Care themselves, but as mentioned before, only nine locations provide IC. The neonatology department has three separated Intensive Care Units of eight beds each. It is important that the newborns are not feeling stressed due to for example high noise levels. A renovation of the units is planned, resulting in more privacy for the families of the children.

The condition of the patients are extensively monitored at the IC. For example heart rate, blood pressure, and oxygen levels are measured. The newborns are placed in an incubator to maintain suitable conditions. This incubator has other equipment attached to it, such as a screen to display measurements, or a machine that offers respiratory support. Birth complications or congenital abnormalities are treated at the IC as well.

Since the IC provides such complex care, expensive equipment and highly educated staff are required. For this reason, the HC has been introduced as a Step Down Unit. Newborns are transferred from the IC to the HC if they do no longer require intensive care. A newborn can be readmitted to the IC, if his condition deteriorates. If the HC is fully occupied, it is possible that patients that no longer require intensive care are still occupying a bed in the IC.

3.1.2 Admission

We distinguish three different origins of a request for a NICU bed. If a patient that requires intensive care is born within the birth centre of the WKZ, the obstetric department informs the neonatology department. The neonatology department is kept up to date on the status of the admitted pregnant women and whether their children might require a bed in the future. A request can also come from a general hospital of the WKZ's own region, or from another NICU.

When the neonatology department receives a request, the coordinating medical specialist and coordinating nurse discuss whether it is feasible to admit this new patient (Oude Weernink (2018)). The current available workforce, the acuity of the already admitted patients, and the origin of the request are taken into account. The patient is assigned a unit and a bed if it is decided feasible to admit him.

Since the IC is often working close to full capacity, the department developed a prioritization scheme for admitting new patients. This way there is less room for discussions when a decision must be made under time pressure. Broadly speaking, the following prioritization is made:

1. Own population (born within the WKZ)
2. Own region
3. Outside of region, but requires specialist care
4. Outside of region, no specialist care

In case of a multiple birth, all children are admitted to the same NICU. Around 55% of the multiple births and 7% of the single births were prematurely born (<37 weeks) in the Netherlands in 2017 (Perined (2019)).

3.1.3 Rejection

When the department deems it infeasible to admit a new patient, taking the previously discussed aspects in Section 3.1.2 into account, the request is denied. Depending on the origin of the denied request, the department must take further action. If the request came from their region, the department is responsible for finding a new place to admit this patient at another NICU. It is possible that a pregnant women whose child will most likely require intensive care is preemptively transferred to another birth centre with a free bed.

Because a request for a bed is communicated by phone, rejections are not automatically registered at the WKZ. Since a couple of years, the department started registering all rejections for IC-beds, including a reason for this rejection.

3.1.4 Discharge

When the medical staff deems the condition of the patient as sufficiently stable and not requiring intensive or high care, the patient can be discharged. This means that usually the patient will be admitted to a local general hospital. However, if that general hospital is unable to admit this patient, he is kept admitted to the birth centre of the NICU. Depending on the organizational structure of the birth centre regarding types of care, the patient now (unnecessarily) occupies a IC, HC, or MC bed. This congests the system and may result in not being to able to admit a new NIC patient.

Furthermore, discharging a patient is not straightforward. Before a patient can be discharged, transport must be prepared for. However, ambulances give priority to emergency calls and may therefore not be available for providing transport at that time. The patient's condition and a summary of his treatment is recorded in a medical correspondence letter which is sent with the patient to the general hospital. In addition, the parents must be present and are heavily involved in this whole process.

3.1.5 Transportation

An ambulance is used for transportation of a child. In addition to the ambulance driver, a nurse and a doctor accompany the newborn in case of a new admission. This turns the ambulance into a mobile NIC. As mentioned before, a NICU is responsible for the transportation of all patients from their own region.

The process of transporting a patient can be divided into multiple smaller steps. In case a patient is admitted from the own region:

1. A doctor and nurse prepare for transportation
2. Drive to the general hospital
3. Move the patient into the ambulance
4. Return to the NICU
5. Admit the patient to the NICU

And, in case a request from the own region is denied:

1. A doctor and nurse prepare for transportation
2. Drive to the general hospital
3. Move the patient into the ambulance
4. Drive to the new NICU
5. Admit the patient to the new NICU
6. Return

When intensive care is no longer required, the patient is discharged and in most cases transferred back to his general hospital. This time, the patient is transported by just the ambulance crew.

3.2 Neonatal Intensive Care

3.2.1 Locations

Since a merge of two hospitals in Amsterdam in June 2018, there are nine *hospital organizations* that provide NIC in the Netherlands. Although Amsterdam UMC currently maintains two locations, these will be merged in the near future. In addition, the two location in Amsterdam share the same catchment area. Therefore, we merge these two locations under the name Amsterdam UMC, using the location of AMC for calculating travel times.

All NICUs except two are located in academic hospitals. Each NICU has their own catchment region, which is composed of a certain number of general hospitals. In Section 3.3.3 we discuss the current assignment of general hospitals to NICUs.

Figure 3.2: The locations of the ten NICUs in the Netherlands. Amsterdam VUmc and AMC are merged in the remaining part of this thesis. Administrative boundaries: GADM (2012)

Figure 3.2 shows the geographical locations of the ten NICUs and how they are dispersed over the Netherlands. Note that in the remaining part of this thesis we merge the two locations in Amsterdam. Table 3.1 contains characteristics of the NICUs, such as the city in which the hospital is located and the number of IC-admissions at each hospital in 2015. The recent operational capacity is included as well. We included the latest response we have received from our capacity survey. Note that capacity may change from week to week due to for example staff shortages.

Hospital	City	IC admissions (2015)[1]	Operational capacity
Amsterdam UMC	Amsterdam	802	28
UMC Groningen	Groningen	525	16
Leiden UMC	Leiden	533	17
Maastricht UMC+	Maastricht	270	13
Radboud UMC	Nijmegen	404	12
Erasmus MC	Rotterdam	510	25
WKZ	Utrecht	555	20
Maxima MC	Veldhoven	252	15
Isala	Zwolle	280	17

Table 3.1: The NICUs and their location, number of IC admissions, and capacity.
[1]Perined (2016)

3.2.2 Specialist care

Prematurely born children have an increased risk of having (major) complications. Since these complications can be complex to treat, not all NICUs have the expertise and equipment to deal with all scenarios. The following four types of specialist care are distinguished:

- Pediatric surgery (PS)
- Pediatric neurosurgery (PNS)
- Pediatric cardiac surgery (PCS)
- Extracorporeal membrane oxygenation (ECMO) [1]

Table 3.2 shows the specialist care each NICU provides. The two NICUs located in a non-academic hospital provide no specialist care.

NICU	PS	PNS	PCS	ECMO
Amsterdam UMC	X	X		
UMC Groningen	X	X	X	
Leiden UMC		X	X	
Maastricht UMC+	X	X		
Radboud UMC	X	X		X
Erasmus MC	X	X		X
WKZ	X	X	X	
Maxima MC				
Isala				

Table 3.2: The NICUs and which specialist care they provide.

[1]ECMO is a method, using a machine, that takes over the functions of the heart and lungs.

3.3 Neonatal Care Network

3.3.1 General hospitals

Primary and secondary birth care are often organized and centered together in local groups called local maternity care consultation and cooperation groups ("Verloskundig Samenwerkingsverbanden") (e.g., see Boesveld et al. (2017)). Since NICUs receive their requests for beds through general hospitals, we use these hospitals to represent NICU demand from the catchment area of the local groups they are collaborating with. We explain how we estimate this demand in Section 3.3.2.

For the situation in 2019, there are a total of 74 hospitals. Figure 3.3 shows all locations in the Neonatal Care Network, which includes all general hospitals in addition to the nine NICUs. Table 3.3 contains a list of all the 74 hospitals.

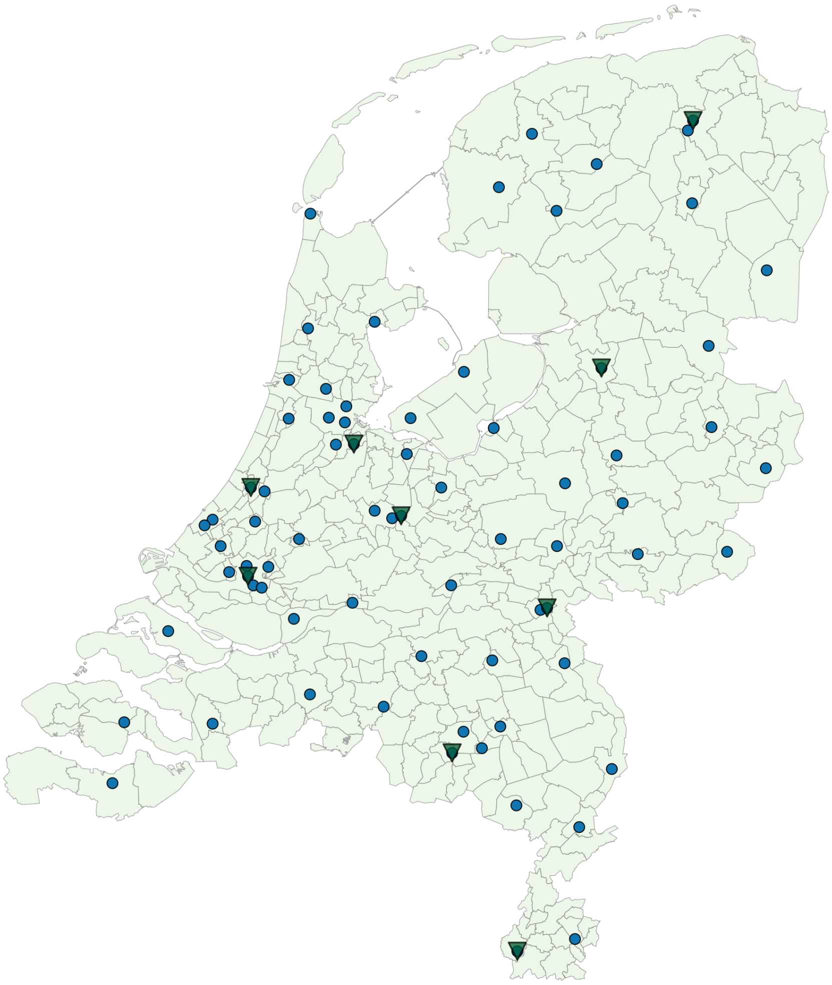

Figure 3.3: The location of all general hospitals and the nine NICUs.

Admiraal De Ruyter Ziekenhuis Goes	Meander Medisch Centrum
Albert Schweitzer Ziekenhuis Dordwijk	Medisch Centrum Leeuwarden
Alrijne Ziekenhuis Leiderdorp	Medisch Spectrum Twente
Amphia Ziekenhuis Breda Langendijk	Noordwest Ziekenhuisgroep Alkmaar
Amsterdam UMC	Noordwest Ziekenhuisgroep Den Helder
Antonius Ziekenhuis Sneek	Onze Lieve Vrouwe Gasthuis locatie Oost
Beatrixziekenhuis	Onze Lieve Vrouwe Gasthuis locatie West
BovenIJ Ziekenhuis	Reinier de Graaf Gasthuis
Bravis Ziekenhuis Bergen op Zoom	Rijnstate
Canisius-Wilhelmina Ziekenhuis	Rode Kruis Ziekenhuis
Catharina Ziekenhuis	Ropcke Zweers Ziekenhuis
Deventer Ziekenhuis	Scheper Emmen
Diakonessenhuis Utrecht	Slingeland Ziekenhuis
Dijklander Ziekenhuis locatie Hoorn	Spaarne Gasthuis locatie Haarlem-Zuid
Elkerliek Ziekenhuis Helmond	St. Anna Ziekenhuis Geldrop
Erasmus MC	St. Antonius Ziekenhuis Utrecht
ETZ Elisabeth	St. Jans Gasthuis
Flevoziekenhuis	Streekziekenhuis Koningin Beatrix
Franciscus Gasthuis	Tergooi locatie Blaricum
Franciscus Vlietland	UMC Groningen
Gelre Ziekenhuizen Apeldoorn	UMC St. Radboud
Gelre Ziekenhuizen Zutphen	Universitair Medisch Centrum Utrecht
Groene Hart Ziekenhuis Gouda	Van Weel-Bethesda Ziekenhuis
Haaglanden Medisch Centrum Westeinde	VieCuri Medisch Centrum Venlo
HagaZiekenhuis Leyweg	Wilhelmina Ziekenhuis Assen
IJsselland Ziekenhuis	Zaans Medisch Centrum
Ikazia Ziekenhuis	Ziekenhuis Amstelland
Isala Zwolle	Ziekenhuis Bernhoven
Jeroen Bosch Ziekenhuis	Ziekenhuis De Gelderse Vallei locatie Ede
LangeLand Ziekenhuis	Ziekenhuis de Tjongerschans
Laurentius Ziekenhuis	Ziekenhuis Nij Smellinghe
Leids Universitair Medisch Centrum	Ziekenhuis Rivierenland
Maasstad Ziekenhuis	Ziekenhuis St. Jansdal Harderwijk
Maastricht UMC+	Ziekenhuis St. Jansdal Lelystad
Maasziekenhuis Pantein	Ziekenhuisgroep Twente Locatie Almelo
Martini Ziekenhuis	ZorgSaam De Honte
Maxima Medisch Centrum Veldhoven	Zuyderland Medisch Centrum Heerlen

Table 3.3: List of included general hospitals

3.3.2 Demand

Perined (Netherlands Perinatal Registry) supplied to us a dataset of the estimated number of births in the catchment area of each previously mentioned general hospital in 2016 and 2017. Using a specific methodology, childbirths at home were allocated to local maternity care consultation and cooperation groups where they likely would have taken place. These numbers may not be published without permission from all hospitals, and therefore will not be shared in this document or elsewhere. Furthermore, we edited the dataset to represent the situation as in 2019, since some hospitals have merged or gone bankrupt since 2016.

Unfortunately we only have two data points (2016, 2017) to work with for each hospital, because we had asked for recent data, in order for it to be applicable to the current situation. In addition to the statistical uncertainty having few data points implies, randomness and policy decisions make it difficult or impossible to accurately predict the number of births of each hospital for a following year. We use the average of 2016 and 2017 as the expected number of births from now on forward.

We would still have to estimate the number of NIC patients using these number of births. Table 3.4 shows the number of births and NICU admissions of the five most recent years that are published by Perined. For calculating a confidence interval of the probability of a pregnancy requiring a NICU admission p, we pool the statistics over those five year. We assume the probability for each pregnancy is the same and independent of others, and there is no trend over the years. Then, using the normal approximation, the confidence interval can be calculated by $\hat{p} \pm \sqrt{\frac{\hat{p}*(1-\hat{p})}{n}}$. With $\hat{p} = 0.2362$ we find a 95% CI of $\left[0.2330, 0.2394\right]$ for p. This interval seems sufficiently small and could be used later during the sensitivity analysis of the final results.

Year	Number of births	Number of NICU admissions	Percentage	Source
2015	169267	4131	2.441%	Perined (2016)
2014	175215	4181	2.386%	Perined (2015)
2013	169884	4100	2.413%	Perined (2014)
2012	176155	4017	2.280%	Perined (2013b)
2011	178607	4099	2.295%	Perined (2013a)

Table 3.4: Number of NICU admissions compared to the total number of births.

There is one more issue with estimating the NICU demand from the dataset. It is likely that the birth centres with a NICU have a larger percentage of newborns requiring NIC than those who have not. Pregnant women whose child possibly needs NIC after birth might be referred to a perinatal center at an earlier stage. The woman then gives birth at that perinatal center, while she actually was referred from a different hospital. Since we do not have the required data to thoroughly investigate this, we will assume that all hospitals and perinatal centers have the same value $\hat{p}$.

3.3.3 Current assignment

Figure 3.4 shows the current assignment of hospitals to the nine NICUs. We refer to Appendix D for the complete list.

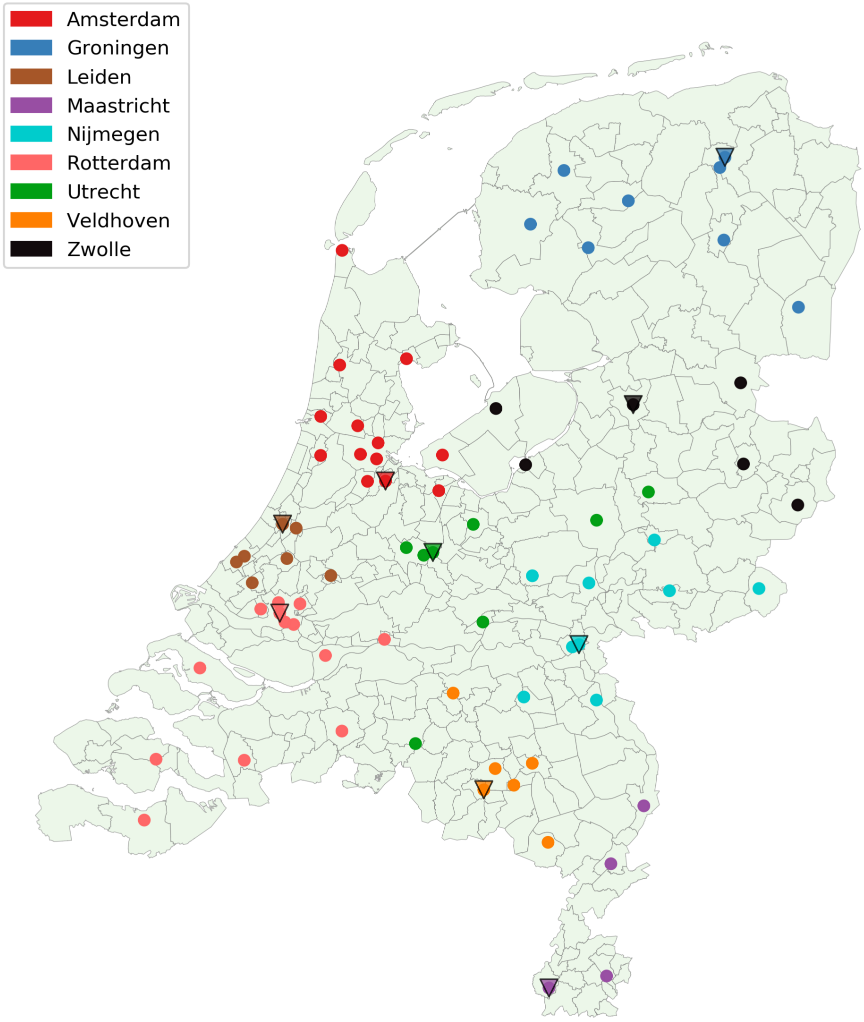

Figure 3.4: The current assignment of general hospitals to NICUs.

Using the assignment and estimated NICU demand of each hospital, we can calculate the total assigned demand to each NICU. Table 3.5 compares our own calculated demand to published data from Perined of 2015.

Perined assigned NICU admissions to a origin region. This is based on the the region of the referring hospital (including transfers), or otherwise the postal code of the woman. However, it is not mentioned or explained how the origin region of a NICU is defined. Therefore it is unclear whether they used, for example, a geographical criterion or the actual assignment that we used. The most severe differences between our own numbers and those of Perined appear to be between geographically 'contested' areas, such as Leiden and Rotterdam, Zwolle and Groningen, and Veldhoven and Maastricht.

Figure 3.5 shows, based on Perined's data, that the ratio between NICU demand and the number of births is not equal in each NICU region. However, for our own calculations, previously discussed in Section 3.3.2, we assumed this ratio *was* equal.

NICU region	**Number of births**	**Normalized number of births (Perined [1])**	**Estimated NICU demand**	**Normalized NICU demand own region (Perined [2])**
Amsterdam	32924	34015	778	781
Groningen	14534	16035	343	491
Leiden	18635	15239	440	526
Maastricht	7182	7792	170	263
Nijmegen	16815	14301	397	377
Rotterdam	28429	30163	671	519
Utrecht	23563	24115	557	488
Veldhoven	11589	13209	274	241
Zwolle	14164	12963	335	279
Total	167833	167833	3964	3964

Table 3.5: Our own calculated number of births and NICU demand for each NICU region, compared to normalized Perined data from 2015. [1] Perined (2016) table 2.1.3; [2] Perined (2016) table 10.5

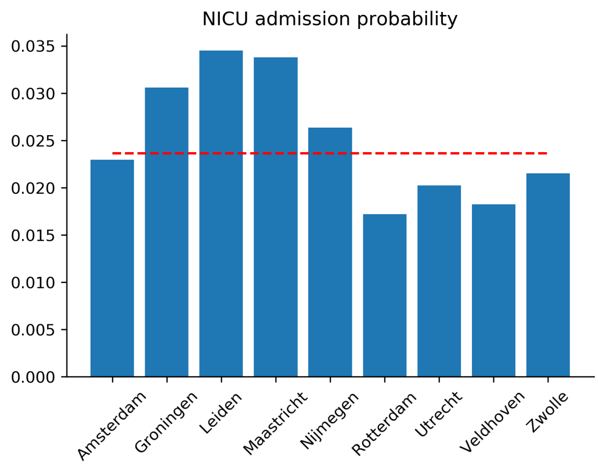

Figure 3.5: The NICU admission probability of each birth allocated to a NICU region, based on normalized Perined data of Table 3.4. The dashed red line represents the average $\hat{p} = 0.2362$

3.3.4 Travel time

We obtained the distance and travel time between the NICUs and the general hospitals from the free-to-use OpenStreetMap (OSM). This was done in R using the package 'osrm'. The travel time is solely based on the speed limit of the road segments. This means that traffic delays are not taken into account. To validate the distances and travel times, we randomly selected 20 routes to manually check using Google Maps.

Distances were the same or at least comparable. Travel times were always either the same or longer in Google maps, compared to OSM. In all cases that the travel time was longer, it was because of extra travel time caused by traffic congestion, as was indicated on the map.

3.3.5 Spatial accessibility

In Chapter 2 we discussed three main developments in Floating Catchment Area (FCA) metrics. Now, in this section we apply these methods and investigate how accessibility to Neonatal Intensive Care is related to travel time from the parents' house.

To recap Section 2.1, Two-Step Floating Catchment Area (2SFCA) measures accessibility to health care services taking *all* service sites that can be reached within a defined time threshold into account. Enhanced Two-Step Floating Catchment Area (E2SFCA) defines multiple travel time zones, each given decreasing weights. And finally, Three-Step Floating Catchment Area (3SFCA) introduces competition between supply locations. If more than one supply location can be reached from a certain demand location, a fraction of the demand is assigned to each supply location. This fraction is proportional to the weight given to this service site and the sum of the weights of all accessible service sites. A set of weights for travel time zones of E2SFCA and 3SFCA can be obtained from a Gaussian function, as suggested by Shi et al. (2012).

Spatial data of municipalities in 2018 are obtained from Kadaster/CBS (2018). The most recent published data on the number of births per municipality is of 2017 (CBS (2019)). Since some municipalities have merged at the start of 2018, we have incorporated these changes into the data set as well. Travel times are obtained using the geographical centroid of a municipality.

The average NICU length of stay was approximately 11.63 days in 2015 (Perined (2016)). Given there are 163 operational beds in the network, there is capacity for a total of 5115 admissions. In 2015, a total of 169 267 children were born (Perined (2016)). This means that, if accessibility were equal in all municipalities, one would expect an accessibility of one NICU admission for every 33 births.

We use E2SFCA with travel time zones {30 minutes, 60 minutes, 90 minutes} and weights {1, 0.6, 0.13}. Figure 3.6 shows the resulting spatial accessibility to NIC. We observe that some areas severely lack NIC coverage, such as Enschede and Zeeland. In addition, as expected, it seems that the rural areas have more trouble accessing NIC than the densely populated areas. If you are living in the Randstad, you are more likely to find an alternative NICU close to home, than if you were living outside of the Randstad.

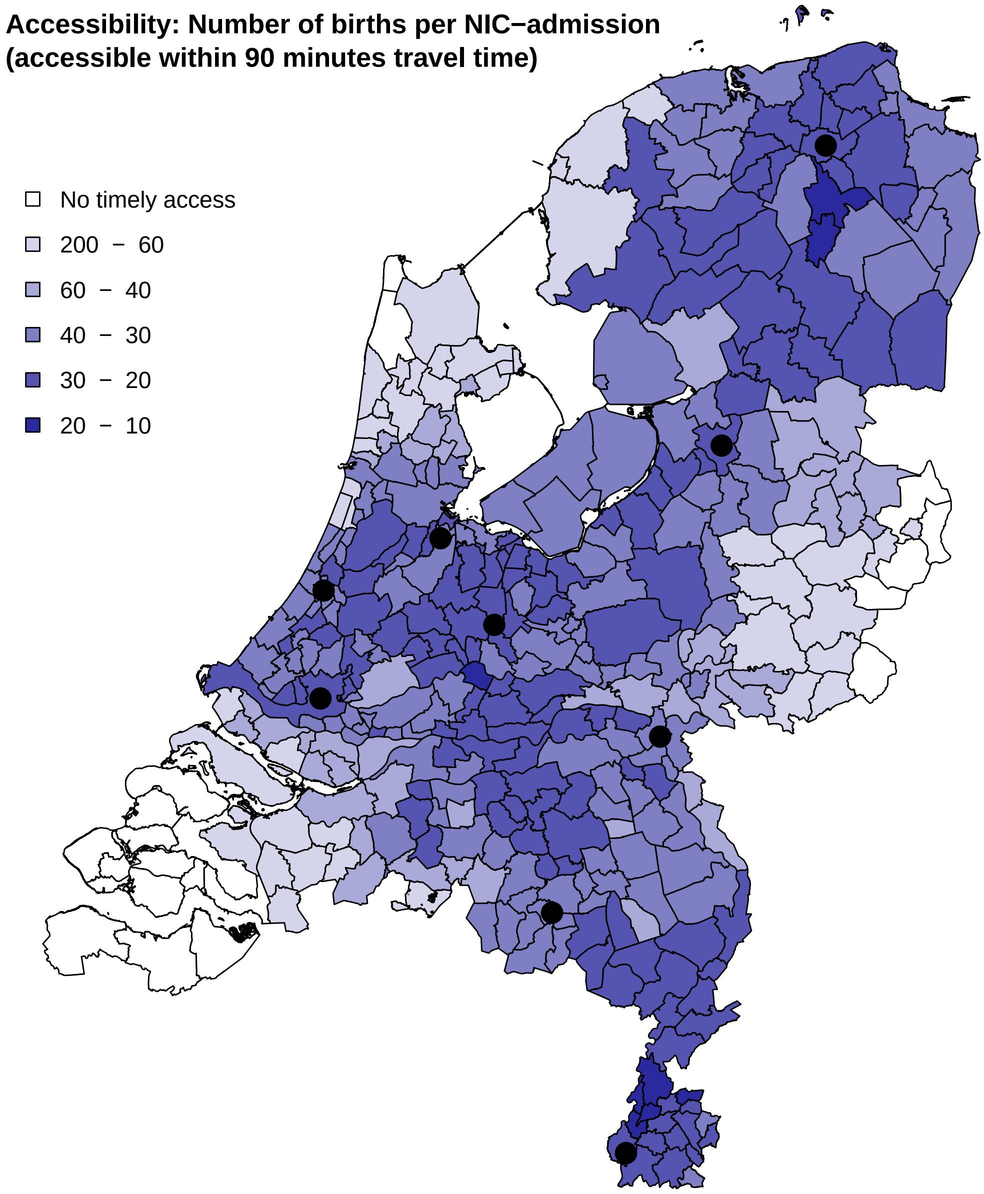

Figure 3.6: Spatial accessibility to NIC measured using E2SFCA.

3.4 Conclusion

The third research question *"What is the current situation in the Neonatal Care Network?"* is answered in this chapter.

The neonatology department of the WKZ provides three types of care: Intensive Care (IC), High care (HC), and Medium Care (MC). The neonatology department has three separated Intensive Care Units of eight beds each. The condition of the patients are extensively monitored at the IC. For example heart rate, blood pressure, and oxygen levels are measured. The newborns are placed in an incubator to maintain suitable conditions. A renovation of the units is planned, resulting in more privacy for the families of the children.

Since the IC provides such complex care, expensive equipment and highly educated staff are required. For this reason, the HC has been introduced as a Step Down Unit. Newborns are transferred from the IC to the HC if they do no longer require intensive care. A newborn can be readmitted to the IC, if his condition deteriorates. If the HC is fully occupied, it is possible that patients that no longer require intensive care are still occupying a bed in the IC.

We described four different processes at the WKZ: admission, rejection, discharge, and transportation. The WKZ prioritizes admissions based on the origin of the request. If a NICU is fully occupied and a new request for a bed comes from a NICUs own region, then that NICU department is responsible for finding a new place to admit this patient at another NICU. Transporting a child is time consuming and expensive. It is possible that a pregnant women whose child will most likely require intensive care is preemptively transferred to another birth centre with a free bed. When the medical staff deems the condition of the patient as sufficiently stable and not requiring intensive or high care, the patient can be discharged. This means that usually the patient will be admitted to his general hospital. However, if that general hospital is unable to admit this patient, he is kept admitted to the birth centre of the NICU. This congests the system and may result in not being to able to admit a new NIC patient.

Since a merge of two hospitals in Amsterdam in June 2018, there are nine *hospital organizations* that provide NIC in the Netherlands. Although Amsterdam UMC currently maintains two locations, these will be merged in the near future. In addition, the two location in Amsterdam share the same catchment area. All NICUs except two are located in academic hospitals.

Each NICU has its own catchment region with a certain number of general hospitals. Using data from Perined on the number of births at each hospital, we estimated the expected NICU demand of each hospital. In addition we gathered all travel times between hospital and NICUs. Furthermore, we obtained the recent operational capacity of each NICU. And finally, we obtained and visualized the current assigned of hospitals to NICUs(3.4.

We used E2SFCA to measure spatial accessibility to NIC. We observed that some areas severely lack NIC coverage, such as Enschede and Zeeland. In addition, it seems that the rural areas have more trouble accessing NIC than the densely populated areas.

Chapter 4

Model formulation

In this chapter we will formulate deterministic and stochastic models to allocate regional hospitals to NICUs. Finding a close-to-optimal solution for the stochastic models will be done using heuristics.

Due to variability in the arrival process and length of stay, not all patients can be admitted to their own NICU. This means that these patient must be admitted to some other NICU. In some rare cases, patients are admitted abroad (fewer than 1% of all patients). Sometimes parents opt to let their child be admitted abroad, even though there is a bed available somewhere in the Netherlands, since the travel time would be shorter. Either way, a transfer decision is always made together with the parents. In the models we formulate in this chapter, we only include Dutch NICUs and make decisions solely based on travel time.

Including stochasticity shifts the focus from individual NICUs towards network behaviour. A network is harder to analyse and it is more difficult to find an optimal solution. A key characteristic of this network is that for rejected patients a new must NICU must be found. Since the number of rejected patient increases when the offered load to the NICU increases, a trade-off can be made in assigning demand to NICUs. For example, for a certain hospital the extra travel time incurred for not choosing the closest, but a less busier NICU, might outweigh the additional transportation time spent on transferring patients to another NICU.

We use the following definitions for travel time, which were described in Section 3.1.5:

- Admission of a patient from the own region:
  NICU → hospital → NICU
- Rejection of a patient from the own region (transfer):
  NICU A → hospital → NICU B → NICU A

We start in Section 4.1 with a deterministic model, in which we assume all allocated demand can be satisfied by a NICU. In Section 4.2 we include stochasticity by formulating the problem as a network of queues. The goal is calculate the number of transfers of patients to other NICUs. In Section 4.3 we use a Continuous-Time Markov Chain (CTCM) to analyze this network. However, this method is only able to solve smaller case studies. So therefore, we use Discrete Event Simulation (DES) to analyze this network as well, in Section 4.4. In Section 4.5 we design an optimization heuristic to find a close-to-optimal solution for the stochastic formulation.

4.1 Deterministic formulation: Integer linear programming

In this Section we propose a deterministic model to assign hospitals to NICUs. In these models we assume there are no transfers of patients between NICUs.

Assigning hospitals to NICUs can be interpreted as a General Assignment Problem (GAP). As mentioned before in Section 2.2, the GAP can be formulated as assigning tasks to agents, such that

- each task is assigned to exactly one agent;
- the required resources of the assigned tasks to an agent do not exceed the agent's capacity;
- and the total cost (profit) of all assignments is minimized (maximized).

With regard to the capacity constraint, we have three options:

1. Uncapacitated, which is equivalent to the closest assignment
2. Capacitated using the maximum throughput per year
3. Capacitated using a maximum allowed rejection probability

In Section 4.1.1 we give the uncapacitated formulation. Afterwards, in Section 4.1.2 we discuss how we can add constraints to limit the assigned demand.

4.1.1 Uncapacitated

The uncapacitated assignment can be interpreted as a best case scenario. We have the following formulation:

Sets

N	Set of NICUs $n, m \in \{1, 2, ..., \vert N\vert\}$
H	Set of hospitals $h \in \{1, 2, ..., \vert H\vert\}$

Parameters

d_h	Average daily demand of hospital h
$t^{h\_to\_n}_{(h,n)}$	Travel time between hospital h to NICU n
$t^{n\_to\_h}_{(n,h)}$	Travel time between NICU n to hospital h
$t^{n\_to\_m}_{(n,m)}$	Travel time between NICU n to NICU m

Decision variables

$$X_{(h,n)} = \begin{cases} 1 & \text{if hospital } h \text{ sends its patients to NICU } n \\ 0 & \text{otherwise} \end{cases}$$

Objective function

$$minimize \quad \sum_{h \in H} \sum_{n \in N} ((t^{n\_to\_h}_{(n,h)} + t^{h\_to\_n}_{(h,n)})\, X_{(h,n)}\, d_h)$$

Subject to

$$\sum_{n \in N} X_{(h,n)} \quad = \quad 1 \qquad \forall h \tag{4.1}$$

$$X_{(h,n)} \quad = \quad \{0, 1\} \qquad \forall (n, h) \tag{4.2}$$

The objective function minimizes the total travel time, according to our own chosen definition of travel time. Formulating a different definition of travel time may be possible, such as only counting the travel time spent on transfers. Constraint (4.1) and (4.2) ensure each hospital is assigned to exactly one NICU.

4.1.2 Capacitated

We can restrict the blocking percentage of a NICU in a linear programming formulation. Pehlivan et al. (2014) formulated a constraint using the Erlang loss function $B(c, a) = \frac{(a)^c / C!}{\sum_{k=0}^{c} (a)^k / k!}$, with offered load $a = \frac{\lambda}{\mu}$. However, this constraint is nonlinear in this form. Therefore Pehlivan et al. (2014) suggested several methods to linearize this Erlang loss function. One of those methods is what the authors refer to as the maximum admissible offered load $\bar{a}(c, \alpha)$, such that $B(c, \bar{a}(c, \alpha)) = \alpha$. The constraint with the Erlang loss function would then have to be replaced by a constraint requiring that the assigned demand is fewer than $a(c, \alpha))$. Since $a(c, \alpha))$ can be precomputed and can be introduced as a parameter, this constraint is now linear.

We can calculate the maximum admissible offered load by iteration. Using the Erlang loss function, we approximate the arrival rate that corresponds to each combination of capacity and rejection probability. For some capacity, we iterate over different arrival rates until its rejection rate is sufficiently close to the target rejection rate.

Additional parameters

mu	service rate
c_n	Operational capacity of NICU n
α	Maximum allowed blocking percentage
$a(c_n, \alpha))$	Maximum admissible load of NICU n with c_n beds

Additional constraints

Either

$$X_{(h,n)}\, d_h \leq c_n \mu_n \qquad \forall n \tag{4.3}$$

or

$$X_{(h,n)}\, d_h \leq a(c_n, \alpha)) \qquad \forall n \tag{4.4}$$

Only one of those two constraints should be added to the formulation of the previous Section 4.1.1. Constraint (4.3) restricts the assigned demand by the maximum throughput per year. Constraint (4.4) limits the demand that can be assigned to a NICU, by requiring that the assigned demand is equal or smaller than the maximum admissible load of that NICU. This maximum admissible load is a substitute for a maximum blocking percentage.

4.2 Stochastic formulation: A network of queues

As mentioned before, we model the NICUs as a network of queues. These queues operate in parallel, as we assume patients only receive treatment at one NICU before leaving the network. Each NICU has their own arrival process for patients of their own region. In case such a patient must be rejected because the NICU is fully occupied, the patient is admitted to another NICU. This new NICU can be found using a predefined prioritization matrix.

Define N as the set of NICUs. We model each individual NICU as an $M|M|c|c$ queue. This queue has a fixed capacity of c beds and has no waiting room or buffer. The arrivals occur according to a Poisson process. We assume the hospital's demand is Poisson distributed, and therefore the arrival rate λ_n is a result of an assignment of regional hospitals to NICU n. The Poisson assumption for NICU arrivals has been used and tested before in evaluating the required number of NICU beds at the WKZ (Oude Weernink (2018)), and is commonly used in context of emergency departments or hospital wards (e.g. in de Bruin et al. (2010)). In addition we assume all patients have an identical average length of stay (service rate $\mu = \frac{1}{LOS}$), independent of treatment location.

By working together in a network, NICUs are able to admit more patients in total, but at the cost of rejecting more patients from their own region. The total number of patients that can be admitted increase because of resource pooling. However, these additional patients occupy beds that previously would have been used for patients from a NICU's own region, increasing the number of transfers.

To illustrate this, we compare the performance of two identical $M|M|c|c$ queues working separately and two working collaboratively. We model each queue with $c = 20$ beds (and no waiting room), a arrival rate of $\lambda = 500$ per year, and a service rate of $\mu = \frac{365}{12}$ per bed per year. Working separately, each queue has an average bed occupancy of 76% ($\frac{\sum_{k=0}^{c+1} k\pi_k}{c}$) and is unable to admit 7.35% of its arrivals. Working together, each queue has a slightly higher average bed occupancy of 79%. Now, each queue rejects 11.2% of their own patients, but the majority of those are admitted to the other NICU. Only 3.34% of all patients leave the network without being admitted somewhere.

We model transfer between NICUs by constructing a prioritization matrix that is used for finding a bed for a patient. This matrix indicates in which order NICUs are asked to admit this specific patient. First, the 1st-priority NICU is asked. If that NICU is unable to admit this patient, then the 2nd-priority NICU is asked, etc.. If all beds in the network are occupied, the patient leaves the network without being admitted somewhere.

Prioritization can be based on a patient's initially assigned NICU (NICU-to-NICU prioritization), or on a patient's regional hospital (hospital-to-NICU prioritization). The second method is preferred, but requires differentiating between arrivals of all assigned regional hospitals to a NICU region.

We prioritize NICUs solely based on travel time (round trip). While this rule might not be optimal with regards to total travel time in all cases, it seems intuitive that it provides a good foundation. It is of course also possible to manually modify this matrix to fit current behavior or choices (e.g., collaborations, a preferred order if difference in travel time is small, etc.). Figure 4.1 shows the NICU-to-NICU prioritization matrix for our case study.

In Sections 4.3 and 4.4 we describe two different methods to calculate and analyze transfers within a network of queues.

NICU region	**Priority 1**	**Priority 2**	**Priority 3**	**Priority 4**	**Priority 5**	**Priority 6**	**Priority 7**	**Priority 8**	**Priority 9**
Amsterdam region	Amsterdam	Utrecht	Leiden	Rotterdam	Zwolle	Veldhoven	Nijmegen	Groningen	Maastricht
Groningen region	Groningen	Zwolle	Amsterdam	Utrecht	Nijmegen	Leiden	Rotterdam	Veldhoven	Maastricht
Leiden region	Leiden	Amsterdam	Rotterdam	Utrecht	Zwolle	Veldhoven	Nijmegen	Groningen	Maastricht
Maastricht region	Maastricht	Veldhoven	Nijmegen	Utrecht	Amsterdam	Rotterdam	Zwolle	Leiden	Groningen
Nijmegen region	Nijmegen	Veldhoven	Utrecht	Zwolle	Amsterdam	Rotterdam	Maastricht	Leiden	Groningen
Rotterdam region	Rotterdam	Leiden	Utrecht	Amsterdam	Veldhoven	Nijmegen	Zwolle	Maastricht	Groningen
Utrecht region	Utrecht	Amsterdam	Rotterdam	Leiden	Zwolle	Veldhoven	Nijmegen	Maastricht	Groningen
Veldhoven region	Veldhoven	Maastricht	Utrecht	Nijmegen	Amsterdam	Rotterdam	Leiden	Zwolle	Groningen
Zwolle region	Zwolle	Utrecht	Groningen	Amsterdam	Nijmegen	Leiden	Rotterdam	Veldhoven	Maastricht

Figure 4.1: The NICU-to-NICU prioritization matrix.

4.3 Analyzing the stochastic formulation: Continuous-Time Markov Chain

4.3.1 Formulation

A Continuous-Time Markov Chain ($CTMC$) is a stochastic model that has many applications and is commonly used to analyze queues. Our formulation is based on the classic 'birth-death process', which means at the time of an event the state variable either increases or decreases by one. The time between events is stochastic as well.

We define w_n as the number of patients admitted to NICU n. To describe the whole system at any point in time, we need to know how many patients are being admitted to each NICU. Therefore we define state $s = (w_1, w_2, ..., w_{|N|})$. The number of patients that can be admitted to one moment is constrained by the number of beds c_n of NICU n. Thus we have a state space $S = \{(w_1, w_2, ..., w_{|n|}) | w_n >= 0,\ w_n \leq c_n \quad \forall n \in N\}$.

We define a set of blocking states B as a subset of S. B includes all states that have at least one NICU fully occupied. When a patient arrives at a fully occupied NICU, a new NICU must be found. Since we know which other NICUs are fully occupied at the moment, we can find the highest priority NICU that is able to admit this patient by iterating through the prioritization matrix. Because the state only contains information about NICUs and not hospitals, we are only able to use NICU-to-NICU prioritization. Hospital-to-NICU prioritization would require to keep track of how many patients of hospital h are being admitted to each NICU. This would lead to a much larger state space. We refer to elements of this NICU-to-NICU prioritization matrix as P_{np}, with indices for NICU n and priority $p \in \{1, .., |N|\}$.

We are interested in the fraction of time spent in each (blocking) state and therefore want to find the steady-state distribution π. Let us define Q as the transition rate matrix. In Q, we denote the rate in which the system moves from state $s \in S$ to a new state $s' \in S$, and define each element as $q_{ss'}$. And similar to Andersen et al. (2017), we define $s' = (..., w_n + 1, ...)$ as an admission to NICU n and $s' = (..., w_n - 1, ...)$ as a discharge.

In Q, the transition rates between states have the following values:

$$q_{ss'} = \begin{cases} \lambda_n & \text{if } s' = (..., w_n + 1, ...) \text{ and } s \notin B \quad \forall n \in N & \text{(4.5a)} \\ \lambda_n + \sum\limits_{m \in M_n^s} \lambda_m & \text{if } s' = (..., w_n + 1, ...) \text{ and } s \in B \quad \forall n \in N & \text{(4.5b)} \\ \mu_n w_n & \text{if } s' = (..., w_n - 1, ...) \quad \forall n \in N & \text{(4.5c)} \\ 0 & \text{otherwise} & \text{(4.5d)} \end{cases}$$

Equation 4.5a indicates an admission of a patient, when there are no fully occupied NICUs in state s. When at least one NICU is fully occupied, the arrival rate to the other

NICUs increase according to the prioritization matrix. Therefore, in Equation 4.1b, we introduce M_n^s to denote the set of NICUs that refer their patients to NICU n when the system is in state s. M_n^s is determined by considering all fully occupied NICUs in state s and finding their highest priority available stand-in NICU using the prioritization matrix. Equation 4.5c indicates a discharge of a patient. Furthermore, all other values in Q are 0 (Equation 4.5d) since their corresponding states are not connected.

After filling Q with the values of Equation 4.5, all diagonal values in Q are set equal to the negative sum of its row, meaning $q_{ss} = -\sum\limits_{s' \in S} q_{ss'} \quad \forall s \in S$. Figure 4.2 presents the algorithm to fill Q.

Now we can find the steady-state distribution π by solving $\pi Q = 0$. Since we have the additional constraint $\sum\limits_{s \in S} \pi_s = 1$, we replace one column of Q with 1's. However, instead of solving $\pi Q = 0$ we can approximate the steady state distribution using the power method. As it converges towards the exact solution, some accuracy might be lost. Using the power method reduces the time spent on finding π enormously.

```
foreach state s ∈ S do
    row_sum = 0 ;
    foreach NICU n ∈ N do
        if w_n < c_n then
            define the new state s' as (..., w_n + 1, ...) ;
            q_ss' += λ_n ;
            row_sum += λ_n ;
        else
            for priority p = 2 to |N| do
                m ← P_np ;
                if w_m < c_m then
                    define the new state s' as (..., w_m + 1, ...) ;
                    q_ss' += λ_n ;
                    row_sum += λ_n ;
                    break;
                end
            end
        end
        if w_n > 0 then
            define the new state s' as (..., w_n − 1, ...) ;
            q_ss' = μ_n w_n ;
            row_sum += μ_n ∗ w_n ;
        end
    end
    q_ss = row_sum ;
end
```

Figure 4.2: Algorithm for filling transition matrix Q

4.3.2 Calculating the number of transfers and total travel time

After obtaining the steady state distribution, we know the fraction of time that is spent over the long run in each state. Using the prioritization matrix, we also know for each state which NICU a patient from a certain region would be admitted to.

The PASTA property of Poisson processes states that on average both an outside random observer and an arriving customer have the same probability of seeing the system in a certain state (Wolff (1982)). This means that the fraction of time that is spent in some state is also the fraction of total customers that find the system in that same state on arrival.

We can use the PASTA property to derive an admission table A ($|n| \times |N| + 1$) from the steady state distribution π. This admission table indicates in each row where patients of a NICU region end up being admitted and includes an additional column indicating the number of patients that left the network without receiving treatment. Figure 4.3 presents an algorithm for deriving A from π. This algorithm iterates through all blocking states $b \in B$ and allocates the value π_b to the NICU where an arriving patient of NICU $n \in N$ would be admitted to. Afterwards, each value in row n of A is multiplied with the arrival rate λ_n.

Define N as the set of all NICUs ;
foreach *blocking state* $b \in B$ **do**
    define M as the set of fully occupied NICUs in this state b ;
    foreach *NICU* $m \in M$ **do**
        find the highest priority NICU h of NICU m that is not fully occupied ;
        if *all NICUs are occupied* **then**
            $A_{m,|N|+1} += \pi_b$;
        else
            $A_{mh} += \pi_b$;
        end
    end
end
foreach *NICU* $n \in N$ **do**
    set A_{nn} equal to $1 - \sum_{m=1}^{|N|+1} A_{nm}$;
    multiply each value in row n of A with the arrival rate λ_n ;
end

Figure 4.3: Algorithm for deriving admission matrix A from the steady state distribution π.

For calculating the total expected travel time we introduce an additional matrix T with the same dimensions as A. In T we calculate the average travel time of a scenario that a patient from NICU region n is admitted to NICU m. In addition, we define a travel time penalty z for patients who are admitted abroad[1]. We find the total travel time TTT through element-wise multiplication of A and T, $TTT = (\sum_{n \in N} \sum_{m \in N} A_{nm} * T_{nm}) + \sum_{n \in N} A_{n,|N|+1} * T_{n,|N|+1}$.

4.3.3 Reducing the state space

During the search for an optimal solution many other solutions are evaluated as well. For each solution, a new admission table must constructed to calculate its performance. Therefore, it is important that the CTMC is solved time efficiently. In this section we discuss several steps we have taken to reduce the state space and increase performance of our algorithm.

The primary indicator of performance and feasibility is the total number of states, which can be calculated using $|S| = \prod_{n \in N} (c_n + 1)$. From this follows that the number of states increase exponentially when the number of NICUs increases. Figure 4.1 shows how the total number of states increases for our case study with up to nine NICUs. For reference, for our computer, a CTMC with more than 2 million states leads to memory issues. Therefore it is crucial that the total number of states is reduced as much as possible.

Number of NICUs	Number of states
1	29
2	493
3	8,381
4	117,334
5	1,760,010
6	45,760,260
7	1,098,246,240
8	14,277,201,120
9	228,435,217,920

Table 4.1: The size of the state space S when including up to some number of NICUs

[1]Note that the total number of patients that are admitted abroad is a function of the total capacity of the network, and is independent of the assignment of hospitals to NICUs

Although there are $c_n + 1$ possible number of admissions for each NICU, not all of those are as relevant. Unless the NICU is severely underloaded with respect to its capacity, little to no time is spent in states with lower number of admissions. Andersen et al. (2017) uses the steady state distribution of a single queue to find a lower bound for the number of admissions. Since a queue in our network will always be at least as busy as if it were to operate on its own, we can use this lower bound to reduce the total state space S. Figure 4.4 shows a shift in probability mass of the steady state distribution for a NICU in the isolated scenario compared to a network scenario (using parameters of Section 4.2).

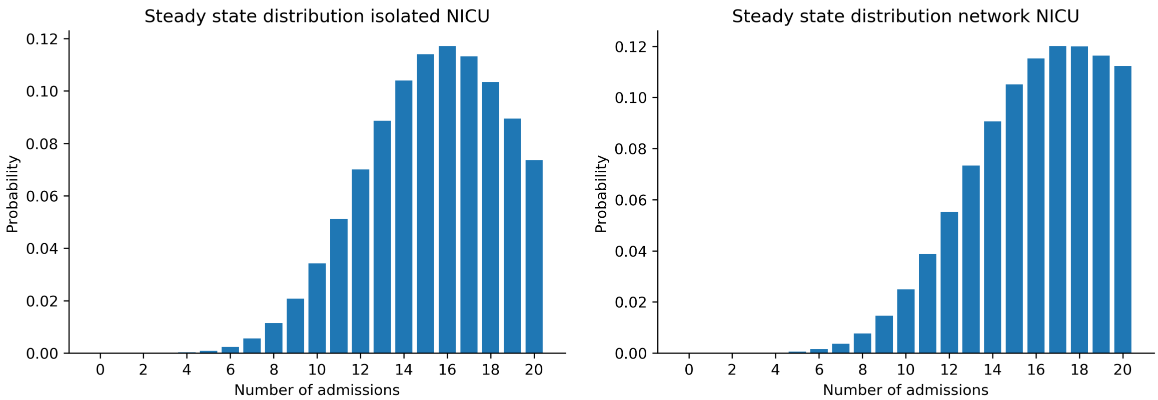

Figure 4.4: The steady state distribution of an isolated NICU (left) and a network NICU (right).

Using a fixed number for the lower bounds is not preferred, because its effectiveness depends on the load on the NICUs, which is determined by the assigned demand. Therefore, we cut off states from the left side of the steady state distribution until we have removed up to a certain total probability mass. For example, removing up to 5% probability mass of the isolated NICU in Figure 4.4 means we can set the LB of admissions to 9. Setting the LB to 10 exceeds the threshold of probability mass that we are allowed to remove, because the removed probability mass would increase from 4.1% (LB = 9) to 7.6% (LB = 10). Figure 4.5 shows that the rejection probability remains approximately equal for a lower bound between 0 and 10, but starts increasing more rapidly afterwards.

It is key to note that increasing the lower bound for the number of admission too aggressively will result in overestimating the number of transfers. This makes comparing different assignments of hospitals difficult, because the solutions might use different lower bounds for admissions. As the total probability mass to be cut off is set higher, states with larger probability mass will be encountered. This makes whether or not the threshold is reached for cutting off an additional state very impactful. For example, setting the total probability mass to be cut off up to 40% might lead to one solution having cut off 39%, and the other just 33%. For a more conservative percentage, such as 5%, the impact of setting the lower bound slightly higher or lower will be much smaller.

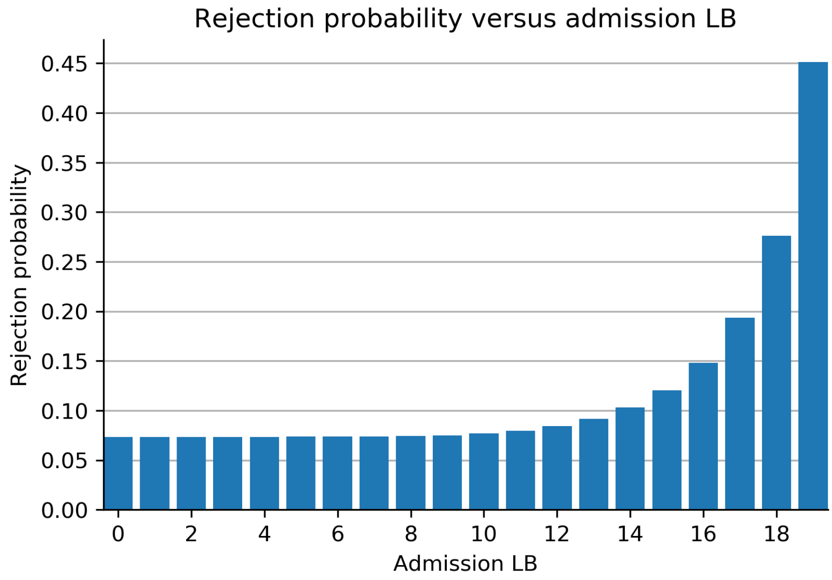

Figure 4.5: The rejection probabilities of $M|M|c|c$ queues with certain lower bounds for the number of admissions.

4.3.4 Model extension: including multiple births

Admissions of multiple births pose a significant challenge for the NICUs. It is highly preferred all siblings are admitted to the same NICU. This means that instead of one bed, multiple beds must be available. For example, if only one bed is available and a twin has to be admitted, the department has to either temporarily increase capacity, or find another NICU to admit the twins. In the models it is not possible to temporarily increase capacity.

Modeling only single arrivals and not accounting for multiple births leads to an under-estimation of the blocking probability. For this reason we propose an extension of our Continuous-Time Markov Chain model to analyze $M|M|c|c$ queues with arrivals of multiple patient types with different capacity requirements modeled as independent Poisson processes.

However, including multiple patient types would make the state space even larger, and is therefore infeasible for this project. Appendix C gives a formulation for one $M|M|c|c$ queue with arrivals of multiple patient types.

This extension of the model could be used instead of the Erlang loss function to calculate the maximum admissible load for the ILP in Section 4.1.2. Therefore it is possible to include the effect of multiples births into an ILP model. However, because we want to compare the results of different models, we decided not to do that.

We apply this model extension to the example $M|M|c|c$ queue we used throughout this chapter to show the impact of including multiple births. In addition to the parameters of Section 4.2, we assume 76.4% of all patients are single births and the remaining 23.6% are twins (obtained from WKZ data of Oude Weernink (2018)). Including multiple births increases the rejection probability from 7.35% to 9.21%, which is a significant difference.

4.4 Analyzing the stochastic formulation: Discrete Event Simulation

Since the CTMC method is unable to solve large instances such as nine NICUs, we use Discrete Event Simulation (DES) as well. DES scales well with the number of NICUs, while CTMC scales well with the number of patients.

The DES application models the same network of queues as the CTMC. As run time of the simulation increases, the total travel time would approach the value found using CTMC. However, we apply two changes to the simulation to make it more realistic. First, we already reserve a bed while the patient is traveling to the NICU. And second, the simulation model allows us to use Hospital-to-NICU prioritization instead of NICU-to-NICU prioritization.

We simulate four replications of 100 years of time. We chose the number of replications using a sequential procedure such that the confidence interval half width would be less than 0.5% of the mean. Since differences between solutions will be small, a high confidence in the true value of a solution is necessary to compare different solutions. However, a trade-off must be made between the total simulation time of one experiment and the accuracy. The simulation time is approximately five seconds per replication. Replications use different random number seeds, but the seeds are the same for all experiments to reduce variance and allow for fair comparison.

Using Welch's graphical method for identifying the warm-up period, we decided to exclude the first fifteen years from data collection. It takes a long time for the average total annual travel time to stabilize. A part of this is due to the random arrival times of patients. A difference of a few patients a year on average can make it unfair to compare the result of different solutions. Therefore, we change the metric to average travel time per patient.

4.5 Optimization heuristics

In formulating an optimization heuristic to find a better (close-to-optimal) solution, we have to take the following characteristics into account:

- The search space is large (74 hospitals with 9 choices each, more than 4.1×10^{70} unique solutions) and contains many bad solutions. For example, assigning a hospital near Groningen to the NICU in Maastricht is unlikely to be optimal.
- Estimating or predicting the impact of a change in a solution is difficult due to the changing flow of patients within the network. A change will not only affect the two NICUs initially involved (e.g. using a move or a swap), but also the others because of competition between NICUs for finding a new place for rejected patients.
- Evaluating one solution is slow and takes at least 20 seconds (only 180/hour). Heavily relying on evaluating all neighbors of a solution is therefore not feasible.

For each solution we want to evaluate we have use the CTMC or DES model to calculate its performance, which can be time consuming. In addition we have to reduce the search space by disallowing certain combinations of hospitals and NICUs. To guide the process of moving from the current solution to a close-to optimal one, we use a metaheuristic. However, the metaheuristic might not have found an optimal solution after running for a specified time. Afterwards we check if we can find better performing neighbors of the best found solution. Our optimization heuristic consists of three steps, which we discuss in each of the following sections:

1. Reducing the search space (Section 4.5.1)
2. Finding a good quality solution using a metaheuristic (Section 4.5.2)
3. Improving the best found solution (Section 4.5.3)

4.5.1 Reducing the search space

We are able to decrease the search space by disallowing the combination of certain hospitals and NICUs. Consider the closest NICU to a hospital as the best choice. We propose to only allow assigning NICUs that have at most t minutes more travel time than the best choice. Doing it this way means more time is spent on evaluating alternatives for hospitals that do not *clearly* fall within a catchment area of one certain NICU.

For our case study, choosing $t = 70$ minutes allows for sufficient alternatives for each hospital, while disallowing many illogical combinations. Most hospitals now have between one to four options, instead of always nine. This results in a solution only having 140 neighbors compared to 592 initially. However, the search space remains extremely large (8.9×10^{29} unique solutions).

4.5.2 Metaheuristic

Many metaheuristics exists to find close-to optimal solutions to optimization problems. Some of the most popular and well-known are Simulated Annealing (SA), tabu search, and Variable Neighborhood Search (VNS). SA approximates a global optimum by incorporating a (time dependent) probability to accept a worse solution. Tabu search uses a local/neighborhood search procedure and prohibits evaluating solutions that have been visited recently. Variable Neighborhood Search uses a local search procedure to evaluate all neighbors of the incumbent solution, while changing neighborhoods to escape local optima when no better solution is found. However, Reduced VNS (RVNS, Hansen et al. (2019)) skips this time consuming local search step.

Since time is a limiting factor, we decide to use RVNS. Xiao et al. (2011) show that RVNS can quickly find a good quality solution. Figure 4.6 shows the steps of RVNS (Hansen et al. (2019)).

Initialization.
Select the set of neighborhood structures $\mathcal{N}_k$, for $k = 1, \ldots, k_{\max}$, that will be used in the search; find an initial solution x; choose a stopping condition;
Repeat the following sequence until the stopping condition is met:
(1) Set $k \leftarrow 1$;
(2) *Repeat* the following steps until $k = k_{\max}$:
(a) *Shaking*. Generate a point x' at random from the kth neighborhood of x ($x' \in \mathcal{N}_k(x)$);
(b) *Move or not*. If this point is better than the incumbent, move there ($x \leftarrow x'$), and continue the search with $\mathcal{N}_1$ ($k \leftarrow 1$); otherwise, set $k \leftarrow k + 1$;

Figure 4.6: The steps of RVNS (obtained from Hansen et al. (2019))

We choose to the define the neighborhood structure as the number of moves applied (e.g., hospital A is moved from NICU B to NICU C). We run RVNS until no better solution has been found after a certain number of consecutive experiments.

4.5.3 Improving the solution

Since the search space is large and RVNS includes a random component, we use steepest descent local search to try to improve the best solution that has been found so far. We evaluate all its neighbors that are within 1-move distance. Afterwards we check whether a better solution has been found. If so, then we set this neighbor as the best solution, and start a new iteration of steepest descent local search. We repeat this until no better solution is found.

4.6 Conclusion

The fourth question *"How can we analyze a network of $M|M|c|c$ queues in which rejected patients must be relocated?"* is answered in this chapter.

We formulated an ILP for both an uncapacitated and capacitated scenario. To model a capacity constraint, we defined a maximum offered admissible load for each NICU, which corresponds to a certain maximum allowed rejection rate. Transfers of patients are not included in these models.

Due to variability in the arrival process and length of stay, not all patients can be admitted to their own NICU. This means that these patient must be admitted to some other NICU. Including stochasticity of demand shifts the focus from individual NICUs towards network behaviour. A network is harder to analyse and it is more difficult to find an optimal solution.

We modeled the NICUs as a network of queues $M|M|c|c$, meaning there is no waiting room. Each NICU has their own Poisson arrival process for patients of their region. In case such a patient must be rejected because the NICU is fully occupied, the patient is admitted to another NICU. This new NICU is found using a predefined prioritization matrix.

This prioritization matrix indicates in which order NICUs are asked to admit this specific patient. First, the 1st-priority NICU is asked. If that NICU is unable to admit this patient, then the 2nd-priority NICU is asked, etc.. If all beds in the network are occupied, the patient leaves the network without being admitted somewhere. Prioritization can be based on a patient's initially assigned NICU (NICU-to-NICU prioritization), or by a patient's regional hospital (hospital-to-NICU prioritization). The second method is preferred, but requires differentiating between arrivals of all assigned regional hospitals to a NICU region.

We used two different methods to analyze this network of queues. The first method we used is the Continuous-Time Markov Chain (CTMC). CTMC is unable to solve large instances in reasonable time, so therefore we introduced Discrete Event Simulation (DES) as a second method.

In the CTMC the prioritization matrix is used for transferring arriving patients to other NICUs, if their primary NICU was fully occupied. Since we are interested in the long term behavior, we calculate the steady state state distribution π. For shorter computation times we could also approximate π using the power method. We used the steady state distribution and the PASTA property of the arrivals to construct an admission table, which we could use to calculate performance indicators, such as the total travel time.

A problem with CTMC is that the number of the states increase exponentially with the included number of NICUs. We can find a lower bound of the number of admission at each NICU using the steady state distribution of individual $M|M|c|c$ queues. This reduces the total state space, but still only allowed us to evaluate networks of at most five NICUs in reasonable time.

We introduced Discrete Event Simulation as the second method to solve larger instances. This model, after a sufficient run time, approximates the result obtained from a CTMC. DES also enabled us to use hospital-to-NICU prioritization.

If the total state space allows for it, the CTMC method is preferred because there is no uncertainty in the mean value of the chosen performance indicator. Furthermore, DES scales well with number of NICUs, while CTMC scales well with number of patients (arrival rate).

When designing an optimization heuristic, we had to take several points into account. The search space is extremely large (4.1×10^{70} unique solutions) and contains many bad solutions. Evaluating one solution is slow and takes at least 20 seconds (only 180/hour).

We introduced an optimization heuristic consisting of three steps. First, we decrease the search space by disallowing the combination of certain hospitals and NICUs. Consider the closest NICU to a hospital as the best choice. We proposed to only allow assigning NICUs that have at most t minutes more travel time than the best choice. In the second step we use the metaheuristic Reduced Variable Neighbourhood Search (RVNS) to quickly find a good quality solution. RVNS evaluates random points from neighborhoods around the incumbent solution. These neighborhoods are defined by a set of neighborhood structures. If either a better solution has been found or all neighborhoods structures have been considered, the process starts again at the first neighborhood structure. These iterations continue until a stopping condition is reached. In the third step, after applying RVNS, we used steepest descent with a 1-move neighborhood search until no improvements can be found.

Chapter 5

Results

In this chapter we will analyse the results we found using the models formulated in Chapter 4. Section 5.1 summarizes the values of the input parameters we have used. Section 5.2 shows the results of the deterministic models, while Section 5.3 shows the results of the stochastic model in which patient transfers are included. Afterwards, we evaluate the impact of including transfers in Section 5.4. Furthermore, Section 5.5 shows how the capacity allocation can be altered to gain further improvements. And finally, Section 5.6 illustrates how spatial accessibility to NIC changes for the optimal capacity allocation.

Since there are 74 hospitals, it is tedious to show the entire assignment of hospitals to NICUs for each alternative solution. Therefore, we only report changes compared to the current solution. Appendix D lists the current assignment of hospitals to NICUs. For convenience, Figure 5.1 visualizes the current assignment once more.

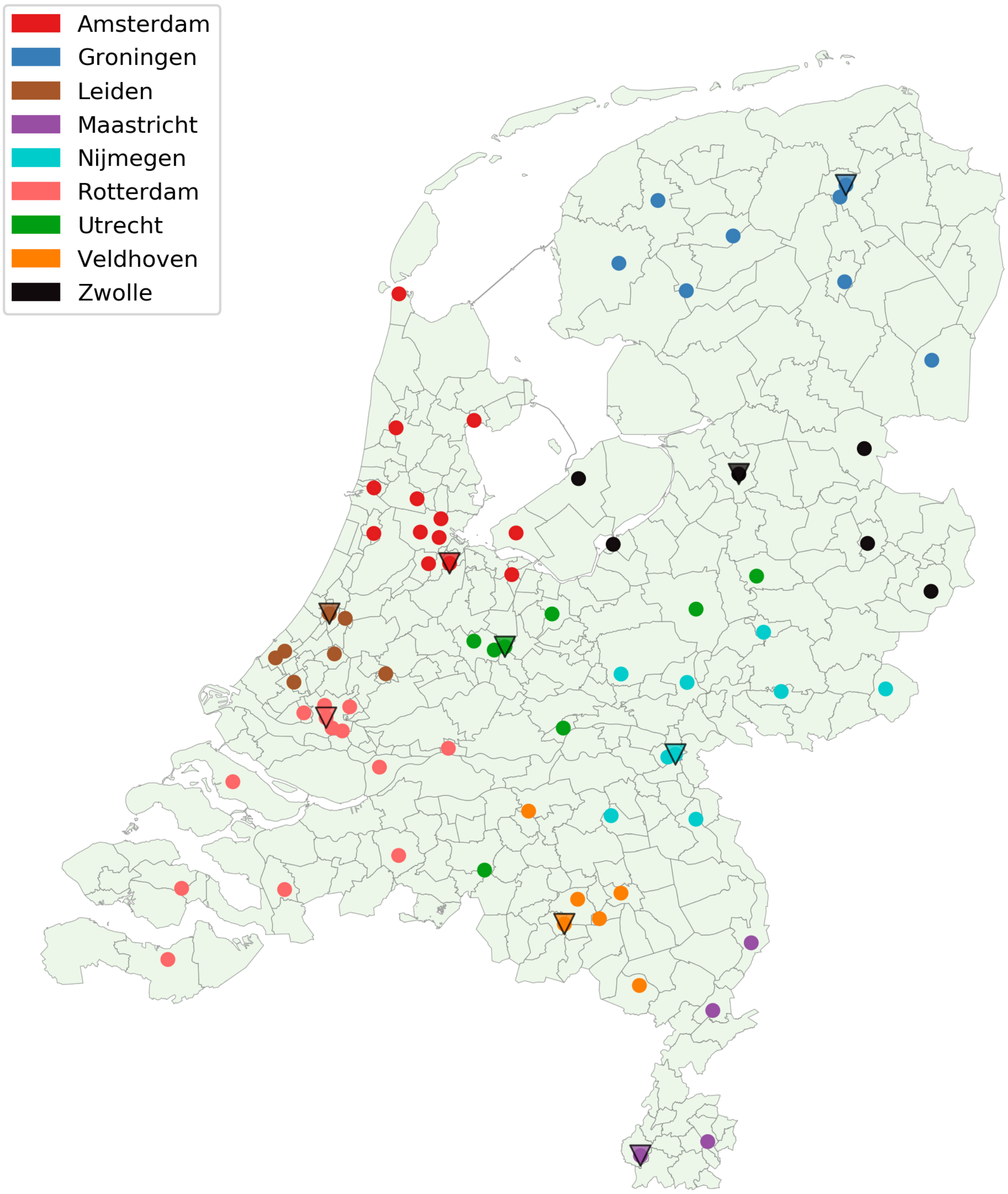

Figure 5.1: The current assignment of general hospitals to NICUs.

5.1 Input parameters

Throughout this report we have gathered input parameters for the models. In this section we briefly summarize the values of the parameters. In remaining part of this chapter we used these values, unless explicitly mentioned otherwise.

The average NICU length of stay was approximately 11.63 days in 2015 (Perined (2016)). Therefore, we used a service rate $\mu = \frac{365}{11.63}$ per year for all NICUs. Table 5.1 shows the recent operational capacity of each NICU.

We obtained a dataset of the number of births allocated to each regional hospital from Perined. As mentioned before, this dataset is confidential and will therefore not be shared. For estimating the number of NIC patients from the number of births of each hospital, we used the upper value of the confidence interval ($\hat{p} = 0.2394$) to provide a robust solution. We expect a total of 4018 NICU admissions a year.

NICU	**Operational capacity**
Amsterdam	28
Groningen	16
Leiden	17
Maastricht	13
Nijmegen	12
Rotterdam	25
Utrecht	20
Veldhoven	15
Zwolle	17

Table 5.1: The operational capacity of each NICU

5.2 Deterministic model

In this section we discuss the results of the uncapacitated and the capacitated variant of an ILP formulation for assigning hospitals to NICUs. For the capacitated model we used the maximum admissible load constraint and compare the effect of choosing different rejection probabilities.

5.2.1 Uncapacitated model

Figure 5.2 shows the solution of the uncapacitated model. Table 5.2 shows the changes compared to the current solution (see Appendix D for the entire list of assignments). We found on average 55.67 minutes of travel time per patient.

According to our definition of the objective function, i.e. the way we define and calculate travel time, the uncapacitated scenario is also the best case scenario. Although the lack of transfers of patients is unrealistic, it provides us with a lower bound (LB) of the objective function, which is used for assessing the quality of other solutions.

The LB provides an estimate of how much improvement is theoretically possible. A higher average travel time than the LB means there is some potential for improvement. Preferably, improvements may be found by reassessing the current assignment of hospitals to NICUs. An even lower travel time can be found by assessing both the assignment and the capacity allocation amongst NICUs. However, we can only truly approach the value of the LB by drastically increasing capacity to absorb variation in arrivals. So therefore realistically, there will always be transfers of patients.

Hospital	**Current**	**Best**
Antonius Ziekenhuis Sneek	Groningen	Zwolle
Beatrixziekenhuis	Rotterdam	Utrecht
Deventer Ziekenhuis	Utrecht	Zwolle
ETZ Elisabeth	Utrecht	Veldhoven
Gelre Ziekenhuizen Apeldoorn	Utrecht	Zwolle
Groene Hart Ziekenhuis Gouda	Leiden	Rotterdam
HagaZiekenhuis Leyweg	Leiden	Rotterdam
Reinier de Graaf Gasthuis	Leiden	Rotterdam
VieCuri Medisch Centrum Venlo	Maastricht	Veldhoven
Ziekenhuis Bernhoven	Nijmegen	Veldhoven
Ziekenhuis De Gelderse Vallei locatie Ede	Nijmegen	Utrecht
Ziekenhuis de Tjongerschans	Groningen	Zwolle
Ziekenhuis St. Jansdal Harderwijk	Zwolle	Utrecht
Ziekenhuis St. Jansdal Lelystad	Zwolle	Amsterdam

Table 5.2: Changes made in the best uncapacitated solution compared to the current assignment.

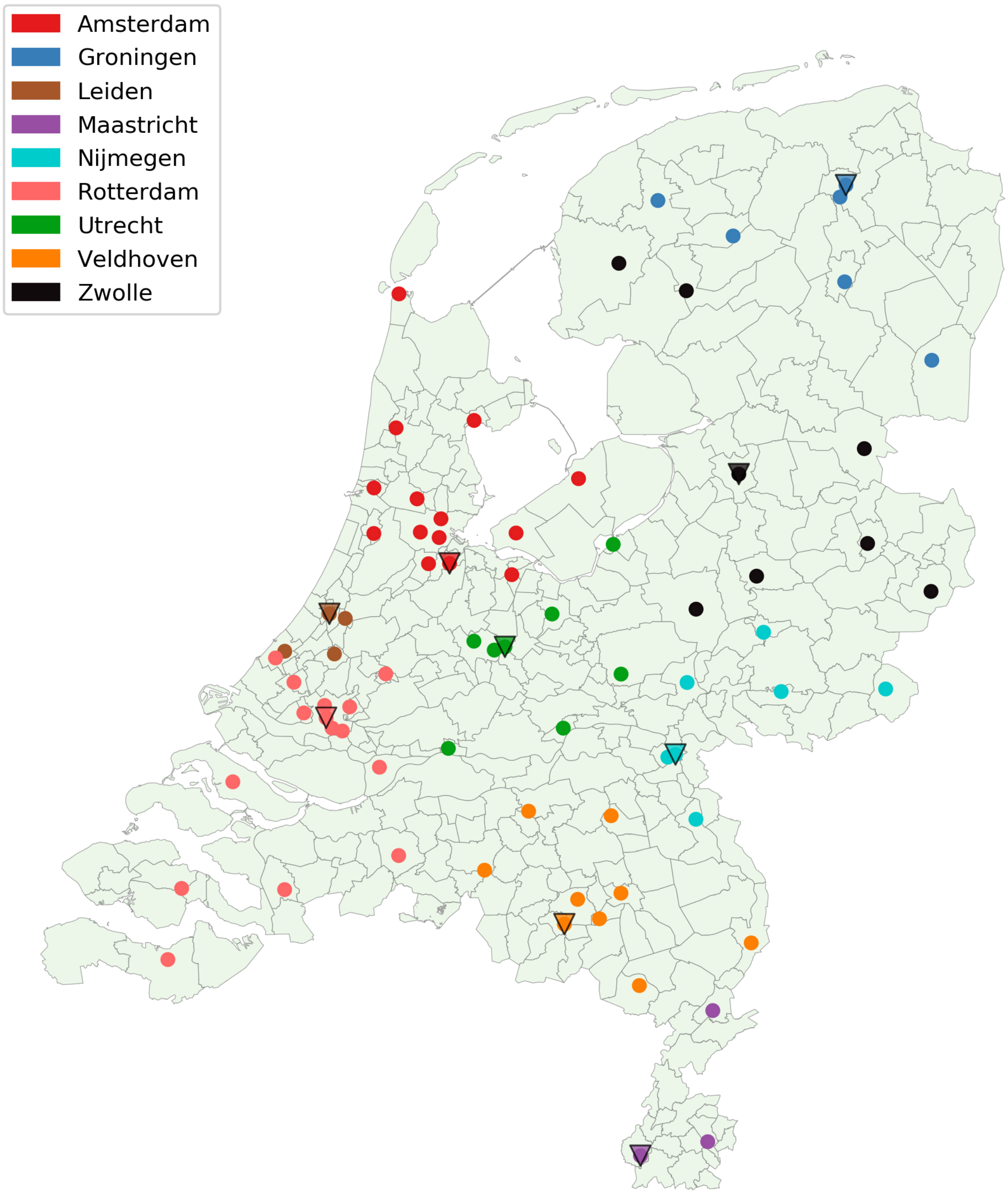

Figure 5.2: The optimal assignment of the uncapacitated scenario.

5.2.2 Capacitated model

As mentioned before, we used the maximum admissible offered load constraint for the capacitated scenario without transfers of patients. Table 5.3 shows that the average travel time decreases when the maximum allowed rejection rate increases. This is as expected since transfers of patients are not taken into account. The higher the maximum allowed rejection probability, the more the solution converges towards the uncapacitated scenario. The solution with a maximum rejection rate of 20% is equal to the result of the uncapacitated scenario.

Not including transfers of patients might be misleading and too optimistic. To analyze the impact of including transfers, we will run the solutions of Table 5.3 in the stochastic model as well. We will compare the results later in Section 5.4.

Max rejection %	**Avg travel time**
7.5	58.17
10	56.43
12.5	56.04
15	55.83
17.5	55.73
20	55.67

Table 5.3: The effect of different maximum rejection probabilities on the average travel time.

5.3 Stochastic model

We use the best assignment resulting from the deterministic model without capacity restrictions as the absolute lower bound of the travel time. Although this is an optimistic value, it gives perspective on how much improvement could be made. We found a lower bound of on average 55.67 minutes of travel time per patient.

We ran our optimization heuristic twice, with different initial solutions. The first run started with the current solution (*current-to-optimal*), and the second run with the solution of the uncapacitated scenario (*uncapacitated-to-optima*). In RVNS we used neighborhood structures of 1-move, 2-moves, and 3-moves. We terminated the RVNS procedure after no improvement of the incumbent solution was found after 300 consecutive experiments. Afterwards, we used steepest descent local search until no more further improvements could be found.

In our optimization heuristic we made the trade-off between run time and accuracy. To reduce the width of the 95% confidence interval of the true mean of the average travel time, we ran additional replications with longer simulation length (10 replications of 1000 years) for the best found solutions.

We found a travel time reduction of approximately 4.6 minutes on average per patient compared the to current situation (Table 5.4). To put in perspective, this equals to a total reduction of about 18423 minutes or 308 hours of travel time a year. Compared to the lower bound, this is a reduction of 25.8% of the undesirable transport time. In addition, the number of transfers are reduced by 15.7%.

Scenario	Travel time per patient		Transfers per year	
	Avg (minutes)	*95% CI*	*Avg (number of)*	*95% CI*
Current	73.47	[73.38, 73.56]	690	[687, 692]
Current-to-optimal	68.87	[68.81, 68.93]	585	[583, 587]
Uncapacitated-to-optimal	68.89	[68.82, 68.96]	578	[576, 581]

Table 5.4: Comparison of the current and best solutions.

Both solutions we found, *current-to-optimal* and *uncapacitated-to-optimal*, perform equally good. Table 5.5 shows the changes to the current situation that both solutions have in common. However, there are three hospitals that have a different assignment (Table 5.6). Figure 5.3 and 5.4 visualize the assignment of the solutions *current-to-optimal* and *uncapacitated-to-optimal*, respectively.

Hospital	Current	Optimal
Beatrixziekenhuis	Rotterdam	Utrecht
Deventer Ziekenhuis	Utrecht	Zwolle
ETZ Elisabeth	Utrecht	Veldhoven
Gelre Ziekenhuizen Apeldoorn	Utrecht	Zwolle
Groene Hart Ziekenhuis Gouda	Leiden	Rotterdam
Reinier de Graaf Gasthuis	Leiden	Rotterdam
St. Jans Gasthuis	Veldhoven	Maastricht
Ziekenhuis Bernhoven	Nijmegen	Veldhoven
Ziekenhuis De Gelderse Vallei locatie Ede	Nijmegen	Utrecht
Ziekenhuis St. Jansdal Lelystad	Zwolle	Amsterdam
ZorgSaam De Honte	Rotterdam	Maastricht

Table 5.5: Changes made that the best solutions have in common, compared to the current assignment.

Hospital	Current	Current-to-optimal	Uncapacitated-to-optimal
NWZ Den Helder	Amsterdam	Amsterdam	Groningen
SKB	Nijmegen	Zwolle	Nijmegen
Ziekenhuis Rivierenland	Utrecht	Nijmegen	Utrecht

Table 5.6: The differences between the best solutions, compared to the current assignment.

Compared to the current situation, we observe that the catchment areas are now more defined. One interesting change in both of the optimal solutions is that hospital "ZorgSaam De Honte" (Zeeland, southwest corner) is now assigned to Maastricht, instead of Rotterdam. Since the expected demand of that hospital is rather low, and Maastricht transfers fewer patients than Rotterdam (Figure 5.5), the additional total travel time via Belgium is smaller than the travel time that would otherwise be spent by Rotterdam on additional transfers.

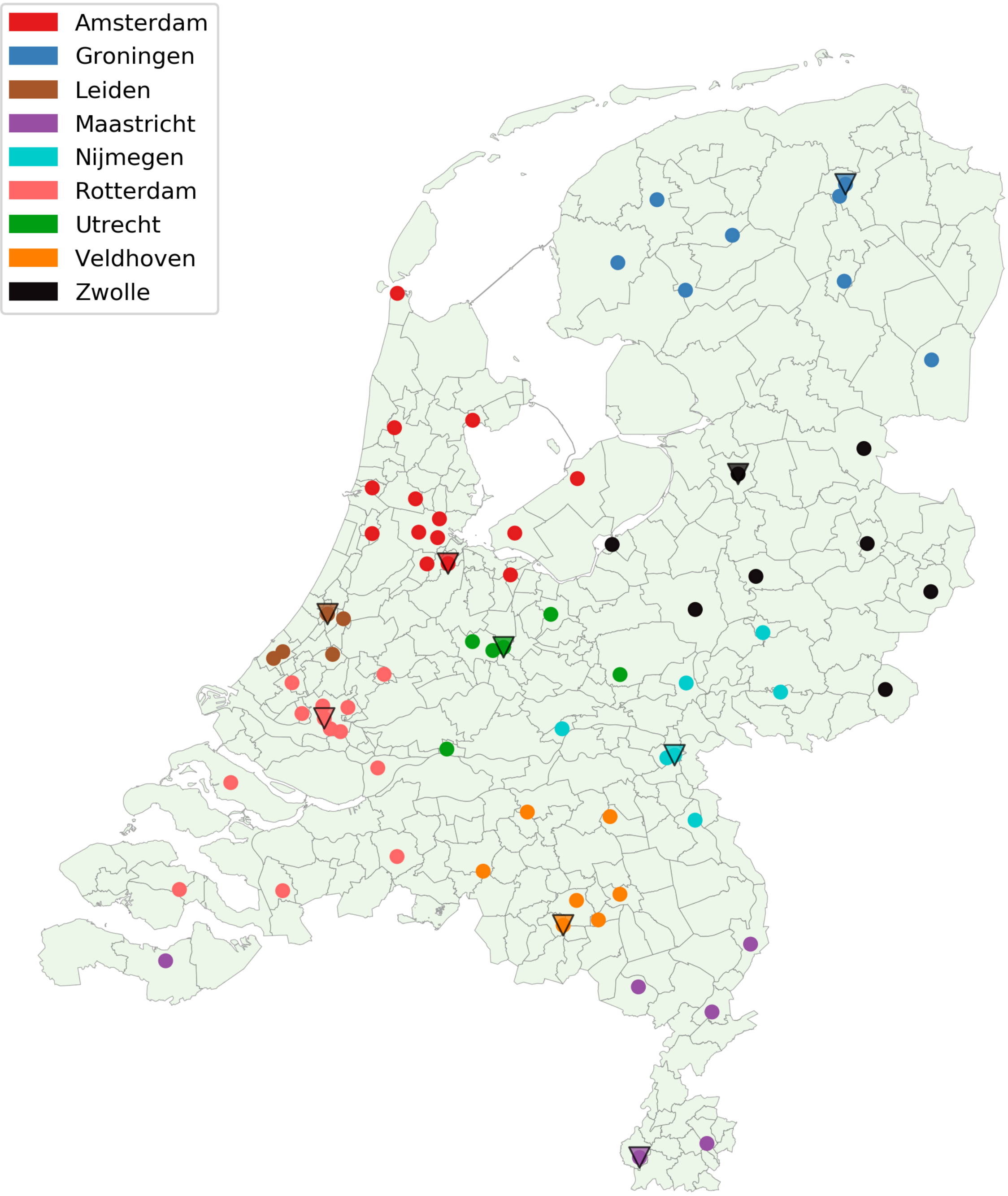

Figure 5.3: The assignment of hospitals to NICUs for the *current-to-optimal* solution.

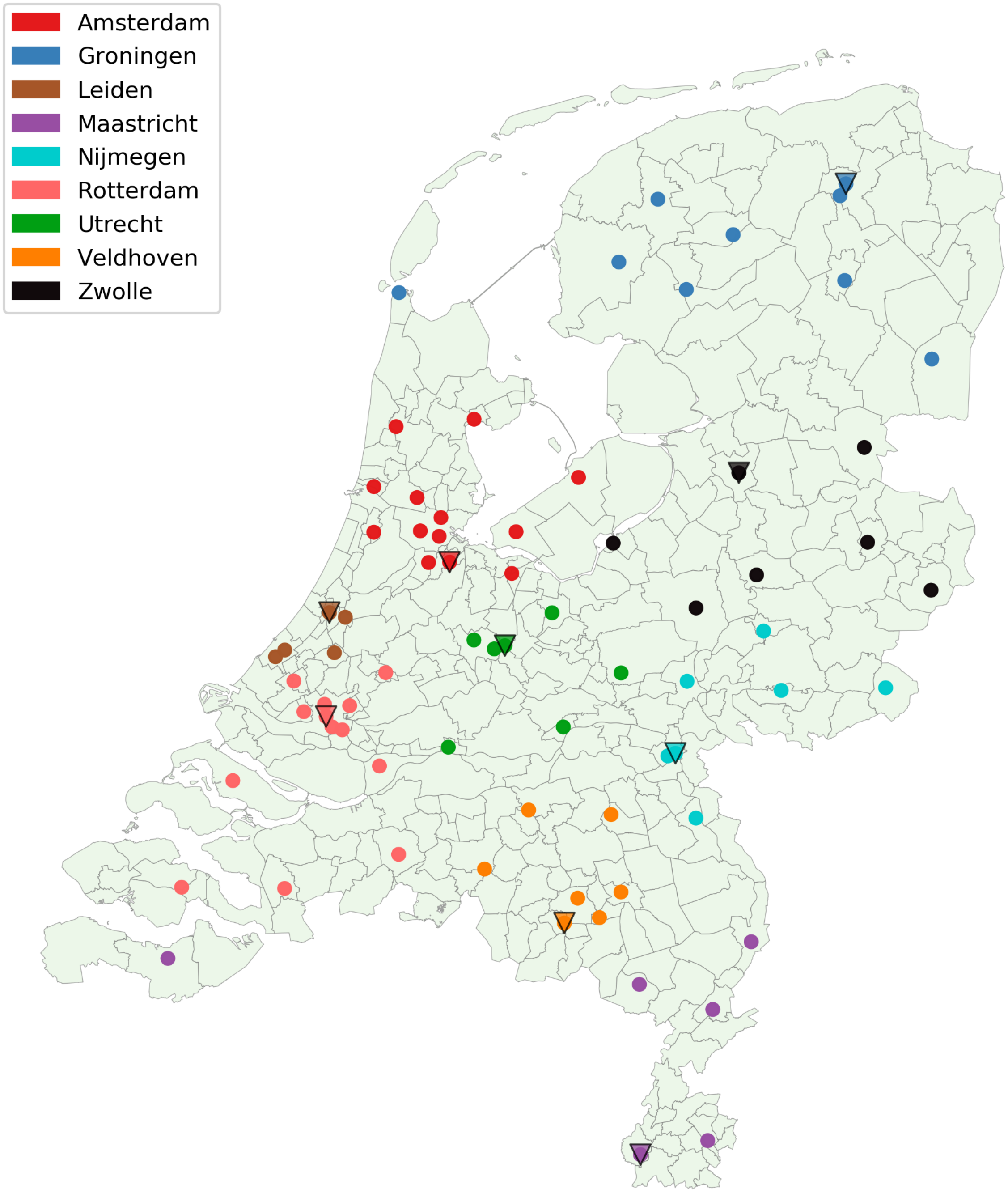

Figure 5.4: The assignment of hospitals to NICUs for the *uncapacitated-to-optimal* solution.

Admission table *current situation*	Amsterdam	Groningen	Leiden	Maastricht	Nijmegen	Rotterdam	Utrecht	Veldhoven	Zwolle	No place found	Total
Demand Amsterdam	643.2	1.9	67.5	0.8	1	23.2	32.9	4.7	12.4	0.4	788
Demand Groningen	2.1	329.1	0.3	0.2	0.3	0.2	0.8	0.1	14.3	0.2	347.6
Demand Leiden	28.4	0.4	342.3	1	1.1	59.3	6.4	5.1	2	0.2	446.2
Demand Maastricht	0.1	0.1	0	170.1	0.3	0.1	0.1	1.1	0.1	0.1	172.1
Demand Nijmegen	6.8	0.4	0.2	1.4	293.5	0.8	37.6	40.5	21.3	0.2	402.7
Demand Rotterdam	19.5	0.4	32.9	2.1	1.8	570.8	29	22	2.5	0.3	681.3
Demand Utrecht	64.3	0.7	6.1	1.5	8.7	13.1	421.9	22.7	25.3	0.2	564.5
Demand Veldhoven	1.2	0.1	0.1	14.5	7	0.5	6.4	247.4	0.1	0.1	277.4
Demand Zwolle	3.4	9	0.6	0.2	1.8	0.2	13.8	0.2	309.2	0.2	338.6
Total	769	342.1	450	191.8	315.5	668.2	548.9	343.8	387.2	1.9	4018.4

Admission table *Current-to-optimal*	Amsterdam	Groningen	Leiden	Maastricht	Nijmegen	Rotterdam	Utrecht	Veldhoven	Zwolle	No place found	Total
Demand Amsterdam	675.9	1.8	75.2	0.6	1.1	18	32.5	3.9	9.3	0.4	818.7
Demand Groningen	2.5	329.3	0.3	0.2	0.3	0.1	1.1	0.1	13.6	0.2	347.7
Demand Leiden	15.3	0.3	278	0.4	0.8	23.3	3.3	2.2	1.1	0.1	324.8
Demand Maastricht	0.2	0.1	0	203.5	0.9	0.4	0.3	2.9	0.1	0.1	208.5
Demand Nijmegen	2.2	0.3	0.2	0.9	241.8	0.5	18.2	14.7	7.6	0.1	286.5
Demand Rotterdam	19.1	0.4	51.8	1.5	2	617.4	32	17.7	1.8	0.3	744
Demand Utrecht	47.1	0.3	5.5	0.8	11.3	13	383.3	2.7	5.6	0.2	469.8
Demand Veldhoven	1.9	0.1	0.1	17.6	19.9	3.3	23.8	332.6	0.2	0.2	399.7
Demand Zwolle	2.3	9.4	0.5	0.2	6.1	0.2	28.9	0.4	370.9	0.2	419.1
Total	766.5	342	411.6	225.7	284.2	676.2	523.4	377.2	410.2	1.8	4018.8

Figure 5.5: The admission tables of the current solution and *current-to-optimal* solution.

From the admission table (Figure 5.5) we can derive other statistics and performance measures than just travel time. Figure 5.6 shows the occupancy, the percentage of patients admitted to their own region, and how much of the capacity is spent on patients from outside a NICU's own region. The results of the solutions *current-to-optimal* and *uncapacitated-to-optimal* are very similar, and therefore we only show one of them.

We observe that NICUs with more beds can operate at a higher occupancy. Furthermore, two of the more isolated NICUs, Groningen and Maastricht, are able to admit almost all of their own patients. Moreover, we see that Leiden and Utrecht still spend a significant part of their capacity on admitting patients from other NICUs. Looking at their geographical location, it seems reasonable that those NICUs often relieve other neighbouring NICUs.

	Current			Current-to-optimal		
NICU	% occupancy	% own region admitted	% capacity spent on outside region	% occupancy	% own region admitted	% capacity spent on outside region
Amsterdam	88%	82%	14%	87%	83%	10%
Groningen	68%	95%	3%	68%	95%	3%
Leiden	84%	77%	20%	77%	86%	25%
Maastricht	47%	99%	5%	55%	98%	5%
Nijmegen	84%	73%	6%	75%	84%	11%
Rotterdam	85%	84%	12%	86%	83%	7%
Utrecht	87%	75%	20%	83%	82%	22%
Veldhoven	73%	89%	20%	80%	83%	9%
Zwolle	73%	91%	15%	77%	88%	7%

Figure 5.6: Performance statistics of the current solution and the *uncapacitated-to-optimal* solution.

5.4 Impact of including transfers

Table 5.7 shows that not including transfers of patients leads to misleading results. As can be seen, a higher maximum allowed rejection probability leads to a lower travel time when disregarding transfers. However, this leads to more transfers per year, which then results in higher travel times.

However, this does not mean the deterministic model should never be used. The assignments obtained from the deterministic model all perform better than the current solution, though none of them are optimal. The deterministic model is less complex, easier to use, and has almost infinitely smaller run time than the stochastic model. However, it is difficult to decide on a value for the capacity constraint, since it is hard to interpret its effect.

Rejection probability	Not including transfers	Including transfers	
	Avg travel time	*Avg travel time*	*# transfers per year*
7.5%	58.17	69.56	556
10%	56.43	69.58	626
12.5%	56.04	71.03	707
15%	55.83	70.91	720
17.5%	55.73	71.70	758
20%	55.67	71.67	755

Table 5.7: The impact of including transfers of patients.

NICU	Assigned demand	Required number of beds
Amsterdam	819	27
Groningen	283	10
Leiden	246	8
Maastricht	136	5
Nijmegen	276	9
Rotterdam	839	27
Utrecht	548	18
Veldhoven	456	15
Zwolle	416	14

Table 5.8: The assigned demand and deterministic required number of beds for the uncapacitated scenario.

5.5 Nationwide capacity allocation

A different capacity allocation than the current one may provide a better match between service and local demand. Therefore, we investigate what the optimal capacity allocation would be and how additional capacity would impact performance.

To find the optimal capacity allocation we used the following method. To minimize the travel time, we used the results of the uncapacitated scenario as a guideline, because that is the solution that we should strive for. Therefore, we used the assignment of the uncapacitated scenario and calculated the required number of beds at each NICU, assuming demand is deterministic (Table 5.8). From this follows that a total of 133 beds would be required. We used the capacity allocation of Table 5.8 as a starting point in the simulation model to iteratively allocate additional capacity. We evaluated at which NICU a capacity increase of one bed would have had the most impact, and allocated an additional bed. We repeated this process until we had allocated all available beds.

By reallocating all 163 beds using this procedure, we found an average travel time of approximately 65.08 minutes per patient. This is a further improvement of 3.8 minutes on the best value we had found in Section 5.3 by merely assigning hospitals to NICUs. In addition to this, the expected number of transfers per year also decreased by another 138. Compared to the current situation, we found a reduction of 47% in undesirable travel time and a decrease of 35.5% in the number of transfers. Table 5.9 shows the optimal capacity allocation for different numbers of total network capacity.

NICU	*Current (163)*	Optimal 163	Optimal 170	Optimal 180
Amsterdam	*28*	32	33	34
Groningen	*16*	12	12	14
Leiden	*17*	10	11	12
Maastricht	*13*	6	7	7
Nijmegen	*12*	11	12	14
Rotterdam	*25*	32	33	34
Utrecht	*20*	24	24	24
Veldhoven	*15*	19	20	21
Zwolle	*17*	17	18	20

Table 5.9: The optimal capacity allocation for some total network capacity.

We checked using steepest descent local search whether the assignment of the uncapacitated model was indeed (likely to be) optimal for this capacity allocation we had found. We could not find an improvement in the assignment of hospitals to NICUs. This indicates that the network has sufficient operational beds available to it, because the optimal assignment of the uncapacitated model is now also optimal after reallocating capacity. Therefore, further improvements, converging towards the lower bound, can only be made by increasing capacity of the network.

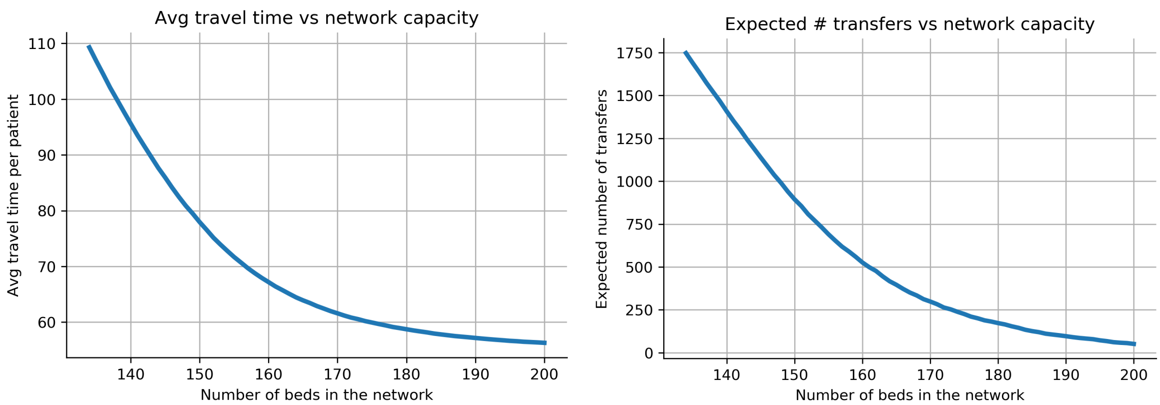

Figure 5.7: The average travel time per patient (left) and expected number of transfers per year (right) compared to the network capacity.

Figure 5.7 shows how the average travel time per patient and the number of transfers per year relate to the capacity of the network. From this it is clear that calculating the required capacity using deterministic demand (133 beds) leads to a severe underestimation of the travel time and number of transfers. Figure 5.8 gives a closer look to how increasing capacity of the network decreases the number of transfers. It seems that increasing capacity moderately from 163 to 170 beds, can still result in a significant and efficient reduction in transfers. Moving towards and beyond 190 beds results in high diminishing returns.

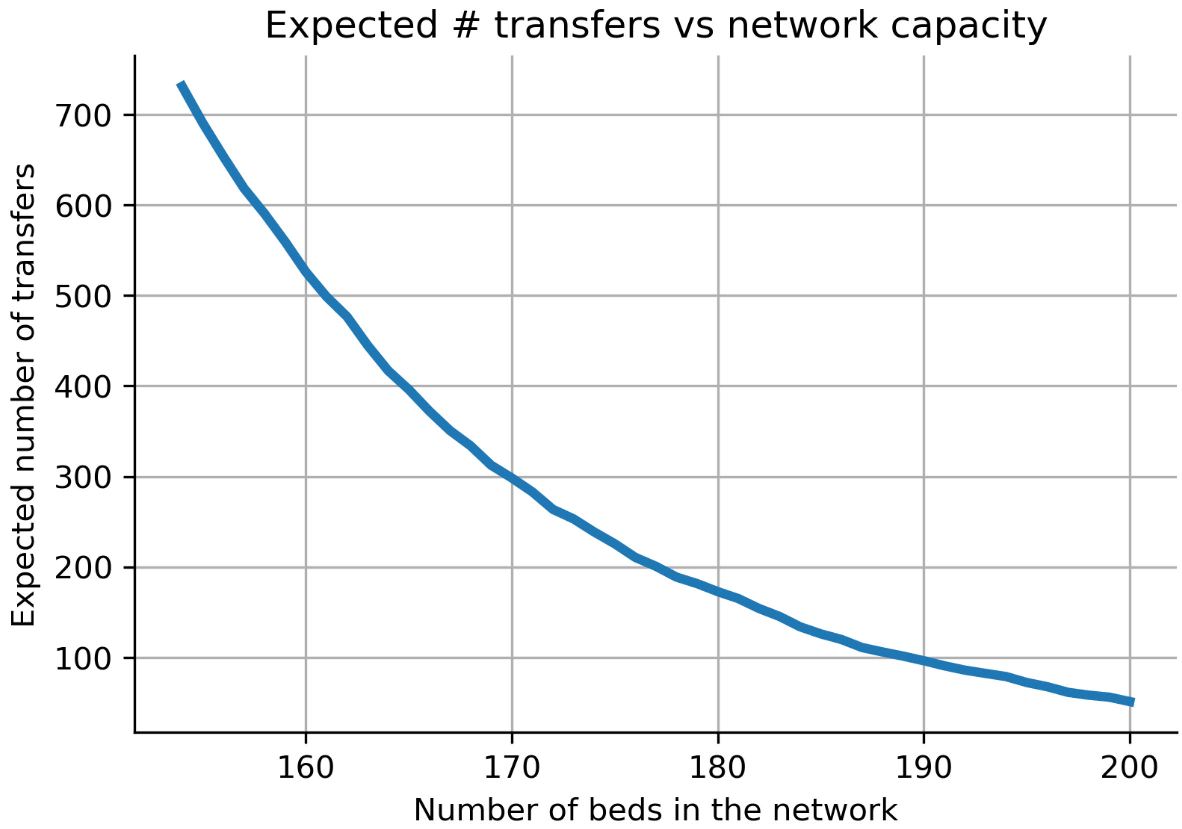

Figure 5.8: A closer look to the expected number of transfers per year compared to the network capacity.

5.6 Spatial accessibility

We discussed spatial accessibility earlier in Section 3.3.5. Now, we compare the difference in spatial accessibility between the current and optimal capacity allocation. Again, we use E2SFCA with travel time zones {30 minutes, 60 minutes, 90 minutes} and weights {1, 0.6, 0.13}. Figure 3.6 shows the resulting spatial accessibility to NIC. From this figure it is clear that more capacity is allocated to NICUs in the Randstad. Overall, it seems that by using optimal capacity allocation there is a better match between service and local demand.

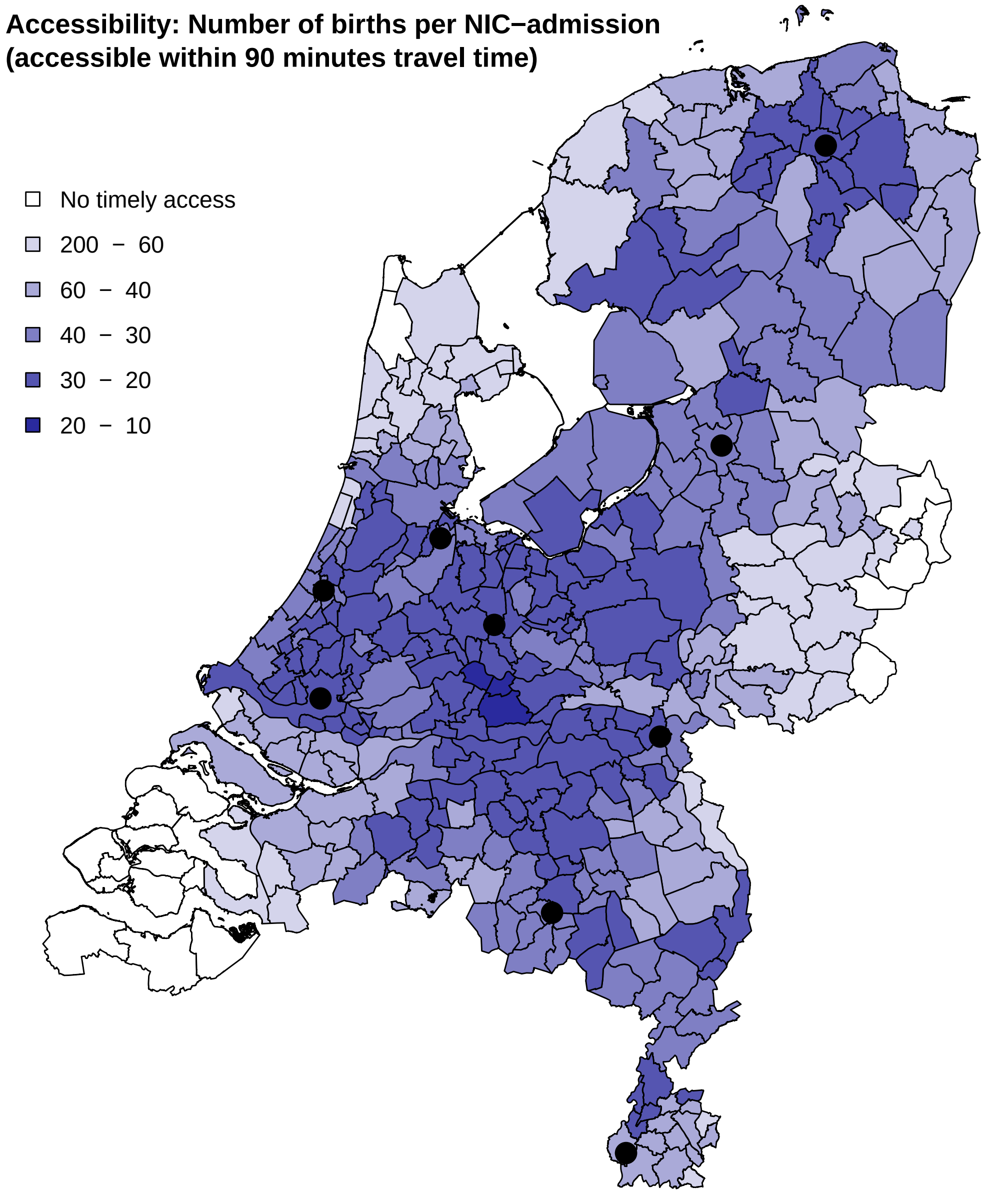

Figure 5.9: Spatial accessibility to NIC measured using E2SFCA.

5.7 Conclusion

The fifth question *"What are the effects of different assignments of general hospitals to Neonatal Intensive Cares?"* is answered in this chapter.

According to the way we defined and calculated travel time, the uncapacitated scenario is the best case scenario. Although the lack of transfers of patients is unrealistic, it provides us with a lower bound (LB) of the objective function, which is used for assessing the quality of other solutions. The LB provides an estimate of how much improvement is theoretically possible. We found a LB of on average 55.67 minutes of travel time per patient.

For the capacitated ILP model, the average travel time decreased when the maximum allowed rejection rate increased. This is as expected since transfers of patients are not taken into account. The higher the maximum allowed rejection probability, the more the solution converges towards the uncapacitated scenario.

For the stochastic model, we ran our optimization heuristic twice, with different initial solutions. The first run started with the current solution (*current-to-optimal*), and the second run with the solution of the uncapacitated scenario (*uncapacitated-to-optima*). In RVNS we used neighborhood structures of 1-move, 2-moves, and 3-moves. We terminated the RVNS procedure after no improvement of the incumbent solution was found after 300 consecutive experiments. Afterwards, we used steepest descent local search until no more further improvements could be found.

We found a travel time reduction of approximately 4.6 minutes on average per patient compared the to current situation. Compared to the lower bound, this is a reduction of 25.8% of the undesirable transport time. In addition, the number of transfers are reduced by 15.7%.

A different capacity allocation than the current one may provide a better match between service and local demand. Therefore, we investigated what the optimal capacity allocation would be and how additional capacity would impact performance.

To minimize the travel time, we used the results of the uncapacitated scenario as a guideline, because that is the solution that we should strive for. In addition, we started with the minimum required capacity for deterministic demand (133 beds). We evaluated at which NICU a capacity increase of one bed would have had the most impact, and allocated an additional bed. We repeated this process until we had allocated all available beds.

By reallocating all 163 beds using this procedure, we found an average travel time of approximately 65.08 minutes per patient. Compared to the current situation, we found a reduction of 47% in undesirable travel time and a decrease of 35.5% in the number of transfers.

It is clear that calculating the required capacity using deterministic demand (133 beds) leads to a severe underestimation of the travel time and number of transfers. Furthermore, it seems that increasing capacity moderately from 163 to 170 beds, can still result in a significant and efficient reduction in transfers. Moving towards and beyond 190 beds results in high diminishing returns.

Chapter 6

Conclusion

In this Chapter conclude our research. In Section 6.1 we answer our main research question. Then, in Section 6.2, we discuss the limitations of our research. Afterwards we formulate our recommendations for the neonatal care network and discuss suggestions for further research, in Sections 6.3 and 6.4 respectively. We end this chapter by mentioning the scientific contributions of this research in Section 6.5.

6.1 Conclusion

In this report we answered our main research questions *"Which assignments of general hospitals to Neonatal Intensive Cares lead to minimized transportation time?"*. This questions was split up into multiple smaller sub-questions.

We analyzed logistical processes at a NICU and learned that neonatology care is not only complex medically, but also logistically. Each NICU has their own catchment region, which is composed of a certain number of general hospitals. In case it is infeasible for a NICU to admit a patient from its catchment area, then a transfer must take place. Transferring a patient is difficult and undesirable for everyone involved.

Using data from Perined on the number of births at each hospital, we estimated the expected NICU demand of each hospital. In addition we gathered all travel times between hospital and NICUs. And finally, we obtained the current assigned of hospitals to NICUs.

To determine our own catchment areas, we first formulated an Integer Linear Programming (ILP) model to assign hospitals to NICUs. We used both an uncapacitated and capacitated scenario. To model a capacity constraint, we defined a maximum offered admissible load for each NICU, which corresponds to a certain maximum allowed rejection rate. Transfers of patients are not included in these ILP models.

To include transfers of patients, we modeled the NICUs as a network of queues $M|M|c|c$ without waiting rooms or buffers. Each NICU has their own Poisson arrival process for patients of their region. In case such a patient must be rejected because the NICU is fully occupied, the patient is admitted at another NICU. This new NICU is found using a predefined prioritization matrix.

We used two different methods to analyze this network of queues. The first method we used is the Continuous-Time Markov Chain (CTMC). CTMC is unable to solve large instances in reasonable time, so therefore we introduced Discrete Event Simulation (DES) as a second method.

To find an optimal solution, We formulated an optimization heuristic consisting of three steps. First, we decrease the search space by disallowing the combination of certain hospitals and NICUs. Consider the closest NICU to a hospital as the best choice. We proposed to only allow assigning NICUs that have at most t minutes more travel time than the best choice. In the second step we use the metaheuristic Reduced Variable Neighbourhood Search (RVNS) to quickly find a good quality solution. In the third step we used steepest descent with a 1-move neighborhood search until no improvements can be found.

According to the way we defined and calculated travel time, the uncapacitated scenario is the best case scenario. Although the lack of transfers of patients is unrealistic, it provides us with a lower bound (LB) of the objective function, which is used for assessing the quality of other solutions. The LB provides an estimate of how much improvement is theoretically possible. We found a LB of on average 55.67 minutes of travel time per patient.

Using the stochastic model, we found for the current situation a travel time of 73.47 minutes on average per patient and 690 transfers per year. By applying our optimization heuristic, we found a travel time reduction of approximately 4.6 minutes on average per patient compared the to current situation. Compared to the lower bound, this is a reduction of 25.8% of the undesirable transport time. In addition, the number of transfers are reduced by 15.7%. The optimal solution includes all changes of Table 6.1 and one set of changes of Table 6.2.

Further improvement can be found by reallocating all 163 beds within the network. We found an average travel time of approximately 65.08 minutes per patient, which is a reduction of 47% in undesirable travel time and a decrease of 35.5% in the number of transfers. In addition, calculating the required capacity using deterministic demand (133 beds) leads to a severe underestimation of the travel time and number of transfers. Furthermore, increasing capacity slightly from 163 to 170 beds, can still result in a significant and efficient reduction in transfers. Moving towards and beyond 190 beds results in high diminishing returns.

Hospital	Current	Optimal
Beatrixziekenhuis	Rotterdam	Utrecht
Deventer Ziekenhuis	Utrecht	Zwolle
ETZ Elisabeth	Utrecht	Veldhoven
Gelre Ziekenhuizen Apeldoorn	Utrecht	Zwolle
Groene Hart Ziekenhuis Gouda	Leiden	Rotterdam
Reinier de Graaf Gasthuis	Leiden	Rotterdam
St. Jans Gasthuis	Veldhoven	Maastricht
Ziekenhuis Bernhoven	Nijmegen	Veldhoven
Ziekenhuis De Gelderse Vallei locatie Ede	Nijmegen	Utrecht
Ziekenhuis St. Jansdal Lelystad	Zwolle	Amsterdam
ZorgSaam De Honte	Rotterdam	Maastricht

Table 6.1: Changes made that the best solutions have in common, compared to the current assignment.

Hospital	Current	Option 1	Option 2
NWZ Den Helder	Amsterdam	Amsterdam	Groningen
SKB	Nijmegen	Zwolle	Nijmegen
Ziekenhuis Rivierenland	Utrecht	Nijmegen	Utrecht

Table 6.2: The differences between the best solutions, compared to the current assignment.

6.2 Discussion

We would have preferred to find an optimal solution using the CTMC method. However, working with network of this size has proven to be difficult. At most 180 solutions could be evaluated per hour, which is very little considering there are 74 hospitals to be allocated to a NICU. Performing a proper sensitivity analysis was therefore infeasible due to time constraints. Smaller networks should be able to use the CTMC method.

Our main limitation is that we lack high quality data for our analysis and input parameters. Obtaining information about catchment areas of hospitals, or about statistics concerning NIC turned out to be either impossible, or much harder than anticipated. Other than that, we would have liked to fully analyze the patient origins and patient flow in the network. However, this would require an extensive dataset of admissions from many different hospital organizations.

Furthermore, we assumed all NICUs operate in the same manner, which is not the case. Not all NICUs explicitly distinguish between IC, MC, and HC. In addition, the average length of stay may be different. There is also a difference in available specialist care, which means the patient mix is different as well. Some patients will need to be transferred to a different NICU for specific care, which is not incorporated into the models. It is unclear how much impact specialist care would have.

Validating our model by comparing our admission table to Perined's publication is difficult, since there are many uncertainties. For example, Perined's methodology of defining the origin region of a NICU admission might be different than ours. The most recent available data is also from 2015, of which we do not have the operational capacities at that time. When calculating transfers of patients, a one ore two beds fewer or more has a significant impact. In addition, our admission table shows the long term patient flow in the network, while available data shows the patient flow for just one year. This means randomness of arrivals, or busy periods, play a significant role. It is also known that NICUs struggle with fluctuating operational capacity, which may also influence the patient flow within a network.

6.3 Recommendations

We recommend the Neonatal Care Network to evaluate which hospitals are assigned to which NICUs as a tactical decision, for example once or twice a year. Doing this will result in a better match of current available capacity and demand, and keep transfers of patients to a minimum. Results and insights of this study could be used as an argument for the *logistical* aspect of the decisions.

We advise to run the models again with more accurate values for the parameters, to obtain a more realistic optimal assignment. These values for parameters could be NICU specific. The prioritization matrix can be modified to represent preferred collaborations between NICUs. In addition, constraints for forcing the assignment of certain hospitals to NICUs could be added. An additional benefit of this is that this also reduces the search space. A new objective function could be formulated as well, on the condition that its value can be calculated using the admission table.

And finally, it is key to note that performance targets such as the number of transfers should be related to the capacity of the network. For example, setting a target of 100 transfers a year is currently unrealistic for a network with 163 beds. Figure 5.8 shows the relationship between the number of transfers and capacity of the network.

6.4 Further research

This thesis could be interpreted as exploratory research and could serve as an introduction to a more elaborate research project in collaboration with all NICUs. This would allow for making a comprehensive overview of patient flows, in which a distinction between pregnant women and ill children should be made. In addition, higher quality data for input parameters could be obtained.

Some further research might be on the topic of where to locate NICUs and/or specialty care. In addition, more research on online operational decisions regarding transferring patients, depending on the state of the network, might prove useful. Real-time decisions could be perhaps be made using actual expected travel time, obtained from for example Google maps.

Furthermore, it might be interesting to investigate how travel times are distributed and if we can use this for modeling quality constraints for assigning hospitals. For example, that 95% of all patients must be admitted within 60 minutes.

6.5 Contribution to science

We formulated a CTMC model and a heuristic to analyze transfers within a network of queues by constructing an admission table from the steady state distribution. We have shown that in same formulation, different patient types can be included as well. These patient types may or may not have different capacity requirements (e.g. two beds for a twin). Instead for a network of NICUs, this model could of course also be used for wards within the same hospital (division). Our approach could be interpreted as an extension of the commonly used Erlang loss model, since both models operate under the same assumptions.

In addition, we also measured spatial accessibility to NIC using E2SFCA. To the best of our knowledge, a FCA method has also not been applied to a situation in the Netherlands.

Bibliography

Ahmadvand, S. and Pishvaee, M. (2018). An efficient method for kidney allocation problem: a credibility-based fuzzy common weights data envelopment analysis approach. *Health Care Management Science*, 21(4):587–603.

Andersen, A., Nielsen, B., and L.B., R. (2017). Optimization of hospital ward resources with patient relocation using markov chain modeling. *European Journal of Operational Research*, 260(3):1152 – 1163.

Bagheri, M., Gholinejad Devin, A., and Izanloo, A. (2016). An application of stochastic programming method for nurse scheduling problem in real word hospital. *Computers and Industrial Engineering*, 96:192–200.

Boesveld, I., Bruijnzeels, M., Hitzert, M., Hermus, M., Van Der Pal-De Bruin, K., Elske Van Den Akker-Van Marle, M., Steegers, E., Franx, A., De Vries, R., and Wiegers, T. (2017). Typology of birth centres in the netherlands using the rainbow model of integrated care: Results of the dutch birth centre study. *BMC Health Services Research*, 17(1).

Calovi, M. and Seghieri, C. (2018). Using a gis to support the spatial reorganization of outpatient care services delivery in italy. *BMC Health Services Research*, 18(1).

Cardoso, T., Oliveira, M., Barbosa-Póvoa, A., and Nickel, S. (2015). Introducing health gains in location-allocation models: A stochastic model for planning the delivery of long-term care. volume 616.

Carello, G. and Lanzarone, E. (2014). A cardinality-constrained robust model for the assignment problem in home care services. *European Journal of Operational Research*, 236(2):748–762.

CBS (2019). Birth; key figures. Retrieved from https://opendata.cbs.nl/.

Cheng, G., Zeng, X., Duan, L., Lu, X., Sun, H., Jiang, T., and Li, Y. (2016). Spatial difference analysis for accessibility to high level hospitals based on travel time in shenzhen, china. *Habitat International*, 53:485–494.

Daldoul, D., Nouaouri, I., Bouchriha, H., and Allaoui, H. (2018). A stochastic model to minimize patient waiting time in an emergency department. *Operations Research for Health Care*, 18:16–25.

de Bruin, A. M., Bekker, R., van Zanten, L., and Koole, G. M. (2010). Dimensioning hospital wards using the erlang loss model. *Annals of Operations Research*, 178(1):23–43.

de Laat, M. W., Wiegerinck, M. M., Walther, F. J., Boluyt, N., Mol, B. W., van der Post, J. A., van Lith, J. M., and Offringa, M. (2010). [practice guideline 'perinatal management of extremely preterm delivery']. *Ned Tijdschr Geneeskd*, 154:A2701.

Delamater, P., Shortridge, A., and Kilcoyne, R. (2019). Using floating catchment area (fca) metrics to predict health care utilization patterns 11 medical and health sciences 1117 public health and health services. *BMC Health Services Research*, 19(1).

Fisher, M. L., Jaikumar, R., and Wassenhove, L. N. V. (1986). A multiplier adjustment method for the generalized assignment problem. *Management Science*, 32(9):1095–1103.

GADM (2012). Gadm database of global administrative areas, version 2.0. [online]. URL: www.gadm.org.

Gilmour, S. (2010). Identification of hospital catchment areas using clustering: An example from the nhs. *Health Services Research*, 45(2):497–513.

Guagliardo, M. (2004). Spatial accessibility of primary care: Concepts, methods and challenges. *International Journal of Health Geographics*, 3.

Hans, E., Van Houdenhoven, M., and Hulshof, P. (2012). A framework for healthcare planning and control. *International Series in Operations Research and Management Science*, 168:303–320.

Hansen, P., Mladenović, N., Brimberg, J., and Pérez, J. A. M. (2019). Variable neighborhood search. pages 57–97.

Hoek, W. (2015). Acuity measurement at the neonatal intensive care unit.

Kadaster/CBS (2018). © kadaster / central bureau for statistics. Retrieved from https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische%20data/wijk-en-buurtkaart-2018.

Karamyar, F., Sadeghi, J., and Yazdi, M. (2018). A benders decomposition for the location-allocation and scheduling model in a healthcare system regarding robust optimization. *Neural Computing and Applications*, 29(10):873–886.

Kim, Y., Byon, Y.-J., and Yeo, H. (2018). Enhancing healthcare accessibility measurements using gis: A case study in seoul, korea. *PLoS ONE*, 13(2).

King, M., Ramnarayan, P., Seaton, S., and Pagel, C. (2019). Modelling the allocation of paediatric intensive care retrieval teams in england and wales. *Archives of Disease in Childhood*.

Klauss, G., Staub, L., Widmer, M., and Busato, A. (2005). Hospital service areas - a new tool for health care planning in switzerland. *BMC Health Services Research*, 5.

Koppka, L., Wiesche, L., Schacht, M., and Werners, B. (2018). Optimal distribution of operating hours over operating rooms using probabilities. *European Journal of Operational Research*, 267(3):1156–1171.

Liu, M., Zhang, L., and Zhang, Z. (2015). Optimal scheduling of logistical support for medical resources order and shipment in community health service centers. *Journal of Industrial Engineering and Management*, 8(5):1362–1379.

Luo, W. and Qi, Y. (2009). An enhanced two-step floating catchment area (e2sfca) method for measuring spatial accessibility to primary care physicians. *Health and Place*, 15(4):1100–1107.

Luo, W. and Wang, F. (2003). Measures of spatial accessibility to health care in a gis environment: Synthesis and a case study in the chicago region. *Environment and Planning B: Planning and Design*, 30(6):865–884.

Ma, L., Luo, N., Wan, T., Hu, C., and Peng, M. (2018). An improved healthcare accessibility measure considering the temporal dimension and population demand of different ages. *International Journal of Environmental Research and Public Health*, 15(11).

Murad, A. (2007). Creating a gis application for health services at jeddah city. *Computers in Biology and Medicine*, 37(6):879–889.

Nauss, R. (2003). Solving the generalized assignment problem: An optimizing and heuristic approach. *INFORMS Journal on Computing*, 15(3):249–266.

Otten, W. (2017). Pooling hospital beds : a capacity allocation study within the wilhelmina kinderziekenhuis.

Oude Weernink, A. (2018). The assessment of pooling intensive care and high care units at the neonatology department of wilhelmina kinderziekenhuis.

Pehlivan, C., Augusto, V., and Xie, X. (2014). Dynamic capacity planning and location of hierarchical service networks under service level constraints. *IEEE Transactions on Automation Science and Engineering*, 11(3):863–880.

Perined (2013a). Perinatale zorg in nederland 2011. www.perined.nl.

Perined (2013b). Perinatale zorg in nederland 2012. www.perined.nl.

Perined (2014). Perinatale zorg in nederland 2013. www.perined.nl.

Perined (2015). Perinatale zorg in nederland 2014. www.perined.nl.

Perined (2016). Perinatale zorg in nederland 2015. www.perined.nl.

Perined (2019). Perinatale zorg in nederland 2017. www.perined.nl.

Radke, J. and Mu, L. (2000). Spatial decompositions, modeling and mapping service regions to predict access to social programs. *Geographic Information Sciences*, 6(2):105–112.

Rocha, T., Da Silva, N., Amaral, P., Barbosa, A., Rocha, J., Alvares, V., De Almeida, D., Thumé, E., Thomaz, E., De Sousa Queiroz, R., De Souza, M., Lein, A., Lopes, D., Staton, C., Vissoci, J., and Facchini, L. (2017). Addressing geographic access barriers to emergency care services: A national ecologic study of hospitals in brazil. *International Journal for Equity in Health*, 16(1).

Ross, G. and Soland, R. M. (1977). Modeling facility location problems as generalized assignment problems. *Management Science*, 24(3):345–357.

Sadatasl, A., Zarandi, M., and Sadeghi, A. (2017). A combined facility location and network design model with multi-type of capacitated links and backup facility and non-deterministic demand by fuzzy logic.

Schuurman, N., Fiedler, R., Grzybowski, S., and Grund, D. (2006). Defining rational hospital catchments for non-urban areas based on travel-time. *International Journal of Health Geographics*, 5.

Shi, X., Alford-Teaster, J., Onega, T., and Wang, D. (2012). Spatial access and local demand for major cancer care facilities in the united states. *Annals of the Association of American Geographers*, 102(5):1125–1134.

Shin, K. and Lee, T. (2018). Improving the measurement of the korean emergency medical system's spatial accessibility. *Applied Geography*, 100:30–38.

Tang, J. and Wang, Y. (2015). An adjustable robust optimisation method for elective and emergency surgery capacity allocation with demand uncertainty. *International Journal of Production Research*, 53(24):7317–7328.

Tansley, G., Stewart, B., Zakariah, A., Boateng, E., Achena, C., Lewis, D., and Mock, C. (2016). Population-level spatial access to prehospital care by the national ambulance service in ghana. *Prehospital Emergency Care*, 20(6):768–775.

Vidyarthi, N. and Jayaswal, S. (2014). Efficient solution of a class of location-allocation problems with stochastic demand and congestion. *Computers and Operations Research*, 48:20–30.

Vieira, B., Demirtas, D., van de Kamer, J., Hans, E., and van Harten, W. (2018). A mathematical programming model for optimizing the staff allocation in radiotherapy under uncertain demand. *European Journal of Operational Research*, 270(2):709–722.

Wan, N., Zou, B., and Sternberg, T. (2012). A three-step floating catchment area method for analyzing spatial access to health services. *International Journal of Geographical Information Science*, 26(6):1073–1089. cited By 72.

Wang, S., Li, J., and Peng, C. (2017). Distributionally robust chance-constrained program surgery planning with downstream resource.

Wang, Y., Tang, J., and Fung, R. (2014). A column-generation-based heuristic algorithm for solving operating theater planning problem under stochastic demand and surgery cancellation risk. *International Journal of Production Economics*, 158:28–36.

Wolff, R. W. (1982). Poisson arrivals see time averages. volume 30, pages 223–231.

Xia, T., Song, X., Zhang, H., Song, X., Kanasugi, H., and Shibasaki, R. (2019). Measuring spatio-temporal accessibility to emergency medical services through big gps data. *Health and Place*, 56:53–62.

Xiao, Y., Kaku, I., Zhao, Q., and Zhang, R. (2011). A reduced variable neighborhood search algorithm for uncapacitated multilevel lot-sizing problems. *European Journal of Operational Research*, 214(2):223 – 231.

Xiong, X., Zhou, S., Huo, Z., and Luo, L. (2018). Gis-based method to delimit spheres of influence for a medical consortium: Experience of a pediatric medical consortium, shanghai. *International Journal of Health Planning and Management*.

Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8(3):338–353.

Zarrinpoor, N., Fallahnezhad, M., and Pishvaee, M. (2018). The design of a reliable and robust hierarchical health service network using an accelerated benders decomposition algorithm. *European Journal of Operational Research*, 265(3):1013–1032.

Öncan, T. (2007). A survey of the generalized assignment problem and its applications. *INFOR*, 45(3):123–141.

Appendix A

Root cause analysis

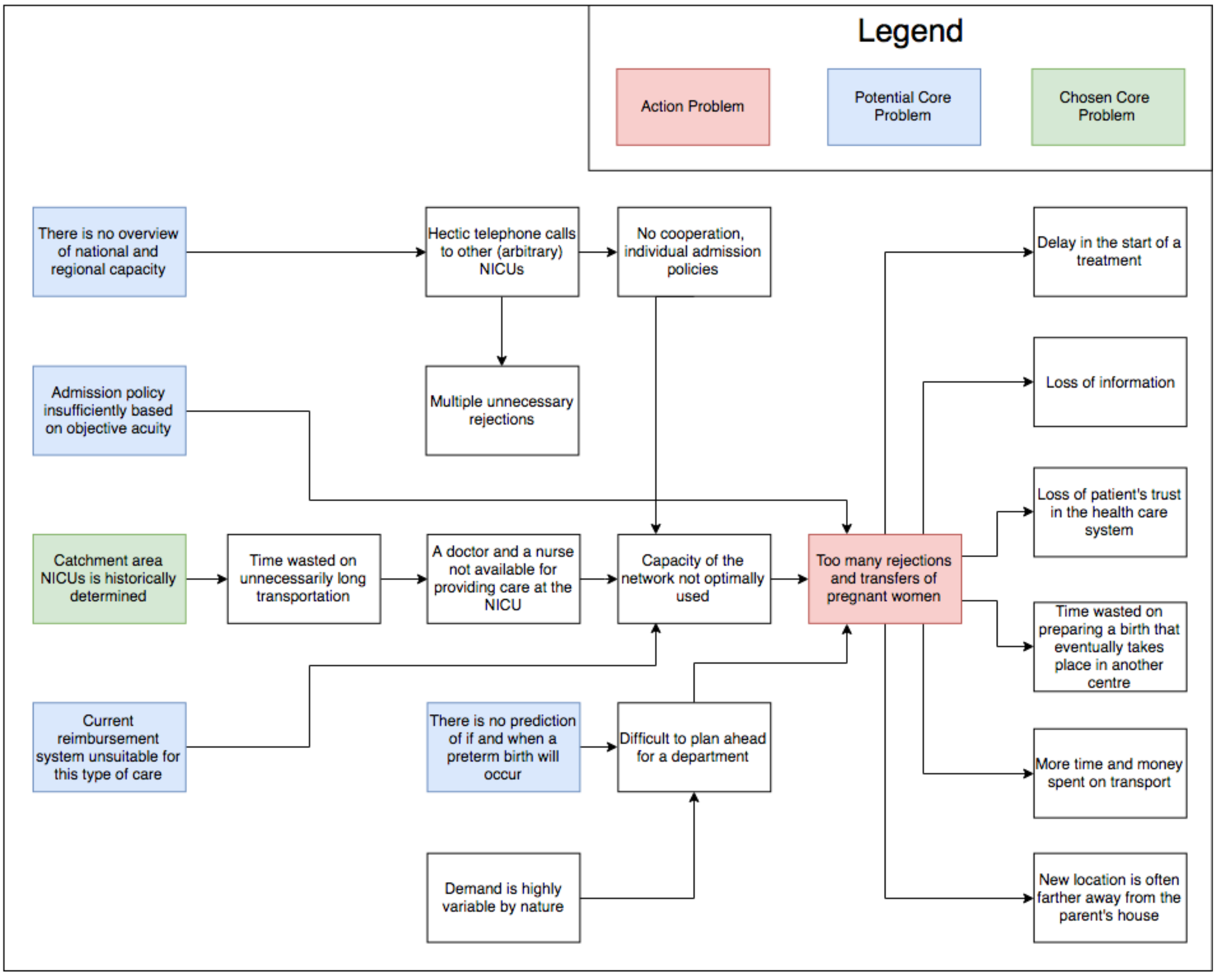

Figure A.1: Underlying root causes and the consequences of the experienced problem (action problem).

Appendix B

Literature search

B.1 Catchment area

The goal of this literature search is to find methods or models that are used to determine catchment areas of hospitals. Several possible synonyms of catchment area are included in the initial search string, such as *service area*. A catchment area might also refer to the place where water is collected, so this is excluded.

Figure B.1 gives an overview of the literature review. Table B.1 contains the selection criteria that were used for selecting articles. At least one of those selection criteria must be true in order for an article to be included or excluded.

Inclusion	**Exclusion**
Catchment area	Clinical studies
Travel time	Focus is purely on equity of accessibility
Emergency care	Non health care applications
GIS	

Table B.1: Selection criteria

B.2 Demand uncertainty

The goal of this literature search is to investigate how uncertainty of parameters can be included in mathematical programming, specifically focusing on demand and health care applications. Inventory models and multistage models are excluded in the search term.

Figure B.2 gives an overview of the literature review. Table B.2 contains the selection criteria that were used for selecting articles. At least one of those selection criteria must be true in order for an article to be included or excluded.

Inclusion	Exclusion
uncertainty	Traffic
allocation/assignment problems	Non health care applications
network	Literature reviews
robust	If no model is given

Table B.2: Selection criteria

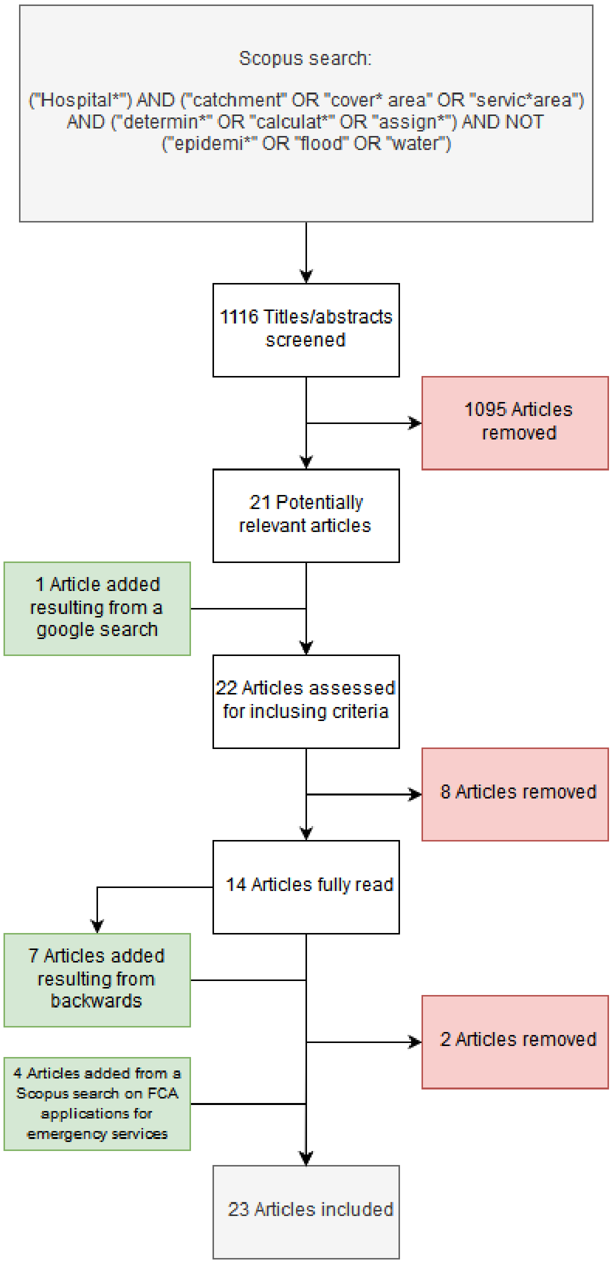

Figure B.1: Flow diagram of a literature search for methods or models for determining catchment areas.

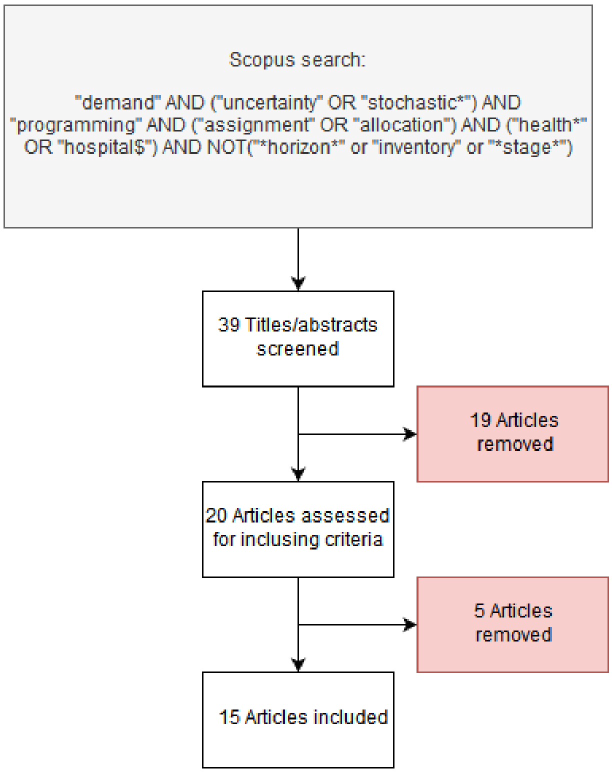

Figure B.2: Flow diagram of a literature search for including stochasticity of demand in mathematical programming, for a health care application.

Appendix C

CTMC multiple births

Using similar notation as Andersen et al. (2017), we include P patient types ($p \in \{1, 2, ..., P\}$) which represent the arrival of a birth of p siblings (e.g, type 1 is a single birth, type 2 a twin, etc.). Then we define w_p as the number of arrivals of type p currently admitted to the NICU. In case of a multiple birth, an arrival is seen as a *single* event for which p beds are required.

The state of the IC can be described as the number of arrivals admitted of each patient type, $s = (w_1^s, w_2^s, ..., w_P^s)$. The state space S is bounded by the available capacity c of the IC. The number of free beds in a state s can be defined as $f^s = c - \sum\limits_{p \in P} w_p^s \cdot p$. Therefore, the finite state space S is described by all unique states s for which $f^s \geq 0$. This Markov chain is irreducible.

We are interested in the fraction of time spent in each state and therefore want to find the steady-state distribution π. We can use two methods to obtain these values. The first method is a matrix formulation, while the second method is a derived formula.

For the first method, let us define Q as the transition rate matrix. In Q, we denote the rate in which the system moves from state $s \in S$ to a new state $s' \in S$, and define each element as $q_{ss'}$, with the following values:

$$q_{ss'} = \begin{cases} \lambda_p & \text{if } s' = \left(..., w_p^s + 1, ...\right), \forall p \in P \\ \mu_p \cdot w_p^s & \text{if } s' = \left(..., w_p^s - 1, ...\right), \forall p \in P \\ 0 & \text{otherwise} \end{cases}$$

Afterwards, all diagonal values in Q are set equal to the negative sum of its row, meaning $q_{ss} = -\sum\limits_{s' \neq s} q_{ss'} \quad \forall s \in S$. Now we can find the steady-state distribution π by solving $\pi Q = 0$. Since we have the additional constraint $\sum\limits_{s \in S} \pi_s = 1$, we replace one of column of Q with 1's.

For the second method, we derive a formula for the steady-state distribution π from the balance equations of neighboring states. For example, using $P = 2$ for sake of readability, the balance equations would be:

$$\begin{aligned} \pi_{w_1-1,w_2} \cdot \lambda_1 &= \pi_{w_1,w_2} \cdot w_1 \cdot \mu_1 \qquad \forall \quad \{s \in S | w_1 > 0\} \\ \pi_{w_1,w_2-1} \cdot \lambda_2 &= \pi_{w_1,w_2} \cdot w_2 \cdot \mu_2 \qquad \forall \quad \{s \in S | w_2 > 0\} \end{aligned}$$

We define $a_p = \frac{\lambda_p}{\mu_p}$ and rewrite these equations to the form:

$$\pi_{w_1 w_2} = \pi_{00} \prod_{p=1}^{P} \frac{a_p^{w_p}}{w_p!} \quad \text{with} \quad \pi_{00} = [1 + \sum_{s \in S \setminus (0,0)} \pi_{w_1 w_2}]^{-1}$$

These formulas work in similar fashion for any number of P patient types. The total number of rejected patients is calculated by $\sum_{p \in P} \sum_{\{s \in S | f^s < p\}} \pi_s \cdot p$.

Appendix D

Assignments

Hospital	**Current**	**Uncapacitated**	**Current_to_opt**	**Uncapacitated_to_opt**
Admiraal De Ruyter Ziekenhuis Goes	Rotterdam	Rotterdam	Rotterdam	Rotterdam
Albert Schweitzer Ziekenhuis Dordwijk	Rotterdam	Rotterdam	Rotterdam	Rotterdam
Alrijne Ziekenhuis Leiderdorp	Leiden	Leiden	Leiden	Leiden
Amphia Ziekenhuis Breda Langendijk	Rotterdam	Rotterdam	Rotterdam	Rotterdam
Amsterdam UMC	Amsterdam	Amsterdam	Amsterdam	Amsterdam
Antonius Ziekenhuis Sneek	Groningen	Zwolle	Groningen	Groningen
Beatrixziekenhuis	Rotterdam	Utrecht	Utrecht	Utrecht
BovenIJ Ziekenhuis	Amsterdam	Amsterdam	Amsterdam	Amsterdam
Bravis Ziekenhuis Bergen op Zoom	Rotterdam	Rotterdam	Rotterdam	Rotterdam
Canisius-Wilhelmina Ziekenhuis	Nijmegen	Nijmegen	Nijmegen	Nijmegen
Catharina Ziekenhuis	Veldhoven	Veldhoven	Veldhoven	Veldhoven
Deventer Ziekenhuis	Utrecht	Zwolle	Zwolle	Zwolle
Diakonessenhuis Utrecht	Utrecht	Utrecht	Utrecht	Utrecht
Dijklander Ziekenhuis locatie Hoorn	Amsterdam	Amsterdam	Amsterdam	Amsterdam
Elkerliek Ziekenhuis Helmond	Veldhoven	Veldhoven	Veldhoven	Veldhoven
Erasmus MC	Rotterdam	Rotterdam	Rotterdam	Rotterdam
ETZ Elisabeth	Utrecht	Veldhoven	Veldhoven	Veldhoven
Flevoziekenhuis	Amsterdam	Amsterdam	Amsterdam	Amsterdam
Franciscus Gasthuis	Rotterdam	Rotterdam	Rotterdam	Rotterdam
Franciscus Vlietland	Rotterdam	Rotterdam	Rotterdam	Rotterdam
Gelre Ziekenhuizen Apeldoorn	Utrecht	Zwolle	Zwolle	Zwolle
Gelre Ziekenhuizen Zutphen	Nijmegen	Nijmegen	Nijmegen	Nijmegen
Groene Hart Ziekenhuis Gouda	Leiden	Rotterdam	Rotterdam	Rotterdam
Haaglanden Medisch Centrum Westeinde	Leiden	Leiden	Leiden	Leiden
HagaZiekenhuis Leyweg	Leiden	Rotterdam	Leiden	Leiden
IJsselland Ziekenhuis	Rotterdam	Rotterdam	Rotterdam	Rotterdam
Ikazia Ziekenhuis	Rotterdam	Rotterdam	Rotterdam	Rotterdam
Isala Zwolle	Zwolle	Zwolle	Zwolle	Zwolle
Jeroen Bosch Ziekenhuis	Veldhoven	Veldhoven	Veldhoven	Veldhoven
LangeLand Ziekenhuis	Leiden	Leiden	Leiden	Leiden
Laurentius Ziekenhuis	Maastricht	Maastricht	Maastricht	Maastricht
Leids Universitair Medisch Centrum	Leiden	Leiden	Leiden	Leiden
Maasstad Ziekenhuis	Rotterdam	Rotterdam	Rotterdam	Rotterdam
Maastricht UMC+	Maastricht	Maastricht	Maastricht	Maastricht
Maasziekenhuis Pantein	Nijmegen	Nijmegen	Nijmegen	Nijmegen
Martini Ziekenhuis	Groningen	Groningen	Groningen	Groningen
Maxima Medisch Centrum Veldhoven	Veldhoven	Veldhoven	Veldhoven	Veldhoven
Meander Medisch Centrum	Utrecht	Utrecht	Utrecht	Utrecht
Medisch Centrum Leeuwarden	Groningen	Groningen	Groningen	Groningen
Medisch Spectrum Twente	Zwolle	Zwolle	Zwolle	Zwolle

Hospital	Current	Uncapacitated	Current_to_opt	Uncapacitated_to_opt
Noordwest Ziekenhuisgroep locatie Alkmaar	Amsterdam	Amsterdam	Amsterdam	Amsterdam
Noordwest Ziekenhuisgroep locatie Den Helder	Amsterdam	Amsterdam	Amsterdam	Groningen
Onze Lieve Vrouwe Gasthuis locatie Oost	Amsterdam	Amsterdam	Amsterdam	Amsterdam
Onze Lieve Vrouwe Gasthuis locatie West	Amsterdam	Amsterdam	Amsterdam	Amsterdam
Reinier de Graaf Gasthuis	Leiden	Rotterdam	Rotterdam	Rotterdam
Rijnstate	Nijmegen	Nijmegen	Nijmegen	Nijmegen
Rode Kruis Ziekenhuis	Amsterdam	Amsterdam	Amsterdam	Amsterdam
Ropcke Zweers Ziekenhuis	Zwolle	Zwolle	Zwolle	Zwolle
Scheper Emmen	Groningen	Groningen	Groningen	Groningen
Slingeland Ziekenhuis	Nijmegen	Nijmegen	Nijmegen	Nijmegen
Spaarne Gasthuis locatie Haarlem-Zuid	Amsterdam	Amsterdam	Amsterdam	Amsterdam
St. Anna Ziekenhuis Geldrop	Veldhoven	Veldhoven	Veldhoven	Veldhoven
St. Antonius Ziekenhuis Utrecht	Utrecht	Utrecht	Utrecht	Utrecht
St. Jans Gasthuis	Veldhoven	Veldhoven	Maastricht	Maastricht
Streekziekenhuis Koningin Beatrix	Nijmegen	Nijmegen	Zwolle	Nijmegen
Tergooi locatie Blaricum	Amsterdam	Amsterdam	Amsterdam	Amsterdam
UMC Groningen	Groningen	Groningen	Groningen	Groningen
UMC St. Radboud	Nijmegen	Nijmegen	Nijmegen	Nijmegen
Universitair Medisch Centrum Utrecht	Utrecht	Utrecht	Utrecht	Utrecht
Van Weel-Bethesda Ziekenhuis	Rotterdam	Rotterdam	Rotterdam	Rotterdam
VieCuri Medisch Centrum Venlo	Maastricht	Veldhoven	Maastricht	Maastricht
Wilhelmina Ziekenhuis Assen	Groningen	Groningen	Groningen	Groningen
Zaans Medisch Centrum	Amsterdam	Amsterdam	Amsterdam	Amsterdam
Ziekenhuis Amstelland	Amsterdam	Amsterdam	Amsterdam	Amsterdam
Ziekenhuis Bernhoven	Nijmegen	Veldhoven	Veldhoven	Veldhoven
Ziekenhuis De Gelderse Vallei locatie Ede	Nijmegen	Utrecht	Utrecht	Utrecht
Ziekenhuis de Tjongerschans	Groningen	Zwolle	Groningen	Groningen
Ziekenhuis Nij Smellinghe	Groningen	Groningen	Groningen	Groningen
Ziekenhuis Rivierenland	Utrecht	Utrecht	Nijmegen	Utrecht
Ziekenhuis St. Jansdal Harderwijk	Zwolle	Utrecht	Zwolle	Zwolle
Ziekenhuis St. Jansdal Lelystad	Zwolle	Amsterdam	Amsterdam	Amsterdam
Ziekenhuisgroep Twente Locatie Almelo	Zwolle	Zwolle	Zwolle	Zwolle
ZorgSaam De Honte	Rotterdam	Rotterdam	Maastricht	Maastricht
Zuyderland Medisch Centrum Heerlen	Maastricht	Maastricht	Maastricht	Maastricht