APRIL – JUNE 2011

EMERGENCY CALL CENTER

FINDING A BALANCE BETWEEN COSTS AND QUALITY OF SERVICE WHEN DEALING WITH EMERGENCY CALLS

Author: *John Puts*

Supervised by: *Rob van der Mei*
Geert Jan Kommer

rivm
National Institute for Public Health and the Environment

vrije Universiteit amsterdam

Preface

Writing a paper is part of the Master Business Mathematics and Informatics (BMI) at the VU University in Amsterdam. The purpose of this paper is to give the student the opportunity to gain experience in doing research on a topic of interest and use the techniques and knowledge that the student has obtained during his/her study. The topic has to concern subjects that are part of the BMI study field since it is part of the compulsory subjects of the BMI master. The student approaches a supervisor for the project that can provide the student with additional information on the topic, give guidance and maybe provide data that can be used in the research part of the project.

The topic of my paper is the most closely related to the "optimization of business processes" part of the BMI study. It regards analyzing an average emergency call center in The Netherlands in order to find out how emergency call centers perform and seek for ways to improve this performance.

This topic was proposed to me by my supervisors, Rob van der Mei and Geert Jan Kommer. They work on the optimization of ambulance services, and gave me the opportunity to undertake some analysis and do research about a topic that is part of their work.

Rob van der Mei works at the *Centre for Mathematics and Computer Science (CWI)*, where he is currently heading the research cluster Probability, Networks and Algorithms. He also has a part-time assignment as a full professor at the VU University Amsterdam, Faculty of Exact Sciences, Division of Mathematics. Geert Jan Kommer is a health care researcher, and is currently working as the manager of the health care research group at the *Rijksinstituut voor volksgezondheid en milieu (RIVM)*. I want to thank them for all their great help and enthusiasm that I experienced throughout the whole project period. They were always coming up with good ideas and provided guidance that kept me on the right track and eventually led to the successful completion on the project. I also would like to thank Sandjai Bhulai for his interest in my project, helpful input and his regular attendance at the meetings with my supervisors. Another person that I would like to thank is the manager of the emergency call center in Amsterdam, Rob Bosman. He made it possible for me to make a visit to the call center and see how things go in reality. Last I would like to thank Joost Bosman for his help on the simulation program Arena that I used for my simulations. He really helped me a lot with getting started with this program.

John Puts
- Amsterdam, April - June 2011

Abstract

This paper addresses the problem of cost-efficient capacity planning of call centers for emergency services. By thorough analysis and the use of historical data and expert guess, we were able to create a model that represents an average emergency call center in The Netherlands very well. By the use of simulation we checked if the current staffing method of an average emergency call center is appropriate, and we found out in which way this staffing method can be done even more efficient. We found that the current staffing method of an average call center in The Netherlands is quite good, but can be done better. We also found that when we exclude a certain part from the incoming calls of a non-urgent nature, we can improve the performance of the call center. Another situation that was analyzed is the situation in which we distinguish two type of call center operators, namely medical and logistic operators. We showed that when we decrease the number of medical operators in the call center by one, we need to staff two logistic operators instead, and provided explanation whether this is more efficient or not. In this paper we also described why the model that we made is a very good starting point for further analysis. With some small adjustments it is possible to find the optimal staffing method for each part of a day, for each day of the week. The model is also a good starting point for further analysis of all the emergency call centers in The Netherlands.

Table of Contents

1. Introduction

1.1 The emergency hotline

Every country in the world has it: An emergency telephone number. It allows a caller to contact local emergency services for assistance. The emergency telephone number may differ from country to country. It mostly is a three-digit number so people can remember it easily. 112 was introduced as a common emergency call number around 1990. It is also the number that is used in The Netherlands.

The emergency hotline plays a central role in our society. It is used to ring the alarm when there is an emergency going on. An emergency call is made to inform the emergency call center co-worker that one or more people need treatment. Mostly the purpose of a call of this medical nature is to ask for an ambulance as soon as possible. The "medical part" of the emergency hotline is a very important part of our health care system that concerns health problems of an urgent nature. Of course nobody wants to land up in a situation where (s)he needs the service of an ambulance. But when you do land up in this kind of situation where your life might depend on the quality of the emergency system, you expect this service to be as good as possible.

Whenever someone calls the emergency hotline, (s)he is connected to an emergency call center. The call center operator asks the caller whether the call concerns a fire-, police or medical issue. They then connect the call to the appropriate emergency service, which then dispatches the appropriate help. In the case of multiple services being needed on a call, the most urgent need must be determined, with other services being called in as needed.

When someone calls the emergency hotline, (s)he expects his call to be answered instantly. This is quite logical, because an emergency mostly requires direct action. Therefore you don´t want to run short on call center operators. But on the other hand, call center operators are costly, so you don´t want to have too much operators working in an emergency call center either. Therefore it is an interesting question how many operators you want to staff in an emergency call center, in order to find the optimal balance between service and costs. Therefore you have to define how fast your phone call has to be answered, and what is meant with 'fast'. This trade-off problem will be discussed in this paper, together with ideas about closely related problems. Our main question will be: *"What is the optimal number of emergency call center employees when at least 95% of all incoming A1 and A2 calls has to be answered within 6 seconds?".* The definition of A1 and A2 calls will be given in chapter 2. We will analyze this question by first describing the structure of the emergency call center as it is in The Netherlands. Afterwards this reality will be modeled in a mathematical way, so that we can use simulation to find interesting results on how the functioning of this system can be improved. We will only be concerned with the medical part of the emergency call center. This means we will not discuss police or fire issues. We will use a lot of emergency call center data that was provided by Geert Jan Kommer. We will use this data to represent an average emergency call center in The Netherlands. The values of the different parameters in our simulation model will be derived from this realistic data.

1.2 LITERATURE

1.2.1 GENERAL CALL CENTER LITERATURE

There have been many publications about call center studies. Some of them concern general call center problems that describe call centers in a very general way and provide mathematical models that can be used to analyze queuing problems. Other publications go more in to depth about call center problems that concern situations that are very specific and often related to a problem that is present in real life. These papers try to analyze this real life problem and use mathematical tools to find optimal strategies to show what strategy is optimal.

When we consider call center literature that is more closely related to the subject of this paper, namely ambulance emergency center problems, we can also find some publications. These contain analysis and interesting findings about many areas that fall in the same category as the subject that this paper addresses. Therefore it is interesting to mention them. Not only because they have provided useful information that helped in writing this paper, but also to give the reader a nice list of references that give the reader the possibility to learn more about this interesting subject.

First let us consider some publications that concern call center studies. A publication that talks about call center theory in general for instance is "*Call Center Mathematics*", KOOLE [7]. This publication provides a scientific method for understanding and improving call centers. It is a very general publication that is very suitable for those who know nothing or very little about call centers, but definitely could use some knowledge on this field. Think of managers that manage operator employees. These kind of people need to know what his workers are doing, so he can do his work better as well. The publication goes more in to detail about staffing, uncertainty of arrivals and service times, queuing models and related topics.

A more in-to-depth publication regarding staffing call centers is "*Staffing a call center with uncertain non-stationary arrival rate and flexibility*" LIAO ET AL. [6]. This publication analyzes the variability of the service level dependent on the length of the aggregation interval of arrivals. It also proposes a new approach for variability-controlled staffing. So for those who want to know more about staffing, this publication is recommended.

Another interesting publication that picks one part of the whole call center theory, and discusses this quite generally is called "*Queueing models of call centers: An introduction*", KOOLE ET AL. [5]. This work starts off with information about data analysis and forecasting, continues with explaining a couple of performance models. It also discusses workforce management and wraps up with a conclusion. It is very handy for those who want to get to know more about queuing models that are often used to analyze call centers.

1.2.2 Emergency Call Center Literature

For those who would like to know more about the static and dynamic ambulance deployment topics; *"Computational Methods for Static Allocation and Real-Time Redeployment of Ambulances*" is a very interesting work, RESTREPO [8]. It presents methods to finding solutions for the two optimization problems.

A very interesting work that talks about the same problem that this paper issues, but in a different country is called *"Simulation Modeling for Staff Optimization of the Toronto Emergency Medical Services Call Centre"*, GILLIAN ET AL. [9]. The goal of this paper is to determine the effectiveness of the current staffing routine as well as calculating the optimal number of call takers for any given hour of the week in the biggest emergency call center in Toronto.

Another publication that is concerned with the same topic, using time-series models instead, is *"The application of forecasting techniques to modeling emergency medical system calls in Calgary, Alberta"*, CHANNOUF ET AL. [10]. In this work, the authors try to develop and evaluate time-series models of call volume to the emergency medical service of a major Canadian city. Its objective is to offer simple and effective models that could be used for realistic simulation of the system and for forecasting daily and hourly call volumes. Of course, the emergency hotline system in Canada is organized quite different than it is in The Netherlands, which is logical since the geographical properties differ greatly. But still, it is similar in a way that an optimization problem is being solved, given many different factors, and often with similar or the same mathematical models that are used in our studies.

For those who would like to read about how the chaotic city of New York addresses the emergency call center optimization problem, we refer to the work called *"Simulation and Cost-Effectiveness Analysis of New York's Emergency Ambulance Service",* SAVAS [11]. The reference to this paper is included because it might be interesting to see how the emergency 911 call center in a unique city like no other is organized. What also makes it more interesting is the fact that this paper dates from the late 60´s, a year in which many mathematical optimization techniques that we know now were not developed yet.

A reference to these and more interesting literature can be found in the reference section at the end of this paper.

1.3 Structure of the Report

We will start with discussing the emergency call center more into depth, and provide clear descriptions of how the emergency system is organized in The Netherlands. This is necessary to understand the problem and be able to see how the results that follow later in this paper apply to this system. Of course, a sketch of the structure of the emergency call center functioning will also be provided to add some visualization.

In the next section, different mathematical models that apply to our model will be discussed together with some explaining why these models fit to the problem. This section makes the transition from the emergency call center system as it is in reality, to a more abstract simplified mathematical model, in order to be able to do some analysis on it and make calculations that can give us information about the functioning of the system. This information is very important to see what are the strong and weak sides of the system, which directly gives you an idea where your optimization can have the most effect.

Then, in chapter 4 we will devote a small part of the paper to a section that talks about simulation. Things like why we use simulation, what simulation program is used, why was this program chosen and what kind of simulations were made are described in this chapter.

Then in chapter 5 we come to the interesting part where the results of our simulation will be presented. This section contains the answer to the main question and sub questions that this paper tries to answer. A small part is also devoted to the discussion of questions that are less abstract but nevertheless very interesting.

The conclusion follows comes right after the results has been presented. It can be found in chapter 6.

The remaining chapters of the paper include a section about further research, an appendix and a bibliography.

2. The Emergency Call Center

2.1 General Info

The emergency call center can be seen as a room where a number of call center operators are awaiting emergency calls. The Netherlands is divided into many zones and each zone has its own emergency call center. In total there are 24 emergency call centers in The Netherlands, divided over the whole country in a way that is geographically efficient. Formally, there are 25 zones, but two of them, namely 'Agglomeratie Amsterdam' and 'Zaanstreek/Waterland' have been merged to one zone solely for EMS. For a visualization of all zones on a map of The Netherlands, we refer the reader to appendix A.2.

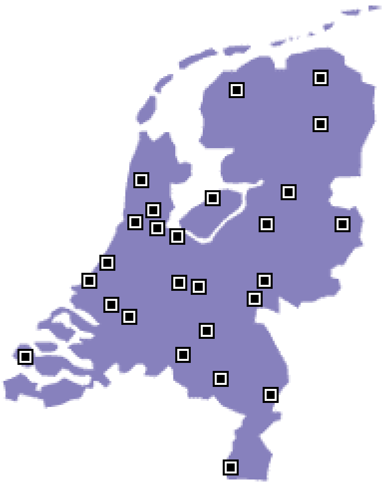


Figure 1 – Emergency Call Centers in The Netherlands

The division of the country in different zones has a reason. If there would be only one nationwide emergency call center where all emergencies would be reported, then it is highly probable that the operators of the call center would not possess any useful knowledge about the area where the emergency is taking place. It's not yet impossible to know all streets in the country. This useful knowledge could for instance be that the operator knows which ambulance station is closest to the place where the emergency is taking place. Another advantage of knowing a specific zone well is the ability to direct an ambulance in the smartest possible way so that it can reach the victim as soon as possible. On the other hand, more call center means higher costs as well, so finding a good balance here also is an interesting question that will be discussed deeper in chapter 7. In The Netherlands we have the following zones, AMBULANCEZORG NEDERLAND [13]:

Groningen	Midden Gelderland	Gooi- en Vechtstreek	Brabant Midden West
Friesland	Gelderland Zuid	Haaglanden	Brabant Noord
Drenthe	Utrecht	Hollands Midden	Brabant-Zuidoost
IJsselland	Noord-Holland Noord	Rotterdam-Rijnmond	Noord en Midden Limburg
Twente	Amsterdam/Waterland	Zuid-Holland Zuid	Zuid Limburg
Noordoost Gelderland	Kennemerland	Zeeland	Flevoland

Table 1 – Overview of the 24 zones in The Netherlands

Not only citizens can make an emergency alert. Think of a general practitioner that is on a visit at a patient and judges the patient needs an ambulance as soon as possible. This is an example of a situation in which a general practitioner, and not a citizen, is the one that makes an emergency alert at the closest emergency call center. The number the general practitioner dials when reporting the emergency is not 112 though. A general practitioner has the direct number to the closest emergency call center, so this call comes in at the same place as calls of a citizen. But there are more type of people that can report an emergency. We distinguish the following type of callers:

- Citizen
- General practitioner
- Police/fireman
- Hospital
- Health care institution

All of these type of callers, except for citizens, have the direct number to the emergency call center. We will discuss the different type of callers mentioned above in the next section of this chapter. Each emergency call center has a number of employees that work in the center, awaiting calls. When a call comes in, an operator has to find out whether it concerns a real emergency or not. This is done by going through a list of questions to determine whether the victim needs an ambulance or not (the triage process). Examples of calls that do not concern a real emergency are calls that are made by children pulling pranks or citizens in panic that over exaggerate a situation. When the emergency is serious though, a second question that has to be answered is whether the emergency requires urgent help or very urgent help. Very urgent help means an ambulance has to be on place as soon as possible. The nationwide standard for very urgent calls says that an ambulance has to arrive at the victim extremely fast. An urgent situation gives the ambulance slightly longer to get to the place of emergency, but still very fast. We distinguish three levels of urgency, namely:

- A1: Very urgent emergency, life threatening or possible invalidity, ambulance drives with ambulance light and sirens on, ambulance needs to be in place within 15 minutes
- A2: Urgent emergency, not life threatening but quick help is very wishful (situation could become life-threatening), ambulance needs to be in place within 30 minutes
- B: Ordered ambulance service, no haste necessary, ambulance needs to be in place at an on forehand agreed time

Ambulances are not parked at the emergency call center. Ambulance stations are chosen geographically smart over the zone that is covered by the emergency call center. Each of the 24 zones have around 7 to 11 ambulance stations, depending on how large (number of inhabitants, amount of hectare ground) the zone is. For a total overview of all ambulance stations per zone, we refer the reader to appendix A.1. When the operator comes to the conclusion an ambulance is required, he or she instantly tries to find out which ambulance station is nearest to the emergency location and directly alerts this station and provides information on where the ambulance needs to go, within how long it needs to be there, and what route it should take. If necessary, the operator stays on the line with the caller in order to provide medical assistance (i.e. how to reanimate). This is only necessary when a citizen calls, since general practitioners, police officers or employees of a health care institution know how this has to be done.

2.2 Type of Incoming Calls

As said before, emergencies are not only reported by citizens. There are more type of callers that report an emergency and request for an ambulance. In this section, we will describe the different type of callers and talk about their properties. Most call processes of the different types are very alike, nevertheless they also differ on various points. Priority is one of the many things that can vary between different type of calls. Also arrival rate and service time may vary. All of the different type of calls can be cut to several pieces, so we can distinguish different states phases in a call. These different states have transition probabilities, which also are vary between different type of calls. All of these parameters per call will be described in this section. In the next chapter, we will explain how we found values for these parameters.

2.2.1 Citizen calls

Whenever a citizen calls the emergency hotline by dialing the number “112” from a landline, the system recognizes in which zone the caller is situated and therefore the call is immediately directed to the emergency call center in that zone. In this way, time can be saved, and saving time when it comes to (life jeopardizing) emergencies is extremely important.

When a call is made from a cell phone though, the operator has to ask the caller where he is calling from first, before he can transfer the call to the emergency call center in the right zone. Another question that needs to be answered is whether the call regards a medical, a police or a fire issue. We will only consider the calls that concern a medical issue. So when we speak of incoming calls in the remainder of this paper, we mean calls that are directed to the emergency department, possibly also meaning an ambulance will be necessary.

When it is clear that the call concerns a medical emergency, we can distinguish a few states in which a conversation can be. Not all of these states necessarily have to occur. It depends on the situation, the medical knowledge of the caller and the nature of the emergency whether a call goes from one state in to another. We can describe the states of a conversation between a citizen and a call center operator as follows:

Call state	Priority	Description
1. Triage process	High	Operator figures out whether it concerns an A1 or A2 emergency by asking a few questions
2. Medical assistance	High	Operator explains how caller can provide medical assistance till ambulance arrives
3. (Temp) end of call	-	Operator is free again, but a coordination call between ambulance driver and hospital might take place later as soon as the patient is carried into the ambulance
4. Coordinating with hospital	Low	As soon as the ambulance arrives, medical help is provided and the patient is carried into the ambulance, the ambulance team might call the emergency call center to inform them about the fact that a patient is coming in soon

Table 2 – The 4 possible states that an emergency call can have

The call starts in the first state and has a certain probability of going to the next state. One could ask why we distinguish a ‘(temporary) end of call’ state. In this state, the operator that handled the first two states is free, but the emergency might need another call, which follows after state 3. So it belongs to the same emergency. Therefore we need to include it in to our model as well. State 3 is the time that an ambulance travels from its station to the patient plus the time that the ambulance crew is providing medical care to the patient. As soon as the patient gets carried in to the ambulance, the ambulance crew might need to call the emergency call center to inform them about the fact that a patient is coming in and tell what kind of preparations have to be made in the hospital so that the patient can be helped as soon as possible, and everything that is necessary for an eventual operation is in place. Because it is highly probable that we are dealing with an A1 or A2 urgency level, citizen calls have a high priority. We could visualize this process as follows:

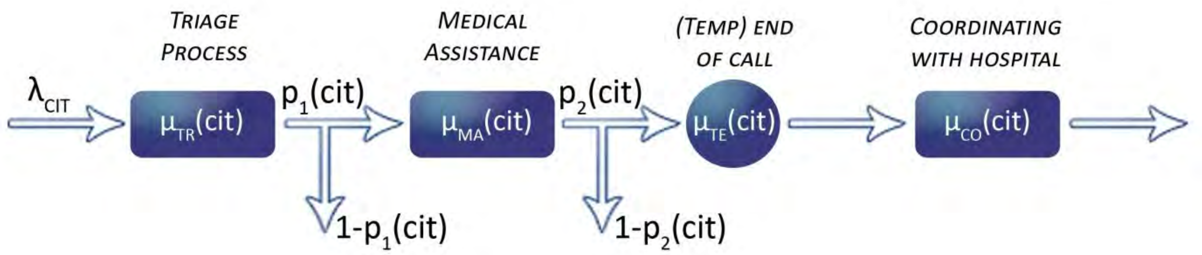


Figure 2 – The different citizen call states together with their corresponding transitions

In this visualization, the parameters have the following meaning:

P_1(cit)	=	The chance that the citizen call goes from the 'triage process' state into the 'medical assistance' state
P_2(cit)	=	The chance that the citizen call goes from the 'medical assistance state into the '(temp) end call' state
1 - P_1(cit)	=	The chance that the call ends for good after the 'triage process' state (so there will be no 'medical assistance' nor '(temp) end call' nor 'coordinating with hospital' state)
1 - P_2(cit)	=	The chance that the call ends for good after the 'medical assistance' state (so there will be no '(temp) end call' nor 'coordinating with hospital' state)
λ_{CIT}	=	The amount of incoming citizen calls per time unit.
μ_{TR}(cit)	=	The amount of time units that the triage process stage of the call takes
μ_{MA}(cit)	=	The amount of time units that the medical assistance stage of the call takes
μ_{TE}(cit)	=	The amount of time units that the (temp) end of call stage of the call takes
μ_{CO}(cit)	=	The amount of time units that the coordinating with hospital stage of the call takes

The values of these parameters and how we came to these values will be discussed more into depth in the next chapter.

2.2.2 General Practitioner Calls

A call between an operator and a general practitioner is somewhat similar to a call between an operator and a citizen. But there are some clear differences too. The biggest difference is the amount of medical knowledge a general practitioner has, weighted against the amount of knowledge a citizen possesses. Therefore, a general practitioner does not need to answer all the questions that a citizen answers in the triage process, since the general practitioner can make a better estimation of how serious a patient is hurt, and therefore how urgent the situation is. With this knowledge, he or she can immediately say whether the emergency is of type A1 or A2.

This gap in medical knowledge also is the reason that in the general practitioner call process, state 2 as described in table 2 will never occur. A general practitioner knows how to provide medical assistance to a patient until an ambulance arrives, so he doesn't have to stay on the line. Because it is highly probable that we are dealing with an A1 or A2 urgency level, also general practitioner calls have a high priority. The fourth type of state as described in table 2 also applies for general practitioner calls. This is quite obvious, since it can also occur that when the patient gets carried in to the ambulance, the ambulance crew might need to call the hospital to inform them about a patient coming in. With these similarities and differences, we come to a similar but yet slightly different model:

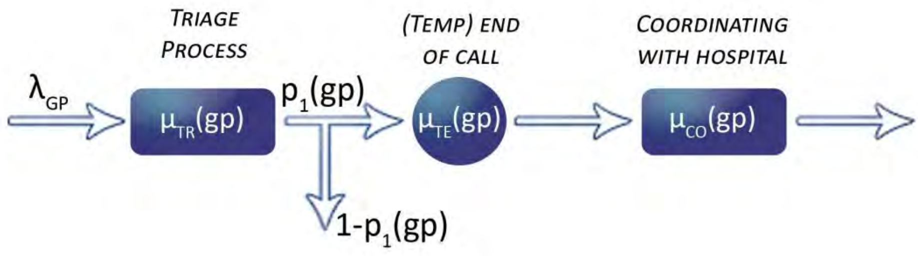


Figure 3 – The general practitioner call process

In this model we see that, as described above, state 2 is not present. Also the arrival rate and expected service time of a call differ from those at a citizen call, but this difference will be discussed more thoroughly in the next chapter. The parameters in this visualization could be described as follows:

$P_1(gp)$	=	The chance that the general practitioner call goes from the 'triage process' state into the 'medical assistance' state
$1 - P_1(gp)$	=	The chance that the call ends for good after the 'triage process' state (so there will be no '(temp) end call' nor 'coordinating with hospital' state)
λ_{GP}	=	The amount of incoming general practitioner calls per time unit.
$\mu_{TR}(gp)$	=	The amount of time units that the triage process stage of the call takes
$\mu_{TE}(gp)$	=	The amount of time units that the (temp) end of call stage of the call takes
$\mu_{CO}(gp)$	=	The amount of time units that the coordinating with hospital stage of the call takes

The values of these parameters and how we came to these values will be discussed more into depth in the next chapter.

2.2.3 Police calls

Police calls are very similar to general practitioner calls. This is because all police officers possess enough first aid knowledge to be able to provide medical assistance to a patient without needing help from an operator. This means that when dealing with police calls, we also do not take the second state as described in table 2 in to account. The fourth type of state as described in table 2 does applies for police calls though. This is quite obvious, since it can also occur that when the patient gets carried in to the ambulance, the ambulance crew might need to call the hospital to inform them about a patient coming in. The incoming arrival rate of police calls and the expected time a call takes differs from the general practitioner and citizen calls, but as said before we will go in to this more in to depth later. Because also here it is highly probable that we are dealing with an A1 or A2 urgency level, the priority of police calls is high as well. The visualization of a police call is presented next:

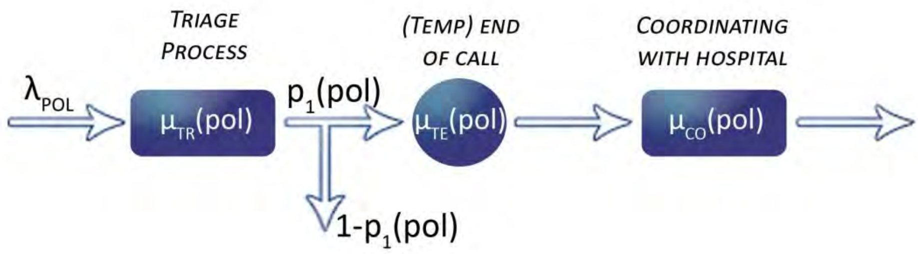


Figure 4 – The police call process

The parameters in this picture have the following meaning:

P_1(pol)	=	The chance that the police call goes from the 'triage process' state into the 'medical assistance' state
1 - P_1(pol)	=	The chance that the call ends for good after the 'triage process' state (so there will be no '(temp) end call' nor 'coordinating with hospital' state)
λ_{POL}	=	The amount of incoming police calls per time unit.
μ_{TR}(pol)	=	The amount of time units that the triage process stage of the call takes
μ_{TE}(pol)	=	The amount of time units that the (temp) end of call stage of the call takes
μ_{CO}(pol)	=	The amount of time units that the coordinating with hospital stage of the call takes

The values of these parameters and how we came to these values will be discussed more into depth in the next chapter.

2.2.3 HOSPITAL CALLS

Hospital calls are a bit different than police, general practitioner and citizen calls. When a hospital calls an emergency call center, it wants to schedule an 'ordered transport' ride. For instance, the hospital might have a patience that is done with his healing process at the hospital and needs to go home, but cannot be transported by a regular car. In this case, an ambulance has to come to the hospital to pick the patient up and bring him home. A hospital may also call the emergency call center to request for an ambulance at the patients address. In this case, the patient doesn´t need transport from the hospital to his home address, but vice versa. This description shows that these type of calls have an urgency level B, so they are not urgent, since these calls don't ask for immediate action. There is a lot of time between the moment the call takes place, and the moment at which the actual action happens. Therefore, hospital calls can be seen as low priority calls. Just like police and general practitioner calls, hospital calls do not include a medical assistance state. A hospital call could have a type 3 and type 4 state, as described in table 2, though. State 3 of a hospital call takes much more time than state 3 of a police, general practitioner or citizen call. State 3 is not the time in which an ambulance drives towards the place of emergency plus the time that the medical team provides treatment on location. State 3 of a hospital call is the time between the moment of scheduling of a patient transport and the moment of the actual patient pick-up. This could take hours. Within short theme, these type of calls will occur less, for instance because there are plans that patient transfer planning will be done through the internet. The above description brings us to the following visualization of the hospital call process:

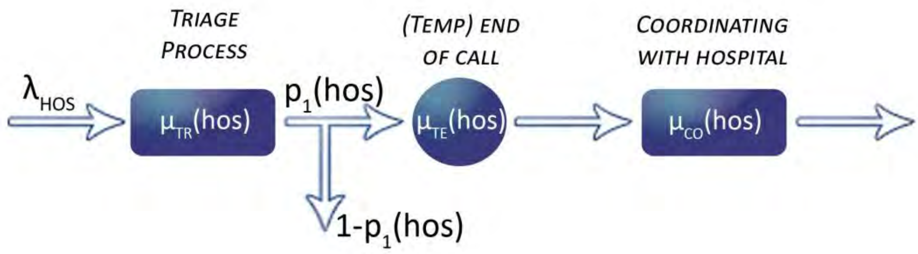


Figure 5 – The hospital call process

We practically see the same process as the police and general practitioner process. But as described above, the amount of time the call could be in the '(temp) end of call' state is much bigger here. The parameters in figure 5 can be described as follows:

P_1(hos)	=	The chance that the hospital call goes from the 'triage process' state into the 'medical assistance' state
1 - P_1(hos)	=	The chance that the call ends for good after the 'triage process' state (so there will be no '(temp) end call' nor 'coordinating with hospital' state)
λ_{HOS}	=	The amount of incoming hospital calls per time unit.
μ_{TR}(hos)	=	The amount of time units that the triage process stage of the call takes
μ_{TE}(hos)	=	The amount of time units that the (temp) end of call stage of the call takes
μ_{CO}(hos)	=	The amount of time units that the coordinating with hospital stage of the call takes

The values of these parameters and how we came to these values will be discussed more into depth in the next chapter.

2.2.3 Health care institution calls

Calls from health care institutions are practically the same as hospital calls. These calls are also made in order to schedule a patient transport from or to the health care institution. Since this type of call has the same properties as the hospital call, we observe the same process (with of course different parameters values):

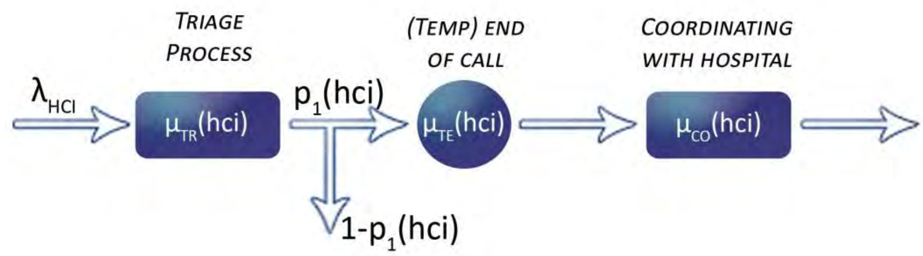


Figure 6 – The health care institution call process

The parameters in this picture have the following meaning:

P_1(hci)	=	The chance that the health care institution call goes from the 'triage process' state into the 'medical assistance' state
1 - P_1(hci)	=	The chance that the call ends for good after the 'triage process' state (so there will be no '(temp) end call' nor 'coordinating with hospital' state)
λ_{HCI}	=	The amount of incoming health care institution calls per time unit.
μ_{TR}(hci)	=	The amount of time units that the triage process stage of the call takes
μ_{TE}(hci)	=	The amount of time units that the (temp) end of call stage of the call takes
μ_{CO}hcil)	=	The amount of time units that the coordinating with hospital stage of the call takes

The values of these parameters and how we came to these values will be discussed more into depth in the next chapter.

For a confirmation of the priorities per type of call, we refer the reader to Appendix A.3. Here an overview of the fractions of all type of calls with their corresponding priorities is given. This shows us how many percent of each type of call concerned an (non) urgent emergency.

3. Mathematical models and assumptions

3.1 The call arrival process

In this section we will go deeper into the arrival process of the incoming calls. We will explain which distribution is the most suitable to use for this process. We also will show what kind of data we used to find the right values for the different parameters in our model, to be able to build a realistic simulation model.

The call arrival process is assumed Poisson. Almost all calls that come in to call centers are Poisson distributed, so it is quite obvious we use the same distribution here. But a good check that can confirm our assumption is important. We will make a histogram of a large number of incoming calls data. For this we use a data set of all incoming emergency calls at the emergency call center in The Netherlands. This data comes from the RIVM, and it is from the year 2009. From this data we can extract the incoming calls at an average emergency call center in The Netherlands, and then easily make a histogram that hopefully will confirm out assumption that the Poisson distribution is the most suitable distribution for incoming calls at an average emergency call center in The Netherlands. We make a histogram of all the calls that were made in the year 2009, sorted per day. We add the fitted Poisson distribution line to it:

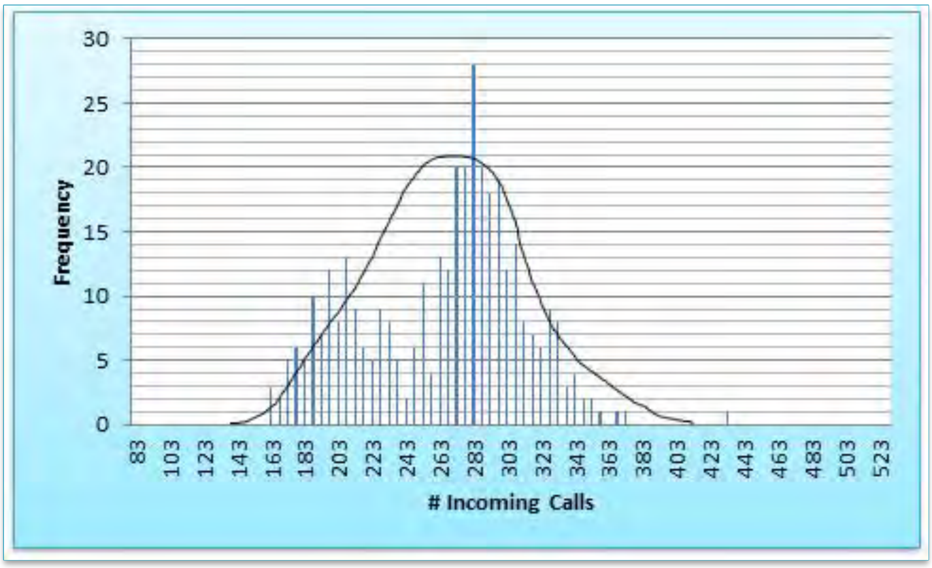
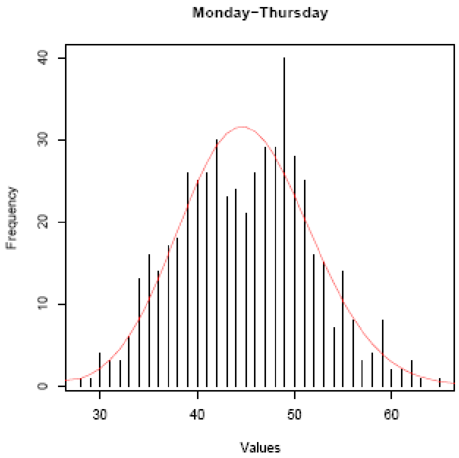
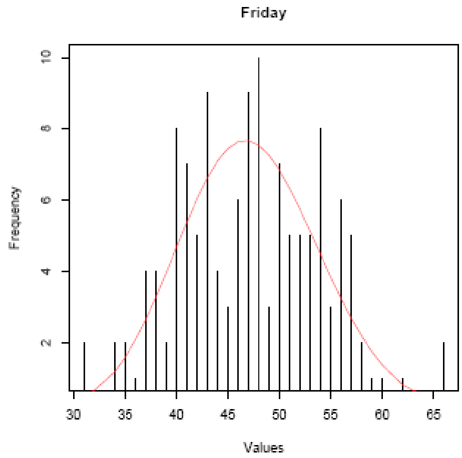


Figure 7 – Histogram of incoming calls at an average call center in The Netherlands

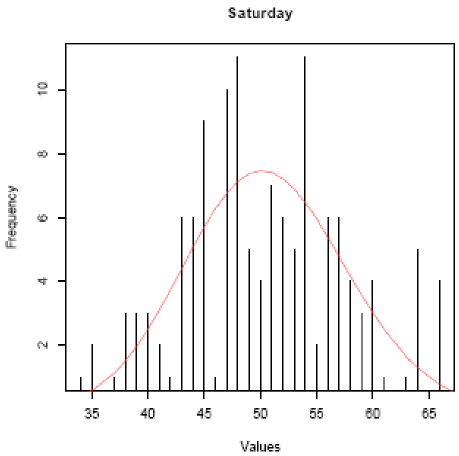
Figure 7 shows that the Poisson distribution is a good approximation for the incoming calls at an average emergency call center in The Netherlands. The two bumps can be explained by the fact that all days of the week were taken together. So for our data, we can use this assumption. Earlier research, CALINESCU. [12], on a large data set of incoming calls at an emergency call center confirms that these type of arrivals can be well represented by a Poisson distribution. On a set of incoming calls from November 1st 2005 to May 31st 2008, a histogram of the frequencies of calls with their fitted Poisson distribution was made:



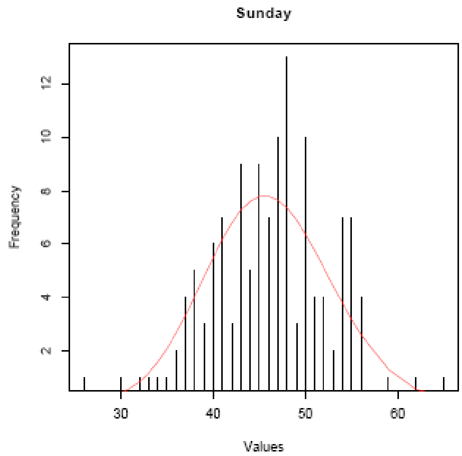
(a) *Set 1*: Monday-Thursday



(b) *Set 2*: Friday



(c) *Set 3*: Saturday



(d) *Set 4*: Sunday

Figure 8 – Call arrival analysis by *[Calinescu, M. (2009)]*

From the same data, we can make an assumption of the average number of calls that come in each day at an average emergency call center in The Netherlands. The total number of calls that were made in 2009 to an average emergency call center in The Netherlands was 96175. This gives us an average of 263 calls per day. These are the calls that led to an ambulance drive. There are also calls that do not lead to an ambulance drive. Think of children that dial the emergency call center as a joke, or people that call in an emergency that was already called in. From expert guess, we know that 17% of all calls is like this, therefore the actual number of average calls per day is 263/0,83 = 317 calls. Of course, on some days like Fridays and Saturdays, more emergencies take place which makes the number of calls on those days higher too. But for our general model, we will consider an average of the yearly number of calls. The model can be easily adjusted by taking the average number of calls on only one specific weekday though.

Then, another thing we need to know is the proportion of all the different type of calls. This tells us what fraction of the 317 calls that are made each day comes from citizens, how many come from general practitioners, and so on. Again, we use data from the RIVM of the year 2009 here. To calculate these proportions, we use the following table that includes this data:

Type of caller	Urgency	Number of rides	Fraction (%)	Cumulative (%)
112	A	303370	25.3	25.3
General practitioner	A	203263	16.9	42.2
Health care institution	B	180078	15.0	57.2
Unknown/not filled in	A	117798	9.8	67.1
General practitioner	B	94320	7.9	74.9
Unknown/not filled in	B	74282	6.2	81.1
Hospital	A	57159	4.8	85.9
Other	A	39021	3.3	89.1
Other	B	34334	2.9	92.0
Police/firemen	A	27289	2.3	94.3
Health care institution	A	25657	2.1	96.4
Hospital	B	25363	2.1	98.5
112	B	10053	0.8	99.4
Police/firemen	B	923	0.1	100.0

Table 3 – All type of calls with their urgency level and number of rides

From the above table, we can derive the fractions of each type of call. One assumption we make in this process, is that 80% of the calls that are made by a citizen, are made from a mobile phone. This means that 20% of the citizen calls is made from a landline. Since there was no data about this percentage, we trust on expert-guess. The proportions we find are:

Citizen (mobile call) = 20.88 %
Citizen (landline call) = 5.22 %
General practitioner call = 24.8 %
Police call = 2.4 %
Hospital call = 6.9 %
Health care institution call = 17.1 %
(total 77.3 %)

When we scale these values in such a way that the total of all fractions is 100%, with proportions being the same as above, we come to the following values:

Citizen (mobile call) = 27.02 %
Citizen (landline call) = 6.75 %
General practitioner call = 32.08 %
Police call = 3.10 %
Hospital call = 8.93 %
Health care institution call = 22.12 %
(total 100.0 %)

With an average number of calls per day of 317 and the above calculated fractions, we come to the following table:

Type of call	Fraction of total calls	Number of calls per day	Arrival rate per hour (λ)
Citizen mobile	27%	86	3.60
Citizen landline	7%	22	0.90
General practitioner	32%	103	4.28
Police	3%	10	0.41
Hospital	9%	29	1.19
Health care institution	22%	71	2.95

Table 4 – Overview of all type of calls with their corresponding λ values

The above table gives us the parameters we need for the arrival part of the mathematical model, that we will eventually convert to a simulation model. A more expanded overview of these and more parameters of our model can be found in Appendix A.4.

3.2 The Service Time Process

3.2.1 Service time of call states

The service time process is assumed exponential (expert guess). Since the service may exist of more than one different states, we have to assign an expected service time to each of the states, so we can put them in the simulation model separately too. Therefore we will have an expected service time for each of the different states in a phone call for each type of call. We will also describe how we got to these values. Most values come from data from emergency call centers in The Netherlands, but some values which are simply not available from data are estimated by the use of expert guess, which is not as reliable as historical data, but still very accurate. We still distinguish the 4 different type of call states as described in table 2.

First we will show what the expected service time is of the first state for all call types, and explain how we found this value. The data that provides the service times comes from the RIVM from the year 2009. We derive our values from the following table that includes information about the average 'triage time' lengths of all calls of each type in 2009:

Type of call	Amount of seconds that call is in 'triage process' state per urgency level	
	A1	A2
112	121.2	273.5
From ambulance	48.5	98.5
General practitioner	114.9	261.6
Hospital	117.2	243.6
Other	120.5	312.3
Police	97.2	200.6
Health care institution	191.3	518.4

Table 5 – Overview of expected service time of 'Triage Process' state for each type of call

So from the average of the amount of time the 'triage process' takes, we take the fraction of urgency level type A1 calls times the expected average time of A1 'triage process' calls plus the fraction of urgency level type A2 calls times the expected average time of A2 ‘triage process’ calls. This can be presented in the following formula:

Expected time that call is in ‘triage process’ state =
fraction A1 calls * A1 ‘triage process’ time + fraction A2 calls * A2 ‘triage process’ time

From historical data, AMBULANCEZORG NEDERLAND [13], we have the following A1 and A2 fractions in 2009:

fraction of A1 emergencies: 65,47%
fraction of A2 emergencies: 34,53%

With these formula and the above percentages, it is easy to calculate the average ‘triage process’ time for each type of call. These calculated values will be pasted in a table later in this section. There is one extra fact we may not forget here. Before a citizen that calls from a mobile phone gets the emergency call center operator of the right region on the line, he or she needs to tell the nationwide call center where the emergency is taking place. The reason for this is that unlike landline calls, it is not possible to know where a mobile call comes from, and therefore that question is asked first by the nationwide call center.

As regards the second state, only incoming calls from citizens can come in this state. Since there is no data available of the expected time that medical treatment on the place of emergency takes, we use expert guess to give an expected value to this parameter. The expected time that citizens receive treatment information of the emergency call center operator by phone is 8 minutes.

Then, if a call comes in the third state, in which the operator is free, but the emergency might be the reason for another call (state 4), we can also need to insert an expected value for this parameter. This state is equal to the time that passes from the moment the ‘triage process’ (or the medical treatment’ for citizens) state ends, till the moment the patient is carried in to the ambulance. So we use the average driving time and the average treatment time on place to make an estimation of this value. For the average driving time of the ambulance and the average treatment time on place, we can look at historical data, RIVM [14].

Average ambulance drive time for A1 emergencies in 2009: 9:44 minutes
Average ambulance drive time for A2 emergencies in 2009: 16:15 minutes

Average treatment time on place 19:32 minutes

By the use of the fractions of A1 and A2 emergencies showed earlier and the above numbers, we can calculate an average time for the third ‘(temp) end of call’ state. As regards the time that hospital or health care institution calls are in state 3, we are talking about a much longer time. As described in section 2.2.3 and 2.2.4 the time between a scheduled patient pick up and the actual pick up can take hours. Therefore we use expert guess to estimate this value. Fortunately, the influence of this situation is minimal, since it concerns a call that goes into a low priority queue, and this situation only occurs a few times a day because of the long interval time. These values for each type of call will be pasted in a table later on in this section.

The average time that a call can be in state 4 is also estimated by the use of expert guess, since there is no data available for this value. The average time state 4 takes is about the same for each type of calls, namely 1,5 minute. Given the description of the service time of each type of call above, we can create the following table:

Type of call	**Amount of minutes that call is in state**			
	State 1 *Triage process*	*State 2* *Medical assistance*	*State 3* *(Temp) end of call*	*State 4* *Coord. with hospital*
Citizen mobile	2.90	8.000	23.29	1.50
Citizen landline	2.90	8.000	23.29	1.50
General practitioner	2.76	State does not apply	31.29	1.50
Police	2.21	State does not apply	31.29	1.50
Hospital	2.68	State does not apply	180.00	1.50
Health care institution	5.07	State does not apply	180.00	1.50

Table 6 – Overview of expected service time of all states for each type of call

As mentioned before, the values of state 3 for hospital and health care institution calls are very high. This is because there is always a long time between the moment of scheduling and the actual pick up of a patient. But since these values hardly influence our model, we can include them without any problems.

3.2.2 Transitions between call states

As described in section 2.2, there are a couple of probabilities that decide whether a call goes from one state to another state, or that the call is over. These probabilities also need to gain a value to complete our model, and be able to make simulations afterwards.

From expert guess of Rob Bosman, employee of the emergency call center in Amsterdam, we know that of every 4 calls of a citizen, state 2 (the 'Medical assistance' state) occurs once. Therefore the probability that the call will go from the 'Triage process' state to the 'Medical assistance' state is 25% for all citizen calls.

Also from expert guess of Rob Bosman we learn that of every 9 calls of any type of call, a 'Coordination with hospital' will need to take place. Therefore the chance that a call will go from the 'Medical assistance' state (for citizens) or from the 'Triage process' state (for all other type of calls) to the '(Temp) end of call' state is 11,11%.

Type of call	**Transition probability from one state into another**		
	'Triage Process' to 'Medical Assistance'	'Triage Process' to '(Temp) end of call'	'Medical Assistance' to '(Temp) end of call'
Citizen mobile	25%	State does not apply	11.11%
Citizen landline	25%	State does not apply	11.11%
General practitioner	State does not apply	11.11%	State does not apply
Police	State does not apply	11.11%	State does not apply
Hospital	State does not apply	11.11%	State does not apply
Health care institution	State does not apply	11.11%	State does not apply

Table 7 – Values of all possible transitions in our model

3.3 THE ENTIRE MODEL

Now that we have described all the separate parts of our model, we can form a better idea about the entire model. Therefore, in this section the entire model is presented. We will call this model the ´Basic Emergency Call Center Model´. The reason for the word ´basic´ in the name is that we will make some variations in the model later on to see how these variations influence our performance. The basic model will be presented as a global visualization of the whole emergency call center process, that was cut into parts in the previous sections:

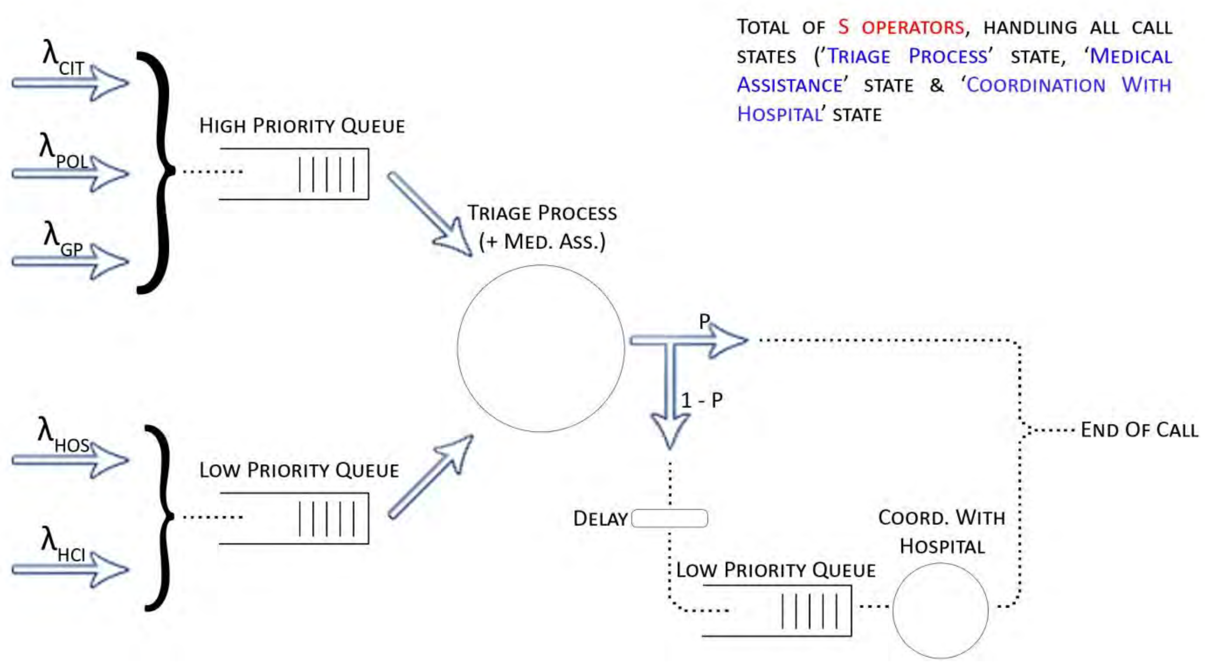


Figure 9 – The basic model of the Emergency Call Center

4. Simulation

4.1 Simulation to mimic reality

Simulation is a very important part of this paper. It is the technique we use to convert reality into a model and mimic reality with it. We use the simulation program 'Rockwell Arena' (version 13.5), a very well-functioning software pack that gives us the possibility to process all variables that were described in chapter 3 and create a model out of it. We chose simulation to model the emergency call center because simple queuing models are not very appropriate for this problem. The reason for this is the large number of separate incoming calls of different types. The call process differs for each type of call as well, and is not a single call process, but exists of 4 states that do not all have to occur. These states are described earlier in the paper (table 2). This incidentals make the emergency call center problem too complex for the use of queuing models. Arena gives us the possibility to include as many queues as we want and decompose the service process of each type of call into parts, which is exactly what we need for this problem. Also Arena gives its user a lot of simulation options. The user can define how long the simulation has to run, and how many iterations have to be simulated. Arena also measures a lot of performance indicators like (average) queue length and (average) waiting times, statistics that are very welcome in our project. But the options of Arena even go beyond that; it gives the user the possibility to define its own performance indicators with the use of some programming knowledge. This makes it possible for us to, for instance, include our main performance indicator (how many % of the incoming callers have to wait longer than 6 seconds?).

These are the main reasons why we chose simulation with Arena to mimic the problem that is analyzed in this paper. We created a quite complex model that is a very good representation of the emergency call center. To give the reader an idea of how our simulation model looks like in Arena, we will present a visualization of it:

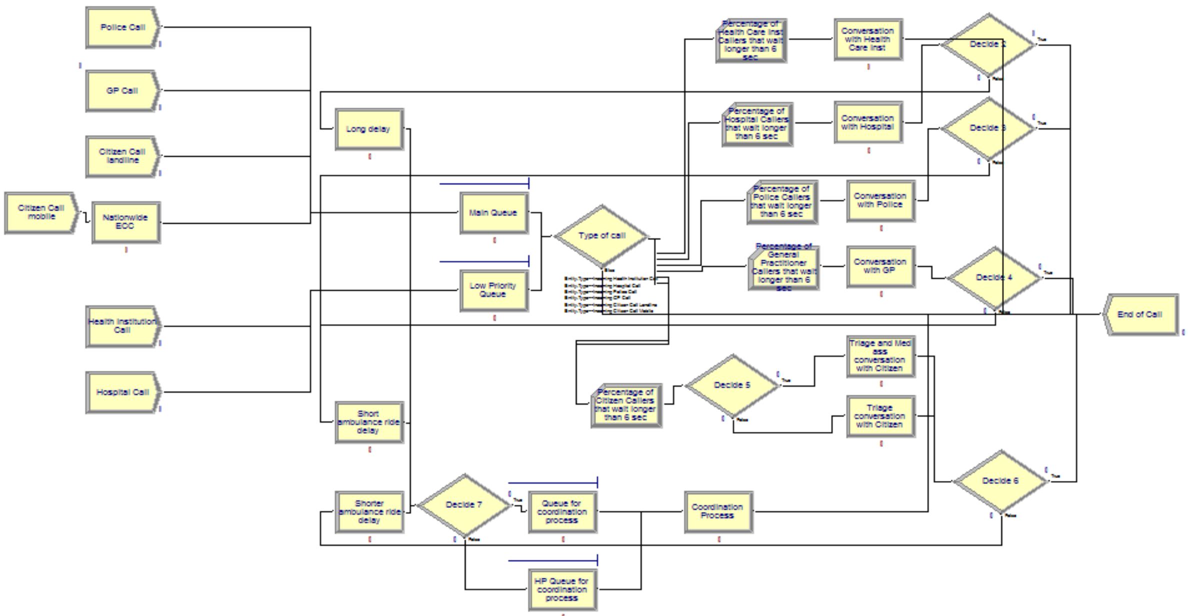


Figure 10 – Our simulation model in Arena

4.2 Simulation Length

In this part we will go a bit more in to depth about the properties of our simulation. Facts like simulation length and number of iterations will be described here. Also an explanation of our choices will be included in this section.

We will simulate for a length of one day, in other words 24 hours. We choose this simulation length because it is the exact length of one day. So 24 hours of simulation tells us how the call center performs on a random day, when an average of 317 calls come in at the emergency call center. An explanation of how we found this number is described earlier in this paper (section 3.1).

But simulating for only 24 hours and presenting the results is somewhat short, and not that reliable. Therefore we will do multiple iterations and take the average of all iterations. The number of iterations that we will simulate depends on when the average over all iterations does not really change anymore when we would add more iterations. So we will present some simulation results with a simulation length of 24 hours and varying numbers of iterations.

We will assume that there are 3 operators working in the call center, and an average of 317 calls arrive per day. This is just an assumption that we will use to find out how many iterations we need to use such that our results will be reliable enough.

Type of Call	% of people that waits longer than 6 sec				
	10 days Simulation	50 days Simulation	100 days Simulation	200 days Simulation	365 days Simulation
Citizen Call	9,35%	8,02%	8,19%	8,35%	8,39%
General Practitioner Call	6,66%	6,77%	6,52%	6,66%	6,61%
Health Care Institution	7,89%	7,82%	8,01%	8,18%	8,20%
Hospital Call	9,39%	7,76%	7,25%	6,62%	6,63%
Police Call	5,05%	5,68%	5,22%	5,34%	5,33%
AVERAGE	**7,67%**	**7,21%**	**7,04%**	**7,03%**	**7,03%**

Table 8 – Simulation results with varying numbers of iterations

We see that when we use more than 100 days of simulation, the average value of our main performance indicator does not really change a lot anymore. This means that it is reliable to use about 100 iterations (100 * 24 hours) or more. Let's check this by looking at the 95% confidence interval of our performance indicator (7.0421%) with this number of iterations:

$$U = P \pm 2 \sqrt{\frac{P(1-P)}{\#\ \text{Sim Minutes}}}$$

95% confidence interval with 100 iterations: [6.91, 7.18] (%)

This looks pretty decent. Let's also look at the confidence interval when we use 365 iterations:

95% confidence interval with 365 iterations: [6.97, 7.11] (%)

This 95% confidence interval is even smaller. Even though 100 iterations is very reliable, we will use 365 iterations, to be absolutely sure our results are reliable, and because of the fact that simulating one whole year is a nice number. It takes about 1 minute to simulate 100 iterations, and it takes about 3.5 minutes to simulate 365 iterations. This confirms it is not too time consuming to choose this number of iterations.

5. Results

5.1 The Basic Model

In the part we will show the performance of the basic model. As explained earlier in this paper, our main question is *"What is the optimal number of emergency call center employees when at least 95% of all incoming A1 and A2 calls has to be answered within 6 seconds?".* Besides this main performance indicator, we are also interested in other performance indicators; think of average waiting time per entity. The previous chapter explained how and why we used simulation. Now we will show the results that our simulation model provided us. Let's first look at the performance of the model when 5 operators are working in the call center. This is the average number of operators that currently work in an average emergency call center in The Netherlands. So we will get an idea of the current performance:

# of operators: 5
# of daily calls (average): 317
# of simulation days: 365

Type of Call	% of people that waits longer than 6 sec
Citizen Call	0.89%
General Practitioner Call	0.32%
Police Call	0.17%
AVERAGE	**0.46%**

Table 9 – Performance of the basic model with 5 operators working

We see that with 5 operators staffed, we meet our service requirement very easily. Not even 0.5% of all incoming urgent calls waits longer than 6 seconds before being answered. This means that 5 operators gives great service, but the question is: Can we still guarantee good service (meeting our requirement) with less than 5 operators? This would allow us to save on the number of staffed operators while not losing on service. Let's see how the model performs with 4 operators:

# of operators: 4
# of daily calls (average): 317
# of simulation days: 365

Type of Call	% of people that waits longer than 6 sec
Citizen Call	1.83%
General Practitioner Call	0.79%
Police Call	0.99%
AVERAGE	**1.60%**

Table 10 – Performance of the basic model with 4 operators working

We see that the percentage of people that has to wait longer than 6 seconds for service is still lower than 5%, so we meet our requirement with 4 staffed operators. Do we still meet it with less operators? Let's see how the model performs with 3 operators:

# of operators: 3
# of daily calls (average): 317
# of simulation days: 365

Type of Call	% of people that waits longer than 6 sec
Citizen Call	8.40%
General Practitioner Call	6.81%
Police Call	5.93%
AVERAGE	**7.05%**

Table 11 – Performance of the basic model with 3 operators working

We see that with 3 operators, the model does not meet our service requirement anymore. These results show that 3 operators in the call center is just not enough. It doesn't make sense to check the model with less than 3 operators since 3 is already not enough, therefore we know the optimal number of operators that should be staffed in an average call center in The Netherlands is 4.

Another interesting performance indicator is the average waiting time of each type of entity. This indicator tells us how long callers actually have to wait on average before being served. The waiting times are displayed in the following table:

Type of Call	Average waiting time (seconds)		
	with 3 operators	with 4 operators	with 5 operators
Citizen Call	10.18	1.57	0.10
General Practitioner Call	8.19	1.32	0.23
Police Call	6.98	0.89	0.18
Hospital Call	11.94	2.05	0.30
Health Care Institution	14.85	3.18	0.23

Table 12 – Average waiting times for each type of call for different numbers of working operators

We clearly see that the average waiting times of each type of caller are higher when we have fewer operators working in the emergency call center. We also see that the average waiting time is higher for hospital and health care institution calls, which is logical since these calls wait in the low priority queue and the rest waits in the high priority queue.

What also is interesting to know is the following: What was the maximum waiting time of each type of call per number of operators. This gives us an idea what the worst possible scenario is. We know that 95% of the incoming calls waits less than 6 seconds, but of these remaining 5% we want to know how long they wait at the busiest moment of the day. Therefore, in the following table an overview of these maximum waiting times is given:

Type of Call	Maximum waiting time (seconds)		
	with 3 operators	with 4 operators	with 5 operators
Citizen Call	50.10	19.77	5.54
General Practitioner Call	32.44	15.60	7.65
Police Call	93.03	33.01	5.76
Hospital Call	100.23	47.53	15.52
Health Care Institution	133.45	57.56	38.44

Table 13 – Worst scenario waiting times for each type of call, given for 3, 4 and 5 operators

We see that with 3 staffed operators, the maximum waiting time of all calls is quite high. It's normal that it's higher for the hospital and health care institution calls, since these come in at a low priority queue. But even for the urgent calls, it's much too long to wait longer than 1.5 minute like in the case of police calls. Even though it happens rarely, we cannot allow that a caller has to wait longer than 1.5 minute in the worst case. The maximum waiting times when we have 4 operators staffed are more acceptable. Of course the maximum waiting times with 5 operators are shorter, but with 4 operators, we meet our quality of service requirement, and we can live with 33 seconds of waiting in the worst possible scenario.

One more interesting fact can be observed when looking at the above numbers. We observe that when we staff 4 operators in the emergency call center, 1.6% of incoming A1 or A2 calls has to wait longer than 6 seconds. But this performance applies for an average of 317 calls per day. This means that when the number of calls on a day is much more than 317 as used in our model, we will not meet our requirement anymore. Therefore we can conclude that on Fridays and Saturdays, when the average number of calls can go up to 400, we might need to staff 5 operators. On days where the average number of calls is much less than 317, maybe 3 operators will even be enough.

The above statistics give us most of the information we need. For an overview of the average operator usage we refer the reader to Appendix A.5.

Some small remarks on the model have to be mentioned:

- Not for all values in our model, data was available. Therefore expert guess was used here. Note that expert guess is a good representation but never exact.
- The value that was used for the number of calls on one day is the average taken of all calls that were made in a year. This means that the results optimize the performance on an average day. But not every day is the same, therefore for example on Friday, where the average number of calls lays higher (400+ calls), we will probably need one operator extra. Also because of the fact that on Fridays it happens more often than normal that several people call about the same emergency at the same time.
- The optimal value that is presented above is the optimal value for the whole day, but on a day the call volumes fluctuate as well, therefore if it is possible to work with part-time employees, it is more efficient to optimize the problem for different moments of a day. This is quite logical since it is highly probable that there are not too much calls in the night on a normal weekday for example.

5.2 The basic model without the low priority queue

In this part we show what happens to the performance of the model when we exclude the low priority queue. This means we will not consider any incoming calls from hospitals and health care institutions in this variation of the basic model.

An interesting question we could ask would be: Can we meet our performance requirement with less operators if we exclude hospital and health care institution calls from the model? This would mean that we would handle low urgency calls by a different call center. This is an interesting question that many emergency call centers in The Netherlands ask themself at the moment. So it´s definitely worth looking at. Let's see how the model performs without hospital and health care institution calls:

Type of Call	% of people that waits longer than 6 sec		
	with 4 operators	with 3 operators	with 2 operators
Citizen Call	0.34%	2.93%	14.52%
General Practitioner Call	0.35%	2.32%	12.97%
Police Call	0.40%	2.11%	12.40%
AVERAGE	**0.36%**	**2.45%**	**13.30%**

Table 14 – Performance of model without calls that are of lower urgency level

We observe that if we leave out the low urgency calls, we can meet our service requirement with one operator less. The main reason for this is the fact that the average number of calls decreases from 317 to 224. But also the fact that busy low urgency calls do not make high urgency calls wait anymore contributes to the decrease in the number of necessary operators.

5.3 Multiple types of call center operators

In this part we show what happens to the performance of our model when we make the model more realistic by not solely considering one type of emergency call center operator, but two. In the basic model we considered one type of operator that is schooled to do all facets of the work. But in this section we will consider two operators, that are specialized in a different part of the work. We distinguish the following two type of operators:

- *Medical operator:* Handles the 'triage process' state and the 'medical assistance' state (see table 2 for a description of the type of states)
- *Logistic operator:* Takes over the logistic part, which exists of finding the ambulance station that is closest to the place of emergency and sending a driving route to the ambulance that will take the ambulance to the place of emergency as soon as possible.

It is interesting to look at this variant because the logistic operator takes some work away from the medical operator, which might make it possible to meet our service quality requirement with one or two medical operators less, by staffing one or two logistic operators. Since logistic operators are cheaper (lower salary), we could save costs by replacing a medical operator by a logistic operator. This situation requires some adjustments in our model. We need to know how many work the logistic operator takes away from the medical operator. From expert guess, we know that this is about 1/3rd of the work. The logistic operator also takes care of the 'coordination with hospital' state.

To give an overview of how much work remains for the medical operator together with how much work now will be done by the logistic operator, we use table 6 to create a table for both type of operators that shows how much work is done by each operator. The work that is left for the medical operator can be found by multiplying the triage state by 2/3rd and taking the coordination state away from the medical operator and place it at the logistic operator:

MEDICAL OPERATOR				
Type of call	Amount of minutes that medical worker works per state			
	State 1 Triage process	State 2 Medical assistance	State 3 (Temp) end of call	State 4 Coord. with hospital
Citizen mobile	1.93	8	no work	0
Citizen landline	1.93	8	no work	0
General practitioner	1.84	state does not apply	no work	0
Police	1.48	state does not apply	no work	0
Hospital	1.79	state does not apply	no work	0
Health care institution	3.38	state does not apply	no work	0

Table 15 – Amount of work that is done by medical operator

LOGISTIC OPERATOR				
Type of call	Amount of minutes that logistic worker works per state			
	State 1 Triage process	State 2 Medical assistance	State 3 (Temp) end of call	State 4 Coord. with hospital
Citizen mobile	0.97	0	no work	1.5
Citizen landline	0.97	0	no work	1.5
General practitioner	0.92	state does not aply	no work	1.5
Police	0.74	state does not aply	no work	1.5
Hospital	0.90	state does not aply	no work	1.5
Health care institution	1.70	state does not aply	no work	1.5

Table 16 – Amount of work that is done by logistic operator

When we use the above two tables as input for our simulation model, we can find some new results that show how the model performs when we distinguish two type of operators instead of one operator that does all the work.

We want to know how what the optimal number of medical operators and logistic operators is to still be able to meet our service requirement. In the following table, we see the performance of our model:

# of Medical operators: 3
# of Logistic operators: 1
# of daily calls (average): 317
# of simulation days: 365

Type of Call	% of people that waits longer than 6 sec	
	For medical operator	For logistic operator
Citizen Call	4.56%	22.12%
General Practitioner Call	2.95%	22.18%
Police Call	2.60%	23.38%
AVERAGE	**3.37%**	**22.56%**

Table 17 – Performance of model with 4 medical operators and 1 logistic operator

We see that when we replace one medical operator by a logistic operator, we meet our performance requirement at the medical operator queue, but not at the queue for the logistic operator. Therefore let's see what happens when we add one logistic worker.

# of Medical operators: 3
# of Logistic operators: 2
# of daily calls (average): 317
# of simulation days: 365

Type of Call	% of people that waits longer than 6 sec	
	For medical operator	For logistic operator
Citizen Call	4.74%	2.38%
General Practitioner Call	3.05%	2.52%
Police Call	2.64%	2.77%
AVERAGE	**3.48%**	**2.56%**

Table 18 – Performance of model with 4 medical operators and 1 logistic operator

We now see that both queues meet our service requirement. So when you distinguish two type of operators in the way we described it in this section, the optimal way to staff the emergency call center is with 3 medical and 2 logistic operators. But is it less costly to staff one medical operator less and staff 2 logistic operators instead? To answer this question, we need to know the hourly earnings of both type of operators. If 2 logistic operators cost less than 1 medical operator, it is profitable to staff the call center this way. If not, then staffing operators that do both jobs is more efficient (in the way that is shown in section 5.3. In the following table, an overview of the average waiting times at both queues is given for each of the above staffing combinations:

Type of Call	Average waiting time (seconds)	
	3 medical operators & 1 logistic operator	3 medical operators & 2 logistic operators
Queue for medical operator	3.20	3.32
Queue for logistic operator	25.38	1.34

Table 19 – Average waiting time of the model with 2 type of staffing combinations

We see that the average waiting time at the queue for the logistic operator is really too long. This confirms our earlier statement that 1 logistic operator is not enough. We need 2 of them.

6. Conclusion

Making emergency call centers more efficient is not a trivial job. There are simply too much factors that influence the performance of the call center. Also the random behavior of many of these factors largely contributes to the difficulty of the problem. We have provided a detailed analysis by the use of simulation with the aim to find a good balance between the quality of service and the costs at an average emergency call center in The Netherlands. The result is an interesting model that mimics reality very closely and provides useful information that can be used to improve performance.

Thorough statistical analysis has been carried out on our data to understand the randomness of the data better. We made assumptions and showed why we can make them. The result is very helpful because, even though there are quite some publications that talk about related problems, they do not concern this detail.

The most important conclusion that we can draw is that, with our given service requirement, it is optimal to staff an average emergency call center in The Netherlands with 4 employees on an average day. This is very important information because of the fact that it is a nice confirmation to most emergency call centers in The Netherlands that the current situation (average of 5 staffed employees) is quite good, but there is room for improvement. When I visited an emergency call center, the manager told me he never really did some thorough research on finding out how many operators you exactly need, but always made some guesses. These guesses proved quite good, so the outcome of our analysis is a good indicator that shows that the average call center doesn’t do a very bad job on staffing. But our results have more meaning; they show that an average of 4 operators is enough to meet the service requirement.

Then, another interesting result of our analysis is that when we exclude the non-urgent calls (hospital and health care institution calls) from our model, we see that our model performs much better and we can even staff one operator less. This is not only because of the fact that less calls are made per day, but also because it never happens anymore that an urgent call has to wait for a non-urgent call to finish. The fact that health care institution calls take much longer on average shows that the time that is saved when urgent calls do not have to wait for non-urgent calls anymore is quite much. So when we exclude the non-urgent calls from our basic model we only need to staff 3 operators.

The third scenario that we analyzed gave us insight whether it is useful to staff two type of employees. From our results we can conclude that when we distinguish a medical and a logistic operator, the optimal staffing combination exists of 3 medical operators and 2 logistic operators. Whether this way of staffing is more efficient than in the basic model, depends on the height of the salary that a logistic operator earns, together with other costs that are involved.

The fact that our model is very easily adjustable shows that it also is a very good starting point for further analysis. This means that with some small adjustments, one could find not the optimal average number of operators we need for a day, but the optimal number of operators we need for each part of a day, for each day in the week. This could optimize emergency call centers even more.

So our results are not only a confirming and helpful tool for emergency call centers in The Netherlands, they also show what the optimal number of operators is to staff an average call center with on an average day. Together with that, with some small adjustments, the model gives the possibility to find the optimal number of operators for each part of a day, for each day of the week. But the model can also be easily changed into a model of an actual call center in The Netherlands. This is because the arrival and service process are almost the same for all call centers, only the number of arrivals and average service times differ. But these values can be easily changed, which makes our model very suitable to find optimal staffing methods for all call centers in The Netherlands.

7. Further Research

This paper deals with optimization problems of emergency call centers, with the focus on an average emergency call center in The Netherlands. A lot of interesting questions that could save costs and improve efficiency were answered in this paper. Nevertheless, a process is never perfect. This means that there is always more room for optimization in different directions. This section describes a couple of these possible directions. The aim of this part is to give the reader a bit of concept of what kind of problems surround the problem that is discussed in this paper, together with some direction towards several problems that are not solved yet.

An example of a possible direction for further research is the analysis of all 24 call centers in The Netherlands. One could answer performance questions that were answered in this paper for all the other 24 call centers in The Netherlands. But one could also add some more main performance related questions and answer those as well.

Another direction where further research might be performed is analysis on a more local level. One could try to find optimal occupancy levels for emergency call centers for each day separately, since not every day is the same (think of Fridays and Saturdays, when there are much more parties going on which means more emergencies too). If we take it at an even more local level, another interesting problem one could try to solve is finding the optimal occupancy levels of emergency call centers on the different moments of a day. This type of analysis could be done by analyzing the hourly call patterns on a day, to find out on which moment of a day the most calls are made (in other words; when do most emergencies happen). This could be useful information when you have the possibility to work with part-timers in an emergency call center.

Last, more further research could be performed on the subject of fusion: Is it useful to merge emergency call centers? This problem could be answered for all possible merges of adjacent call center zones. This could save a lot of costs on unnecessary call center buildings but also on operator salaries.

Bibliography

[1] Tijms, H.C. (2004) - *Inleiding in de Operationele Analyse*, Second edition, Epsilon.

[2] Liao, S., Koole, G., van Delft and C., Jouini, O. (2010) - *Staffing a call center with uncertain non-stationary arrival rate and flexibility.*
http://www.math.vu.nl/~koole/articles/2010report2/art.pdf

[3] Roubos, A, Koole, G. and Stolletz, R. (2009) - *Service level variability of inbound call centers.*
http://www.math.vu.nl/~koole/articles/2009report6/art.pdf

[4] Gans, N., Koole and G., Mandelbaum, A. (2003) - *Telephone call centers: Tutorial, review, and research prospects*.
http://www.math.vu.nl/~koole/articles/2003msom/art.pdf

[5] Koole, G., and Mandelbaum, A. (2002) - *Queueing models of call centers: An introduction*.
http://www.math.vu.nl/~koole/articles/2002aor/art.pdf

[6] Liao, S., van Delft, C., Koole, G., Dallery, Y., and Jouini, O. (2009) - *Call center capacity allocation with random workload*.
http://www.math.vu.nl/~koole/articles/2009cie/art.pdf

[7] Koole, G. (2007) - *Call center mathematics.*
http://www.math.vu.nl/~koole/ccmath/book.pdf

[8] Restrepo, M. (2008) - *Computational Methods for Static Allocation and Real-Time Redeployment of Ambulances*.
http://legacy.orie.cornell.edu/shane/theses/MateoRestrpo.pdf

[9] Chin, G., and Coke, J. (2008) - *Simulation Modeling for Staff Optimization of the Toronto Emergency Medical Services Call Centre.*
http://www.mie.utoronto.ca/undergrad/thesis-catalog/files/163.pdf

[10] Channouf, N., L'Ecuyer, P., Ingolfsson, A. and Avramidis, A. (2001) - *The application of forecasting techniques to modeling emergency medical system calls in Calgary, Alberta.*

[11] Savas, E. S. (1969) - *Simulation and Cost-Effectiveness Analysis of New York's Emergency Ambulance Service.*

[12] Calinescu, M. (2009) - *Forecasting and Capacity Planning for Ambulance Services.*

[13] Ambulancezorg Nederland (2009) - *Ambulances in-zicht 2009.*

[14] RIVM: Rijksinstituut voor Volksgezondheid en milieu (2009) - *Tijdsduren in de ambulancezorg, analyse van spoedinzetten in 2009.*

Appendix

A.1 All ambulance stations in the Netherlands in 2009

1 Groningen	2 Friesland	3 Drenthe	4 IJsselland
Groningen Noord	Ameland	Annen	Zwolle
Groningen Zuid	Buitenpost	Assen	Raalte
Nuis	Dokkum	Beilen	IJsselmuiden
Leens	Drachten	Borger	Ommen
Uithuizermeeden	Harlingen	Coevorden	Hardenberg
Appingedam	Heerenveen	Dwingeloo	Deventer
Winschoten	Koudum	Emmen	Balkbrug
Stadskanaal	Leeuwarden UMCG	Hoogeveen	Steenwijk
Ter Apel	Leeuwarden Kijlstra Groep	Klazienaveen	Zwartsluis
Veendam	Lemmer/Balk	Meppel	Lichtmis
Sappemeer	Oosterwolde	Roden	
	Schiermonnikoog	Eelde	
	Stiens		
	Sneek		
	Terschelling		
	Vlieland		
5 Twente	**6 Noordoost Gelderland**	**7 Midden Gelderland**	**8 Gelderland Zuid**
Almelo	Apeldoorn	Arnhem	Nijmegen
Enschede	Borculo	Barneveld	Tiel
Haaksbergen	Doetinchem	Dieren	Wijchen
Hengelo	Elburg	Ede	Druten
Markelo	Ermelo	Elst	Geldermalsen
Nijverdal	Heerde	Renkum	Kesteren
Oldenzaal	Varsseveld	Zevenaar	Zaltbommel
Tubbergen	Winterswijk		Culemborg
Vroomshoop	Zutphen		
9 Utrecht	**10 Noord-Holland Noord**	**11 Amsterdam/Waterland**	**12 Kennemerland**
Utrecht - Noord	Veiligh. Reg. NHN Den Helder	Valckenierstraat	Heemskerk
Maarssen	Veiligh. Reg. NHN Hoorn/Wognum	Karperweg	Velsen-Zuid
Utrecht - Centrum	Veiligh. Reg. NHN Den Burg	Purmerend	Haarlem
Nieuwegein	Veiligh. Reg. NHN Hoogkarspel	Zaandam	Hoofddorp
Vinkeveen	Veiligh. Reg. NHN Wieringerwerf	Amstelveen	

Woerden	Amb. dnst Kennemerland Alkmaar N	Amsterdam ZO	
Amersfoort - Centrum	Amb. dnst Kennemerland Alkmaar Z	Aalsmeer	
Amersfoort - Noord	Amb. dnst Kennemerland De Mossel	Monnickendam	
Zeist	Amb. dnst Kennemerland Schagen	Westzaan	
Doorn			
Rhenen			
13 Gooi- en Vechtstreek	**14 Haaglanden**	**15 Hollands Midden**	**16 R'dam-Rijnmond**
Zuid	Centrum Westeinde	Moordrecht	Capelle aan den IJssel
Noord	Mariahoeve Bezuidenhoutseweg	Nederlek	Rotterdam Centrum
Vechtstreek	Wassenaar Hofkampweg	Gouda	Rotterdam Noord
	Mangaanstraat	Alphen a/d Rijn	Barendrecht
	Naaldwijk Westland Zuideinde	Nieuwveen	Schiedam
	Forepark Moldau	Leiden	Spijkenisse
	Zichtenburglaan	Leiderdorp	Haringvlietdam Nrd
	Zoetermeer Blauwroodlaan	Noordwijk	Brielle
	Delft Engelsestraat	De Zilk	Dirksland
17 Zuid-Holland Zuid	**18 Zeeland**	**19 Brabant Midden West**	**20 Brabant Noord**
Dordrecht	Middelburg	Bergen op Zoom	Den Bosch
Gorinchem	Goes	Roosendaal	Boxtel
Klaaswaal	Zierikzee	Breda-Zuid (Ulvenhout)	Oss
Meerkerk	Tholen	Breda-Noord	Haps
Papendrecht	Rilland	Tilburg-Noord	Uden
Zwijndrecht	Neeltje Jans	Tilburg-Zuid	Veghel
	Oostkapelle	Waalwijk	
	Oostburg	Oosterhout	
	Terneuzen	Giessen	
	Hulst		
21 Brabant-Zuidoost	**22 Noord en Midden Limburg**	**23 Zuid Limburg**	**24 Flevoland**
Eindhoven	Venlo	Geleen	Almere
Helmond huidig	Venray	Heerlen	Lelystad
Best	Helden-Panningen	Maastricht	Emmeloord
Eersel	Bergen	Gulpen	Zeewolde
Deurne	Roermond		Dronten
Valkenswaard	Weert		Urk
Maarheeze	Echt		

(Source: AMBULANCEZORG NEDERLAND [13])

A.2 THE 24 EMERGENCY FACILITY ZONES IN THE NETHERLANDS IN 2009

(Source: AMBULANCEZORG NEDERLAND [13])

A.3 FRACTION OF TYPE OF CALLS WITH URGENCY LEVELS IN 2009

Type of call	***Average of A1 & A2***	***A1***	***A2***	***B***
Unknown/Not filled in	15.2	13.5	17.5	17.5
Citizen call	39.1	51.6	21.3	2.4
General practitioner	26.2	21.2	33.3	22.3
Hospital	7.4	3.1	13.5	6.0
Other	5.0	3.3	7.4	8.1
Police/firemen call	3.5	4.3	2.4	0.2
Health care institution	3.3	2.6	4.3	42.5
Total	**100**	**100**	**100**	**100**

(Source: AMBULANCEZORG NEDERLAND [13])

A.4 ALL RELEVANT PARAMETERS WITH VALUES

ARRIVAL PROCESS	
Parameter name	**Parameter value (arrivals p/h)**
λ_{CIT} (mobile)	3.60
λ_{GP} (landline)	0.90
λ_{GP}	4.28
λ_{POL}	0.41
λ_{HOS}	1.19
λ_{HCI}	2.95

SERVICE PROCESS	
Parameter name	**Parameter value (minutes)**
μ_{TR}(cit mobile)	2.90
μ_{MA}(cit mobile)	8.00
μ_{TE}(cit mobile)	23.29
μ_{CO}(cit mobile)	1.50
μ_{TR}(cit landline)	2.90
μ_{MA}(cit landline)	8.00
μ_{TE}(cit landline)	23.29
μ_{CO}(cit landline)	1.50
μ_{TR}(gp)	2.76
μ_{TE}(gp)	31.29
μ_{CO}(gp)	1.50
μ_{TR}(pol)	2.21
μ_{TE}(pol)	31.29
μ_{CO}(pol)	1.50
μ_{TR}(hos)	2.68
μ_{TE}(hos)	180.00
μ_{CO}(hos)	1.50
μ_{TR}(hci)	5.07
μ_{TE}(hci)	180.00
μ_{CO}(hci)	1.50

TRANSITIONS	
Parameter name	**Parameter value**
P_1(cit mobile)	25%
P_2(cit mobile)	11.11%
P_1(cit landline)	25%
P_2(cit landline)	11.11%
P_1(gp)	11.11%
P_1(pol)	11.11%
P_1(hos)	11.11%
P_1hci)	11.11%

A.5 OPERATOR UTILIZATION STATISTICS

Number of operators	Non-stop operator utilization	
	Amount	*Percentage*
5	0.91 out of 5	18%
4	0.91 out of 4	23%
3	0.91 out of 3	30%