Influence of different sources of variation on the capacity of an NS service location

Bachelor thesis

Author Vanja Zomer

Education Civil Engineering
University of Twente

Date 26-03-2019

Abstract

The Dutch railway system is becoming more crowded, resulting in a need for more train units. This also means that more train units need to be treated at the service locations of the Nederlandse Spoorwegen (NS). To handle this increase, the NS wants to make better use of the available resources.

This research investigates the influence of certain sources on the capacity of service locations. The sources investigated are the arrival process, allowing coupling and decoupling and the topology of a service location. Ultimately, the location layout turned out to have a lot of influence. Therefore the arrival process and allowing coupling and decoupling are also investigated when specified for the location topology.

For these sources, survival functions are made based on model data from the NS. Those survival functions show the chance that no solution is found (chance of failure) for a work package per number of train units. Namely, in a work package a certain number of train units are planned. The chance of failure increases while increasing the number of train units in a work package, since it is harder to find solutions for work packages in which less train units need to be handled. These survival functions are made based on the Turnbull Algorithm and the Product Limit Method.

Once the survival functions of different scenarios of the sources are composed, influence is determined by testing those survival functions on significant differences. This is done by executing the Kolmogorov Smirnov test. More information about the cause of impact is obtained by comparing the medians and the variances of the survival functions. The source of variation in median in this research was the arrival process, especially when specified per service location layout. The sources of variation in variance of this research were the topology of the locations, the arrival process at the a carousel layout and allowing coupling and decoupling at a shuffleboard layout.

Terminology

In this research, terms will be used which need explanation. Below, some definitions are given to prevent confusion.

Allowing coupling/decoupling	Allowing coupling and decoupling is an option that can be set in the behandelcalculator. This will be investigated as a potential source of variation.
Arrival period	A period in which the train units arrive. This period starts at 17.30 and ends at 2:20 the next day.
Capacity	The number of train units that a percentage of the cases can be treated during a shift at a service location.
Cutting loss	The remaining railway that cannot be used because no train can fit there.
Departure period	A period in which the train units depart. This period starts at 2.20 and ends at 8:00.
Instance	A work package. This includes all information of what has to happen on a service location that particular shift.
Sawing	The action that a machinist needs to walk from one end of the train to the other end of the train because the train needs to drive in opposite direction.
Service location	A location where a train unit receives its planned maintenance. In this research service location the Binckhorst will be investigated.
Shift	A period in which a work package is planned. This period is from 17:30 until 8:00 the next day.
Source of variation	A source of variation is a factor that has influence on the capacity of a NS service location (e.g. the arrival process could have influence on the capacity so this could be a source of variation).
Topology	The layout of the railway at a service location.
Train	Ride in revenue service for passenger transport.
Train unit	Train set, fixed composition with driver cabins at both ends.
Variation in capacity	The number of train units that can be handled in a shift at a service location differs per observation, so the observations for the capacity vary per measure. This variation is caused by sources of variation.
Work package	A work package is a package which includes all the maintenance tasks and shunting movements of the train units that need to be treated in that particular shift. The model name of a work package is 'instance'.

Table of contents

Abstract 2
Terminology 3
Chapter 1 Introduction 6
1.1. Problem context 6
1.2. Company 6
1.3. Aim of the project 7
1.4. Limits and boundaries 7
1.5. The ‘behandelcalculator’ 7
1.6. Research framework 9
1.7. Scientific relevance 12
1.8. Report structure 15
Chapter 2 Variation in capacity 16
2.1. Method description 16
2.2. Methods for constructing a survival function 16
2.3. Cumulative distribution function and the probability density function 19
2.4. Quantification of variation 20
Chapter 3 Selection of sources 22
3.1. Treatment process 22
3.2. Choice of sources 26
Chapter 4 Influence of sources 29
4.1. Methodology 29
4.2. Statistical tests 30
4.3. Arrival process 32
4.4. Coupling and decoupling 38
4.5. Topology 42
4.6. Influence of the topology on the capacity 48
Chapter 5 Conclusion and discussion 56
5.1. Conclusion 56
5.2. Discussion 57
5.3. Recommendations 58
References 59
Appendices 61
A. Matlab script constructing a survival function 61
B. Matlab script random generator 64
C. Matlab script filter 65

D. Critical values χ^2-test 67
E. Matlab script adding two analyses 68
F. Matlab script statistical tests 72

Chapter 1
Introduction

In this chapter an introduction to the assignment is given. First, the problem context will be discussed. Then, the role of the company is described, just like the problem aim. After that, the limits and boundaries and terminology are stated. Then the scientific relevance will be discussed as well as a description of the model that will be used in this research. Finally, the research framework and the report structure will be sketched.

1.1. Problem context

The Nederlandse Spoorwegen, the NS, has to deal with an increasing number of travellers. Therefore, the NS has to deploy more trains to the railway system, which results in an increasing number of trains and train units that need to be checked for safety and comfort, cleaned and washed. The checking and cleaning process is called the service process. This process occurs on certain service locations. However, since more train units need to be handled at service locations due to the increasing use of train units, the capacity of those service locations has become a bottleneck. In this case, the capacity of a service location is considered as the number of train units which in 95% of the times can be handled in a certain period.

It happens frequently that not every train unit that was planned for that shift, can finish the process on time. However, the NS wants to deliver all trains perfect, but now choices have to be made. Delivering an unsafe train is not an option, but delivering a dirty train might be better than a cancelled train. However, the NS does not want to make those choices at all, so capacity improvement is a very important issue.

The capacity is built up from different sources. But which sources may contribute to the capacity and what is their influence? The NS wishes to better understand the influence of different sources on the capacity of service locations. (NS en ProRail: het spoor is bijna vol, 2018)

1.2. Company

This assignment is commissioned by the NS. The NS is a company that regulates the biggest part of the public transport of trains. The department that provides this project is the department of Maintenance Development. This department works on the improvement of maintenance processes.

The vision of the NS is to make the Netherlands assessible for everybody. The NS has a clear ambition: they want to deliver mobility of the highest level. Always close, always affordable and always sustainable. It is important to pay attention to this because the railway system will become even busier in the future.

This research contributes to the vision of the NS, since this project will investigate which sources have influence on the capacity of service locations. Without good functioning service locations, the railway network cannot expand, since the train units need to be safe and clean for delivering the highest level of mobility. The capacity of service locations needs to improve so more train units can be treated in the future. Besides that, it is better to use the service locations optimally instead of constructing extra service locations, since that is less sustainable and more expensive.

The NS wants to deliver a good product to their customers, the travellers. This includes clean, safe and comfortable trains that depart and arrive on time. (Boxtel, 2018)

1.3. Aim of the project

Since the capacity of service locations seems to be the bottleneck in the possibility of expanding the railway system, it needs to be improved. The number of train units that can be handled in a shift at a service location differs per observation, so the observations for the capacity vary per measure. It is important to get insight in which sources cause this variation in capacity and to what extend their influence is.

It is important to investigate which sources may have influence and require further research. By making hypotheses and testing them, the influence of different characteristics will mapped better. In the end, the aim of the project is to better understand how the sources of variation influence the capacity of service locations, so that the NS can work further with these results and more train units get treated each shift.

This research will focus on one service location where a lot of research is being done in the department of Maintenance Development. This service location is the Binckhorst which is located at The Hague Central station.

This research will use the model data obtained from a large model; the 'behandelcalculator'. This is a tool for executing this research. The 'behandelcalculator' is a simulation model of service location the Binckhorst. The model searches solutions for hypothetical work packages, which not necessarily have taken place. This data will be used in this research instead of empirical data.

1.4. Limits and boundaries

To execute this research, some limits and boundaries have to be set. In this research there are four limits or boundaries set, stated below.

- The influence on the capacity of only a number of sources will be investigated.
- Research on service location the Binckhorst will be taken into account only.
- Only model data will be used for the main research.
- The capacity can be expressed in different ways, but in this research is will always be expressed in number of train units so that different results can be compared with each other.

1.5. The 'behandelcalculator'

The behandelcalculator is a model that is made by employees of the department of Maintenance Development. This model is used to determine the capacity. It is important to understand the model before the research framework is made. Therefore, a flowchart is shown below. It is a complex model, so an explanation of the flowchart is also given. Besides that, this section will explain what the input and output look like and how to process the output into interesting information.

1.5.1. Flow chart of the behandelcalculator

A global flow chart of the model can be seen in the figure below.

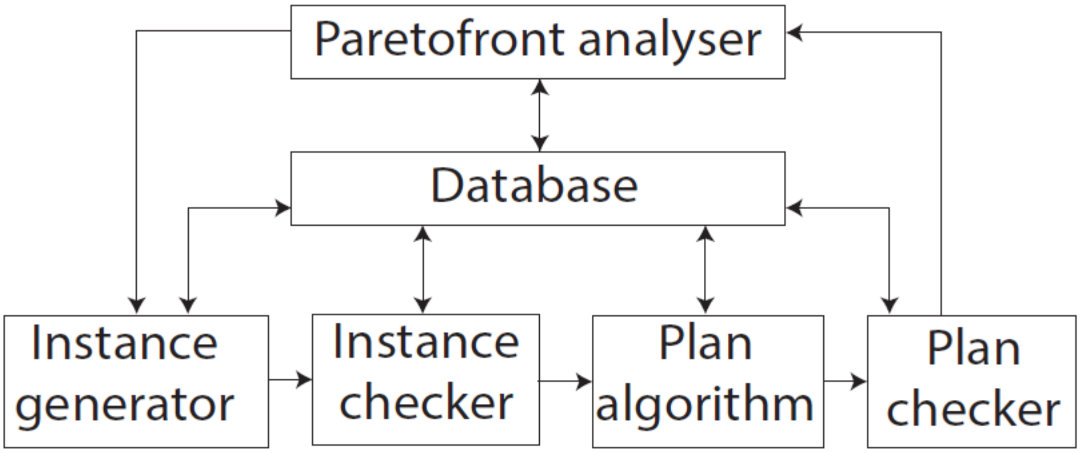

Figure 1: Flowchart behandelcalculator

1.5.1.1. Database

In the database all the input is stored. The input consists of the characteristics of the service location, but also of all the chances and statistical distributions that maintenance has to take place or not. For example: train units need to be washed once every twelve days. The chance that a train unit needs to be washed is therefore $^{1}/_{12}$. Another important thing that can be installed in the database is the characteristic that is going to be variable and thus in what terms the capacity will be expressed. In this research this variable will always be the number of train units. It is also possible to express the capacity in more variables, but this will not be done in this research. The output is also written in the database again.

1.5.1.2. Instance generator and instance checker

Every iteration a combination of variable settings is chosen and sent to the instance generator. In this research, the variable that is analysed is the number of train units. Therefore a number of train units value is sent to the instance generator. For these settings a work package will be generated by the instance generator using the corresponding characteristics. All miscellaneous settings are determined by chances of occurrence which are in the database. Such a work package is called an instance. This comes down to all information that has to be executed, such as whether a task needs to be executed or not and how long the tasks in those particular cases take.

Once a work package is set, it is sent to the instance checker. This instance checker looks whether the work package is feasible. In case that in the work package more train length needs to be at a service area, than railway is present, a plan cannot be found at all. In this case, the plan algorithm does not need to run. So, this instance checker saves a lot of calculation time.

1.5.1.3. Plan algorithm and plan checker

The plan algorithm is a program that searches for a solution of the work package. Once a plan is found, the plan algorithm quits its job, because the only thing the behandelcalculator wants to know is whether a solution is found or not. The plan algorithm looks for a solution by using heuristics like the local search technique.

In the plan checker, the plan found by the plan algorithm is checked on feasibility. When a resource at one moment is exceeded, the plan is not feasible, so actually no real plan is found.

1.5.1.4. Pareto-front analyser

The pareto-front analyser is a tool to determine at which points in at the chance of failure is at most 0,05. This means that in at least 95% of the cases a solution is found for a work package. The pareto-front analyser searches for this front by generating a lot of setting values around that 95%-front. However, in this research will not be worked with a fixed percentage for the capacity. Instead the chance of failure between 0 and 1 will be investigated. Therefore, the pareto-front analyser will not be used. To get information about the chance of failure between 0 and 1, spread data is needed. In this research will be worked with 30 instances per number of train unit value.

1.5.2. Output

The output differs from the number of characteristics that are varied. In this research only one variable will be varied, namely the number of train units. Therefore only the one dimensional output will be discussed.

The output in this research is a list of instances with the information whether a solution is found or not. Per instance it is also known of how much train units it existed. These data is shown in a lot different tables in the program SQL Server Management Studio.

The data can be analysed in excel, but a more advanced tool is the program Power BI. This is a program that is used in the department and will be interesting to use for analysing the output results.

1.6. Research framework

To be able to execute the research, a research framework is made. First the research questions are stated. Then, the methodology to answer each of these research questions is discussed.

1.6.1. Research questions

To be able to carry out the research the aim of the project is translated into one main research question. Thereafter, this question is split into some central questions. Those central questions are split into sub questions. The sub questions will answer the central questions and the central questions will answer the main question. In this way, the aim of the project will be achieved.

Main question

How do different sources influence the variation in capacity of the NS service location the Binckhorst?

Sub questions

[1] Which sources may have influence on the variation in the capacity?

- [A] Can the variation in capacity be quantified?
- [B] Where could be bottlenecks in the service process?
- [C] Can the already existing instances be used to decide which sources are interesting to investigate?
- [D] Which sources are taken into account?
- [E] Do the sources cause variation in capacity?

[2] How do the sources influence the variation in capacity?
- [A] What could be the reason of the influences hypothetically?
- [B] How can the hypotheses be tested?
- [C] What influence do the different sources of variation have on the capacity?
- [D] Are there any interactions between sources?

1.6.2. Methodology

To answer the research questions formulated above, both qualitative and quantitative research will be done. To reach the answer to the main question, different methods will be applied. In this project, data analysis will take place. Data review and statistical analysis should be a part of it too. Below, for the different questions is stated which method likely is to use for answering the question.

[1] Which sources may have influence on the variation in the capacity?
[A] Can the variation in capacity be quantified?
In reality a lot of sources can have influence on the capacity. Using a lot of instances which are already available, a survival function can be made by using the Turnbull Algorithm and the Product Limit Method. which expresses the chance of failure for an instance that contains a particular number of train units. It is assumed that the instances used for compiling this graph give a good estimate for the capacity. From this survival function, the probability density function can be determined. Probably, this graph will not be a smooth function in the beginning. This is caused by sources of variation which influence the capacity. By identifying the sources and filtering them out, the graph gets rid of outliers and becomes homogeneous, theoretically. So by constructing a probability density function with the data available can be shown that there really is variation in capacity.

[B] Where could be bottlenecks in the service process?
Literature study will be done on the treatment process, so the process at service locations becomes more clear. By looking for potential problems in the service process, potential bottlenecks can be found. Improvement possibilities will can be mentioned here too.

[C] Can the already existing instances be used to decide which sources are interesting to investigate?
Initially, the already existing instances can be investigated. It is possible that for some instances with many train units no solution is found, and for some instances that consist of a low number of train units a solution is found. Maybe striking differences will come up. This can be used to identify possible sources of influence.

The database with already existing instances consists of instances with a lot of different characteristics. From this database, the instances with specific characteristics can be filtered. This results in smaller datasets from which survival functions can be composed. The expected value and variance of the sets can be calculated and compared to the expected value and variance of the large database.

[D] Which sources of variation are taken into account?
By constructing the probability density function for the capacity based on the available instances, a non-homogeneous graph with a bunch of outliers arises. By filtering out influencing sources, the graph gets more and more homogeneous theoretically. However, lots of sources can have influence on the capacity of service locations. It is not reasonable to investigate all sources in this research. That is why a selection of sources will be made in this research question. In sub question 1B and 1C

possible sources that might have influence on the capacity of service locations are defined. From these sources a selection will be made which sources are going to be taken into account. The miscellaneous sources are seen as one group from which the influence is neglected.

[E] Do the sources cause variation in capacity?
To investigate whether there actually is variation in capacity caused by the sources selected in sub question 1D, this has to be tested. For every source the behandelcalculator needs to be executed with differences in only that particular source. When the expected value and the variation of the cumulative function (the chance of failure plotted against the number of train units in a work package) remain constant, this means that this source does not cause variation in capacity and the following research does not need to be executed. However, when significant differences occur, the source does have influence.

[2] How do the sources influence the variation in capacity?
[A] What could be the reason of the influences hypothetically?
In question 1C different sources that might have influence on the variation in capacity have are selected. These sources are going to be invested. In this question hypotheses will be stated about the influence of those sources. This will be done based on the influence they are likely to have.

[B] How can the hypotheses be tested?
In question 1E hypotheses about the influence of different sources on the capacity are stated. In this part of the research, a plan will be made to test those hypotheses. By changing characteristics of the investigated source in the input of the behandelcalculator, the model will show differences in capacity. Statistical tests will prove whether the influence on capacity is significant or not and thus whether the hypothesis has to be rejected or not. The results of those tests will support or will not support those hypotheses.

First the survival functions of scenarios, analysed with the behandelcalculator, must be made. In a survival function the chance on failure of an instance is plotted against the number of train units in an instance. To check whether significant differences exists between two scenarios Kolmogorov-Smirnov test will be performed. Based on differences in the cumulative distribution function that occur while differencing a source, tests can be executed to test the hypotheses with. Statistical tests that can help investigating an hypothesis are for example the student-t-test, the F-test or the paired samples test. The test that seems to be likely to execute for a particular source, will be chosen in this research question.

[C] What influence do the different sources of variation have on the capacity?
To get to know what the influence of the different sources is on the capacity, the hypotheses will be actually tested. The results can be analysed in the program Power BI, which consists of a suite data analytics tools. The results will be obtained in this research question.

[D] Are there any interactions between sources?
It is possible that some sources correlate with other sources or that some sources interact with other sources. To determine possible interacting sources a causal loop diagram will be made in which the selected sources in sub question 1D are included.

To analyse whether there exist interactions between the sources, scenarios will be set for possible interacting sources of variation. For these scenarios the behandelcalculator will execute analyses with corresponding settings, so the data that is generated corresponds with the scenarios. This can also be done by filtering the already available data, however, not all data for the different scenarios is available.

About the interactions hypotheses will be set and these will be tested in the same way as the single sources of variation in sub questions 2A, 2B and 2C.

1.7. Scientific relevance

The research that will be carried out in this project is relevant for the NS, but it is also relevant in other disciplines. In other disciplines research on capacity variation is done as well. In those papers methods are used which might be useful in this research as well. Those are described here. In this heading the capacity expression that will be used in this research is described, but also papers about the capacity variation in other disciplines are discussed.

1.7.1. Capacity expression

One of the bottlenecks in the logistic planning process at Dutch railways is the capacity of the infrastructure at the larger railway stations. To provide passenger trains with the right composition of rolling stock, many shunting movements between platform tracks and shunting areas are necessary, especially just before and after the peak hours. (van den Broek, Hoogeveen, & van den Akker, 2012) The gross capacity of a service location can be measured as the number of meters of track available at the service location. However, several aspects need to be taken into account when determining the capacity, because the trains cannot just be parked directly behind each other because not at every parking all resources are available. (Lentink, 2006)

At a service location like Zwolle, the capacity in terms of the total length of the shunt tracks is scarce. However, in Zwolle relatively many shunting processes take place. This implies that capacity expressed in other terms or more terms, may better reflect this. (Freling, Lentink, Kroon, & Huisman, 2002) Therefore a pareto-front can be used to express the capacity. A pareto-front is the front at which you cannot increase one value without decreasing another value. In the example in Figure 2, the percentage of train units that need to be washed cannot increase without decreasing the number of train units. In this example the pareto-front is two dimensional, since it is expressed in two characteristics.

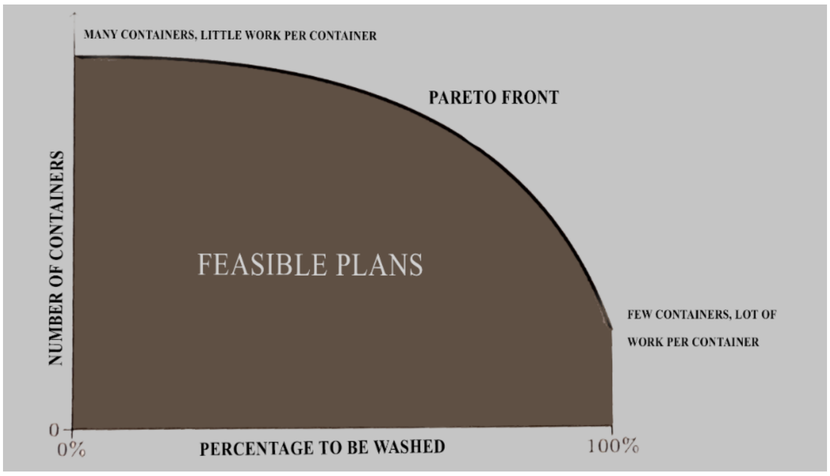

Figure 2: Pareto-front (NS Techniek, 2018)

The capacity of a service location can be expressed as the pareto-front of distribution parameters where the composition of the daily work packages is characterised statistically, so that for at least 95% of the shifts a feasible plan can be found. (NS Techniek, 2018) In this research, only one characteristic will be used to express the capacity, namely the number of train units that can be treated in a shift. Thereby, there will not be worked with the percentage of at least 95%. Instead, all success rates will be considered.

1.7.2. Predictability of the capacity

This research will help to declare differences in capacity, but it will not really contribute to prediction of the capacity. It can also make a prediction of the capacity based on the number of train units a work package exists of, but then only one characteristics is taken into account. For a better prediction of capacity, the capacity should be expressed in more characteristics. Based on what the characteristic in a work package are, a prediction of the capacity can be made with taking into account the values of the characteristics. The more characteristics the capacity is expressed in, the better a prediction will be.

1.7.2.1. Capacity in other disciplines

Capacity can be defined in different ways. This depends on the research goal. (Olba, Daamen, Vellinga, & Hoogendoorn, 2017) It can be defined as the maximum number of vehicles on a road or the maximum number of passengers at an airport or a train station for example. However, it would be better to express the capacity per unit time which can be accommodated under given conditions with a reasonable expectation of occurrence. This will be used in the research too since in the research will be looked at the number of trains that can be handled during a shift.

Capacity says something about the physical amount of vehicles or passengers a road, train station or airport can afford. It depends on traffic conditions, geometric design of the road etc. Capacity is expressed in terms of units of characteristics, for example traffic composition and the environmental conditions too. Capacity is a probabilistic measure and it varies with respect to time and position. (Mathew & Krishna Rao, 2007) In the examples below is described what sources might cause capacity variation in their disciplines, but also how this is applicable on the research in this proposal.

Capacity variation in traffic

In the capacity of freeways in traffic also variation exists. One of those sources that cause variation is the quasi-random nature of the occurrence of road congestion. Congestion has a negative impact on the capacity of the road network. Such congestion occurs at the delay point sections of the road network on which the vehicle has to stop or slow down. Effective control of the capacity of delay points can be considered an important and urgent scientific problem. However, this is very complex especially since some congestion is unpredictable, like car accidents. Since the occurrence of congestion cannot be predicted yet, this causes variation in the capacity of the road network. (Kucherov, Rogozov, Lipko, & Elkin, 2018)

Another factor that has influence on the variation of capacity in traffic is the type of the day. First, there was a suspect that there are differences in capacity on a workday relative to weekend-days and holiday-days. This is tested by making hypotheses and it turned out to be true. (Calvert, Taale, & Hoogendoorn, 2015)

In the paper 'Quantification of motorway capacity variation: influence of day type specific variation and capacity drop' a probability function of breakdown and discharge capacity is made. This is basically a function where the probability of failure is plotted against the traffic flow. At increasing traffic flow, the change on failure gets larger. This principle will be used in the research on capacity of NS service locations.

Another principle that will be used in this research that is also used in this paper is the way the distribution of the probability of failure fit is determined. This is done using the Kolmogorov-Smirnov test, since the KS-test is best suited to test capacity distributions in such a way as it quantifies a distance between the empirical distribution of the sample and the cumulative distribution of a reference distribution. More importantly, the KS-test is distribution free and therefore makes no assumptions with respect to the underlying distribution. (Jia A, 2010) (Chakravarti IM, 2009) The KS-test is also an exact test while some other commonly applied tests, such as the chi-squared-test, depend on an adequate sample size to validate approximations. (Ross SM, 2009). Since in the assignment in this proposal it is also about capacity, this method seems likely to use. However, model data is used instead of empirical data, but the model data can be approached as empirical data in this research.

The paper of Calvert, Taale and Hoogendoorn makes use of scenarios. The data which they want to investigate is filtered based on the scenarios. This seems a good method, but in the research for NS not just three scenarios from one source which can be separated very easily are investigated. In this research sources like the arrival process will be investigated a distribution of the arrival process cannot be determined based on an instance. However, in this research will be made scenarios too. These scenarios will be used to show interactions between sources if possible. This will not be done by filtering the existing data but by executing new analyses based on the scenarios, and in this way generating data based on the scenarios.

Capacity variation in cargo shipping

Roll-off-roll-on (Ro-Ro) terminals, at which wheeled cargo is shipped, require a growth in terms of capacity since the Ro-Ro traffic increases worldwide. At Ro-Ro terminals the capacity variation occurs. This is caused by different variables, such as the number of vehicles arrived at a terminal, ship capacity, number and layout of terminal gates, terminal traffic, local traffic and security checks. Some of those characteristics are fixed for a certain location, but some are not. The number of vehicles arrived to a terminal for example. This can be compared to the arrival process of trains at a service location. Differences in this arrival process will affect the capacity and will thus cause variation in capacity of the terminal. Another source that is variable is the capacity of the ships which is comparable with the type of trains. The terminal traffic also influences the capacity of the terminal. On the terminal the cargo needs to make movements which can be compared with the shunting movements at service locations. In short: the process at terminals is in many respects commensurate with the process at NS service locations just like the capacity variation problem. (Özkan, Nas, & Nil, 2016)

In the paper 'Capacity Analysis of Ro-Ro Terminals by Using Simulation Modelling Method' analysis on the capacity is done by making a simulation model. The data from the simulation model is used to investigate interactions between sources by making scenarios and run the model, with the settings conform the scenarios, for outcomes. In the research on the capacity of NS service locations will be scenarios set to investigate interactions between sources. In this research also a simulation model is available which generates the data. The model can be installed in such a way that the scenarios are varied and the rest of the settings are constant. In this way differences and interactions between factors can be found.

Capacity variation in the aircraft business

Maintenance on aircrafts needs to take place just like maintenance on train units. In the process of heavy maintenance the capacity varies a lot. This mainly has to do with the uncertainty of the maintenance. This maintenance exists of predictable and unpredictable maintenance. Predictable maintenance is the maintenance that protects the airplanes from failing and prevents worse problems. In this part of the maintenance are characteristics that can cause variation in capacity, such as a delay that occurs because a task needs more time. However, the main impact on the capacity is caused by the lots of unscheduled maintenance activities that have to take place. Because it is an unpredictable source which can have large impact, it also contributes to variation in the capacity of the maintenance process. (Rosales, 2015) The paper 'Analysing delays and disruptions in Aircraft Heavy Maintenance' focusses on building a model. This is not relevant in the analysis of the capacity of service locations. However, in the paper is made use of a causal loop diagram to get insight in interactions of variables. This could be a clever method to get insight in interaction of variables in the treatment process at NS service locations.

1.8. Report structure

In this research, first the variation in capacity will be evinced. This will be done in chapter 2. To do so, a survival function needs to be constructed. Therefore, the construction of the survival function will be described in this chapter as well. This method for constructing survival functions will be used in the rest of the research too every time a survival function needs to be constructed.

In chapter 3, the sources that are going to be investigated are chosen. To do so, it is necessary to get a better insight in the treatment process at service locations. Therefore, an overview of the treatment process will be given. After that, the sources that are going to be investigated in chapter 4 will be chosen.

In chapter 4, the sources are investigated. The first source is the arrival process, the second source is allowing coupling and decoupling and the last source is the topology of the service location. For the topology, the Kleine and the Grote Binckhorst are used, since those two have totally different topologies. After that, the interaction of the sources with the location topology will be investigated.

In the last chapter, chapter 5, the conclusion of the experiences of chapter 4 will be stated. Also the methods used in the research will be discussed. Finally, a couple of recommendations will be described in this chapter.

Chapter 2
Variation in capacity

In reality a lot of sources can have influence on the capacity. These sources cause variation in capacity. This research will look at sources that cause this variation in capacity. In this chapter will be shown that there is variation in capacity. Therefore a method is sought to show that variation in capacity exists and whether this can be quantified.

2.1. Method description

Using a database in which many different instances are already available, a cumulative graph can be made which expresses the chance of success for an instance that contains a particular number of train units. This graph is a survival function. A survival function can be composed by executing the Product Limit method. It is assumed that the instances used for compiling this graph give a good estimate for the capacity. From this cumulative graph the probability density function can be determined. In case variation in capacity does exist, this graph will not be a clear function in the beginning, since this is caused by sources of variation which influence the capacity. By identifying the sources and filtering them out, the graph gets rid of outliers and becomes homogeneous, theoretically. So by constructing a probability density function with the data available can be shown that there really is variation in capacity.

2.2. Methods for constructing a survival function

Data is generated with the 'behandelcalculator'. This data consists of instances, which have characteristics. For those instances a solution is or is not found. With these data a graph can be made in which the chance of success is plotted against the number of train units in an instance. This is called a survival function.

$$S(t) = 1 - F_c(t) = p(c \leq t) \text{ where } S(t) \text{ is the survival function,}$$
$$F_c(t) \text{ is the capacity distribution function, } c \text{ is the capacity and } t \text{ is the number of train units.}$$

How survival functions will be constructed in this research will be explained in this paragraph.

2.2.1. Intuitive method

A simple way is calculating the number of failures and the total number of instances generated with a particular number of train units. With this information the chance on failure can be calculated for every number of train units. However, these chances can fluctuate. In practice, the chance of failure cannot become lower than before since the chance that an instance with less train units will be solved is always higher than that an instance with more train units will be solved. A possible way to get rid of this problem and construct a cumulative function, is to assume that the chance of failure is at least the chance at number of train units minus 1. An uncertainty that this method brings with it, is when a calculated chance of failure is higher due to few measurements, the chance will not be lower anymore.

In the next example this problem occurs. An analysis is done and it should give a capacity of 20 train units. However, there is been done one measurement consisting of just one single train unit and for this instance no solution could be found in the run period. In this case, the chance of failure for one train unit is 1, and since this chance cannot become lower than before, the chance of failure will falsely be 1 for all number of train units. However, this is a very specific case and this will not happen soon since the pareto-front analyser is built in the 'behandelcalculator'.

2.2.2. Product Limit Method

Another way to construct a survival function is by using the Product Limit Method based on lifetime data. In a survival function, the number of train units is plotted against the chance of success. This can be easily translated in a function in which the number of train units is plotted against the chance of failure, since chance of failure is 1 minus the chance of success. The Product Limit Method estimates the survival function based on data. The Product Limit Method has several estimators. Three well-known estimators are discussed here. The Kaplan-Meier estimator, the Nelson-Aalen estimator and Turnbull's estimator.

2.2.2.1. Censoring data

Before explaining the three Product Limit estimators, censoring needs to be discussed. In survival analysis events on individuals are investigated during a time period. However, it is possible that the event already happened for an individual before the investigation period, but it is unknown when. The only thing known is that the event happened before the beginning of the investigation. This data is called 'left-censored data'. It is also possible that the event has not yet happened at the end of the investigation. The only thing known in this case is the fact that the moment of the event is later than the end of the investigation. This is data is called 'right censored'. These censored data consists of interesting information that you want to use, but it cannot be used directly. (Klein & Moeschberger, 2003)

In this research, there is not such a variable as time, but it is possible to look at the number of train units as the time. The number of train units in an instance increases, until the capacity drops. This happens when no solution can be found for the first time in that lifetime. In this case, a lifetime is an increasing number of instances, from zero train units in an instance till the number of train units in an instance where the capacity drops.

This capacity drop is the event. However, the data available are individual instances consisting of a particular number of train units. There is not such a thing that every time a solution is found one train is added to the instance, and the behandelcalculator calculates whether this new instance can be solved. The data can be seen as a snapshot. At the moment that 15 train units are in the instance and no solution is found, this means that the capacity has dropped somewhere before the 15 train units, which is left censored data. At the moment that 15 train units are in the instance and a solution is found, this means that the capacity is not dropped yet, and therefore the capacity will drop when the instance consists of more than 15 train units, but it is unknown at how many train units this will be. These data is right censored.

2.2.2.2. Estimators

The Kaplan-Meier estimator is an estimator used a lot as estimator for the Product Limit Method. This estimator is used to estimate the traffic breakdown capacity for example. The traffic flow increases every time interval until congestion occurs. At every interval a measure is done, where occurrence of congestion is a success event. The failures are right censored because the success event did not take place yet and will take place in the future. (Brilon, Geistefeldt, & Regler, 2005) This is a typical example of the use of the Kaplan-Meier estimator since in this example only right censored data is present. The Kaplan-Meier estimator handles right censored data, but it does not handle left censored data. In problems that have left censored data, a solution is to apply left truncation. This principle is ignoring the left censored data, but this is combined by throwing useful information away (Klein & Moeschberger, 2003). Another way to get along with left censored data by using the Kaplan-Meier estimator is making assumptions about the lifetime of left censored data.

The Nelson-Aalen estimator is similar to the Kaplan-Meier estimator. The Nelson-Aalen estimator can be approached with the Kaplan-Meier estimator. (Wolfe, 2000) Since the Kaplan-Meier method is more common, the Nelson-Aalen estimator will not be preferred over the Kaplan-Meier estimator.

Turnbull's estimator differs from the Kaplan-Meier estimator and the Nelson-Aalen estimator since this estimator is an algorithm. Turnbull estimates the survival function by using an iterative process. In this process the survival function gets more and more stabilized until there is nearly no difference between the last two iterations. Turnbull's estimator handles right censored data, but also left censored data. (Klein & Moeschberger, 2003)

2.2.3. Application in this research

This research focusses on the capacity of a service location. This capacity is expressed in a number of train units that can be handled. The number of train units, can be seen as the time in a lifetime analysis. An instance of 1 train unit for which a solution is found, will be extended with another train unit, until there is no solution found for the instance. At that number of train units, the instance fails and this was the maximum feasible 'lifetime' for this instance with these characteristic.

However, the data available in this research consists only of an instance with a particular number of train units, but the instance is not extended. There is only one measurement of an instance lifetime at a certain 'time', which is in this case a number of train units. The measurement at this moment can be a failure or a success. A failure means that no solution could be found. This comes down to the fact that the success should have taken place with less train units in the instance, which is left censored data. A success on the other hand implies that perhaps there could be more train units handled in the instance, but it is unknown how many. This means that successes are exact data or right censored data. A success at a number of train units implies that the number of train units at the success minus 1 is right censored.

Since the Product Limit Method is scientifically based and the intuitive method is not, the Product Limit Method will be used in this research. The estimator that will be used is Turnbull's estimator since the data in this research is doubly censored and this estimator handles right censored data and left censored data. This will cause less uncertainty than making assumptions while using the Kaplan-Meier estimator.

2.2.4. Turnbull's algorithm in Matlab

Turnbull's estimator makes use of an iterative process. To estimate the survival function in this research by using Turnbull's estimator, a script is written in Matlab, which executes Turnbull's algorithm and estimates the needed survival function. This can be used for constructing survival functions. A survival function can be used to show variation in capacity, which will be done in this chapter, but Turnbull's algorithm will be used for constructing many other survival functions later in this research.

The iterative process programmed in Matlab, starts with an initial estimate of the survival function S. Any legitimate estimate will work. Turnbull suggests to use the Kaplan-Meier estimator, but since in the data no exact measures are available, this will not work. Therefore, the intuitive method is used to construct an initial estimate of the survival function.

The first step of the iterative process, is calculating p_{ij} values using the current estimate of S. This can be done by using the following formula:

$$p_{ij} = \frac{S_k(t_{j-1}) - S_k(t_j)}{1 - S_k(t_i)} \text{ for } j \leq i \text{ where } S_k \text{ is the current estimate of the survival function.}$$

This results in an upper triangular matrix with i columns and j rows. This p_{ij} matrix will be used for executing step two. In step two, an estimate of the number of events at a particular number of train units t_i is made, which is $\dot{d}_i$. $\dot{d}_i$ can be calculated by the following formula:

$$\dot{d}_i = d_i + \sum_{i=j}^{m} c_i \, p_{ij} \text{ where } c_i \text{ is the number of left censored observations at } t_i.$$

Since $\dot{d}_i$ is an estimate for the events at a number of train units t_i that follow from the left censored data, only right censored data still has to be handled. This can be done by the Kaplan-Meier estimator, since left censored data is not an issue anymore. By executing the Kaplan-Meier estimator, a new estimate for the survival function arises. (Klein & Moeschberger, 2003)

The Kaplan Meier estimator gives an estimate of the survival function by calculating the following formula at every t (Brilon, Geistefeldt, & Regler, 2005):

$$S(t) = \prod_{j:\, t_j \leq t} \frac{n_j - d_j}{n_j} \text{ where } S(t) \text{ is the survival function,}$$
$$n_j \text{ is the number of train units for which the capacity can be reached at that point and}$$
$$d_j \text{ is the number of times that the capacity is reached at } t_j \text{ number of train units.}$$

The iterative procedure stops if the estimate $S_{k+1}(t)$ is close to $S_k(t)$. Then, not much differences will occur anymore. In the script it is assumed that the process stops when the sum of the differences is below 0,001. (Klein & Moeschberger, 2003)

The script can be found in appendix A.

2.3. Cumulative distribution function and the probability density function

The probability density function will be constructed to show the variation in capacity based on a large set of data. This probability density function can be derived from the cumulative distribution function, which can be found from the survival function.

A survival function is a decreasing function from 1 to 0, since the chance of success becomes lower over the time. A cumulative distribution function is an increasing function from 0 to 1. This function shows the chance that an event will happen at that time or has happened already. The chance of failure will increase while increasing the number of train units in a work package. The cumulative graph about the chance of failure can be found from the survival function, since the following holds:

Cumulative graph P(fail) = 1 – P(success) = 1 – Survival function S.

The probability density function can be derived from the cumulative distribution function. This includes the probability density function of failure. The probability density function is the derivative from the cumulative distribution function. Therefore, the probability density function of failure follows from the cumulative distribution function P(fail).

This results in Figure 3 are the cumulative distribution function (blue) and the probability density function (black), based on instances with a lot of different characteristics.

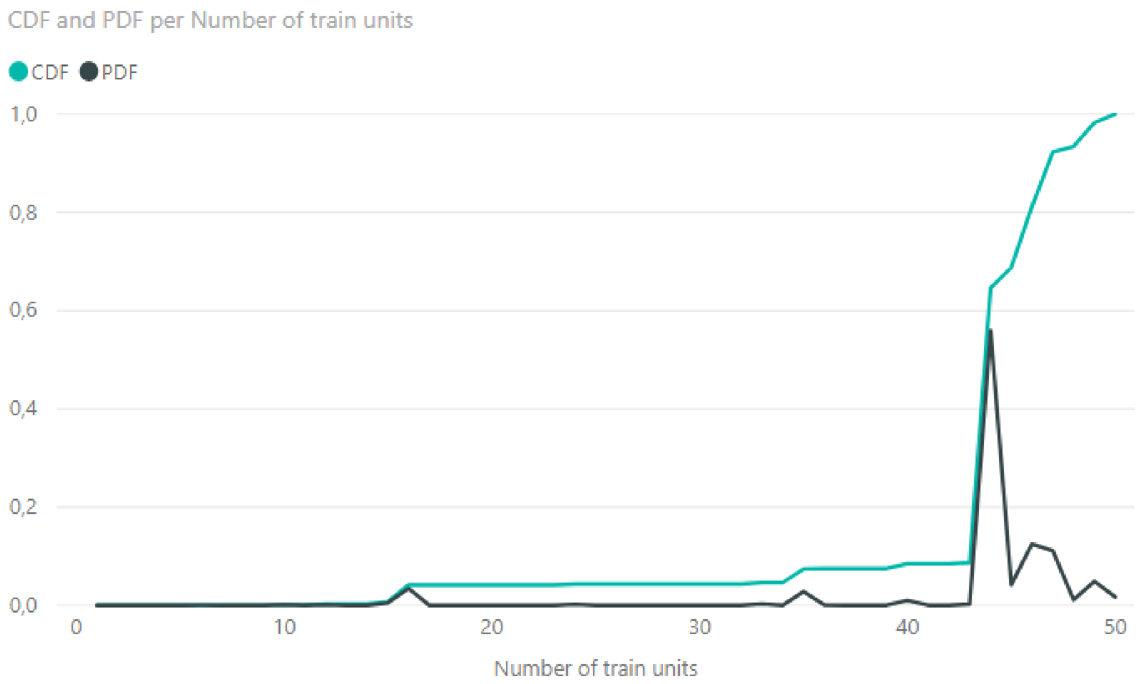

Figure 3: CDF and PDF per Number of train units of the database

In Figure 3 can be seen that the probability density function of the capacity is not a clear function, like a normal distribution. However, a lot of variation can be seen due to the different peeks. Therefore, variation in capacity exists.

2.4. Quantification of variation

By identifying factors that will have influence on the variation in capacity, the variation will decrease theoretically. This will be checked in the remainder of this research by giving a measure for the variation in capacity to a set of data. In this chapter is shown that variation in capacity exists, but to compare this variation with other sets of data which are based on specific characteristics, a measure for the degree of variation is needed.

The variation can be expressed as the variance, since the variance is a measure for the scatter of a set of values. The variance is a characteristic of a sample. Therefore, a sample will be generated based on the survival function shown in Figure 3. A sample with a sample size of 100.000 measures will be generated by a random generator. A script for making a sample with measures based on the survival function is made Matlab and can be found in appendix B. This sample will be used to calculate the variance of the dataset, by using the following formula:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - x_{avg})^2}{n} \text{ where } x_i \text{ is the value of observation i,}$$

$$x_{avg} \text{ is the mean of all values } s_i \text{ and n is the number of observations}$$

Executing this formula, based on the probability density function in Figure 3, the variance becomes $s^2 = 40{,}543$.

A source can also cause variation in capacity mean. In this research will be worked with the median of the datasets instead of the mean. This will be done since the distributions of the survival functions are not known. To test the samples on significant differences in mean, the assumption has to be made that the survival functions are normally distributed. This is a dangerous assumption. Therefore it is chosen to work with a more robust test, which tests significant differences between medians. The median is the central value of a sample. The median of this sample is $m = 43$.

Chapter 3
Selection of sources

In this chapter, a selection of the sources will be made. Those sources are going to be investigated in the remainder of this research. The selection will be made by first investigating the treatment process. Based on this, a couple of sources that potentially have influence on the capacity variation will come up. From these a selection will be made, substantiated with available instances in the data.

Some sources will interact with other sources. To track down those sources, a flow chart will be made of the sources that are going to be investigate and the potential interactions between them. The interaction between those sources will be explored by making scenario's and run those scenario's in the behandelcalculator. Therefore, scenario's will be chosen based on the flowchart, substantiated with available instances in the data.

3.1. Treatment process

The treatment process is investigated to select sources which potentially can cause variation in capacity. This treatment process starts with the arrival of a train and ends with the departure of a train at a service location. Between those events many things can happen to the train. Shunting movements are necessary for every time a train needs to be replaced. Some maintenance tasks need to be executed, at which some of the tasks need specific tracks. Tasks that can be necessary to execute are cleaning of the internal or external of the train, A- and B-checks or small repairs. Trains also need to depart in the right composition, which can differ from the arrival composition. In that case, the matching problem comes up.

3.1.1. Arrival process

A source that probably influences the capacity, is the arrival process. The way the trains arrive at service locations is important since there is a difference between evenly distributed arriving trains and a bunch of trains which arrive simultaneously and have to be cleaned checked and matched in a short period. Many trains at once will end up in a lower capacity than evenly distributed arriving trains throughout the arrival period.

3.1.2. Departure process

The departure process of the trains could also have influence on the capacity, since the moment of departure determines the time a train is present at the service location. The longer a train is present at a service location, the more time is available to execute all the tasks.

3.1.3. Shunting problem

The Train Units Shunting Problem is concerned with the assignment of tracks to trains. In this problem, it is important find a feasible solution to plan the arriving trains in such a way that they are ready to depart in the right composition, with the right maintenance executed. Finding a solution becomes more complicated when more tasks need to be executed, especially since for some tasks particular tracks are required.

Besides this, there is a lot of uncertainty. This uncertainty is caused by the fact that a shunting plan cannot be made a long time before the execution. Sometimes a train ends up somewhere else because of a problem at the rails for example, or maybe defect has occurred so a certain repair needs to take place. This uncertainty in terms of modelling means that the calculation time cannot be long, which results in less time to find a solution, so less solutions will be found. This comes down to a lower capacity. (van den Broek, Hoogeveen, & van den Akker, 2012) (Haahr, Lusby, & Wagenaar, 2017) (Freling, Lenting, Kroon, & Huisman, 2005)

3.1.4. Tasks

Many tasks need to be executed at service locations and the goal is to make use of the resources at service locations optimally. The tasks are divided into three groups: daily maintenance, short-term maintenance and refurbishment & overhaul. This research will only look at the daily maintenance at one service location. Daily maintenance takes place at 35 service location throughout the country. This consists of thorough cleaning of the inside and outside of trains, safety checks and minor repairs if necessary. For some of the tasks particular tracks are required which are not available at all service locations. (Apallius de Vos & Van Dongen, 2015) The different tasks are described below. (Hoepel, 2017) (Beerthuizen, 2017) (Huizingh, 2018)

3.1.4.1. Cleaning

The cleanness of rolling stock directly influences the perceived quality of the offered service to passengers. Providing clean rolling stock is one of the five main objectives of the NS. The train units need to be cleaned both inside and outside.

Internal cleaning

There are different types of internal cleaning. Cleaning at the end of a railway passenger line which consists of the fast cleaning of the interior of the train and emptying trash cans. Modular cleaning, which consists of all standard interior cleaning activities. Those are separated in different modules, which all have prescribed frequencies. Periodic thorough cleaning is typically scheduled once every few months at the same time as large maintenance activities. Modular cleaning is the most important internal cleaning type that is executed at service locations.

The process of internal cleaning takes place along dedicated platforms. If it is impossible to clean a train unit internally at a track along such a platform, one can consider cleaning a train unit at some other track. (Lentink, 2006)

External cleaning

External cleaning involves washing the outside of the train. This is less urgent than internal cleaning because it is secondary priority for passengers, but it needs to happen too. The process of external cleaning needs a train-wash, which is available at 15 service locations in the Netherlands. Because it requires specialized equipment, cleaning the outside of the trains cannot be carried out at other tracks. Sometimes urgent external cleaning needs to be executed, for example when graffiti needs to be removed from the front window or after incidents. (Lentink, 2006)

3.1.4.2. Checks

Different checks need to be executed, to prevent trouble. A distinction can be made between the checks. The A-check and the B-check. The A-check is a larger inspection and takes about an hour to execute. This inspection needs to take place approximately once every 12 days. The B-check is a shorter check and takes about 20 minutes to execute. This inspection needs to take place more often, namely once per two days. Inspection B consists of inspection preparation proceedings, brake testing proceedings and external checks on the streamers of the train unit. Inspection A consists of a B check but also of some extra safety checks like doors, entry steps and the train driver cabin. The contents of the A- and B-checks can vary per train type. (NedTrain, 2014)

3.1.4.3. Repairs

In the current production concept, two categories of train failures are known. Simple failures are handled nightly at the service locations. Complex failures are repaired at three large maintenance depots, but those repairs cost a lot of time and are very expensive. These repairs need to be prevented as much as possible, so it is important to carry out the minor repairs very seriously. (Busstra & van Dongen, 2015) The minor repairs can take place at the service locations. Examples of these minor repairs are door repairs or changing a light bulb. The main characteristic of minor repairs is that they can be executed at the service location. In some cases, special equipment is needed and that is available at special platforms, such as an aerial platform or a technical centre.

3.1.4.4. Required time and frequency of tasks

The time that the several tasks take has influence on the capacity, since the longer the tasks, the fewer tasks can be executed. However, this is not the only point. It can be that the duration of a task lies on a broad spectrum. This implies that the standard deviation of the duration of the tasks is large. The standard deviation of different tasks might have influence on the certainty of the occurrence of particular values for capacity.

The frequency of different tasks may also have influence on the capacity. If a task needs to be executed every two days instead of every day, the capacity will probably improve. However, requirements are made about this topic. A safety check for example has to be executed daily, so not just a frequency of once per two days can be set here.

3.1.5. Matching problem

The matching problem is about deciding when and where to decouple and couple which train units, to create the train unit configurations that are required for the departing trains.

As input for this matching problem, the planner receives a work package with planned arrivals and departures of all train services at the station under consideration. In addition, this timetable also prescribes the configuration of each train, which follows from the rolling stock circulation. More specifically, if the train configuration consists of different train types, the order of the different types of train units in the train is given by the timetable. In general, train units of the same type can be used interchangeably. However, the number of the train units are important, because it is important to know what maintenance has to take place for that particular train unit the next night at a service location. (Lentink, 2006) (Freling, Lentink, Kroon, & Huisman, 2002)

3.1.6. Train types

On the railway net different train types are deployed. These trains have to go through the maintenance system at service locations. Common train types are the SLT, the ICMm and the VIRM(m). It could be that at a service location the different train types also have influence on the capacity. (NS, 2018)

3.1.7. Layout of the service location

The layout on service locations varies. Good accessible service tracks will cause a more efficient use of tracks, so this could influence the capacity. Two different layout types are common. The carousel layout and the shuffleboard layout. (Beerthuizen, 2017) (Huizingh, 2018)

3.1.7.1. Carousel layout

The carousel layout is a service location layout at which most tracks are free track, which means that a train unit can enter and depart the track at both sides. The train enters the track at one side and departs at the other side, so the first-in-first-out (FIFO) system is used. A carousel route is created in this way. The Kleine Binckhorst has a carousel layout. In the figure below, the carousel layout at the Kleine Binckhorst is shown.

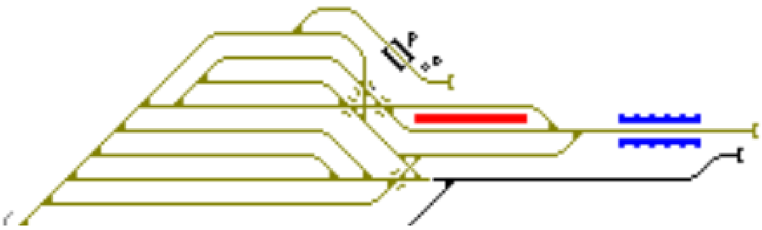

Figure 4: Carousel layout Kleine Binckhorst (Emplacementstekeningen van NS, 2018)

3.1.7.2. Shuffleboard layout

The shuffleboard layout is a service location layout at which most tracks can be entered from only one side. The other side had a dead end. This means that the last-in-first-out (LIFO) system is used at shuffleboard layouts. The Grote Binckhorst has a shuffleboard layout. In the figure on the right, the shuffleboard layout at the Grote Binckhorst is shown.

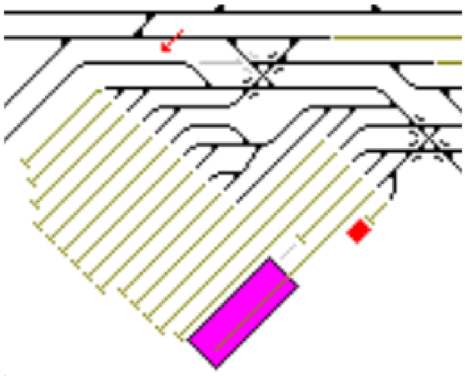

Figure 5: Shuffleboard layout Grote Binckhorst (Emplacementstekeningen van NS, 2018)

3.2. Choice of sources

In the description of the treatment process, some sources that potentially influence the capacity have come up. Some of those sources will be investigated and some will not to demarcate the research assignment. Below, for every potentially influencing source is substantiated whether and why it will be investigated in this research or not.

3.2.1. Arrival process and departure process

To decide whether the arrival process is interesting to investigate, the dataset used in chapter 2 is filtered. This is done by making a program in Matlab, given in appendix C. The dataset is filtered based on two different kinds of arrival processes. One process is gradually spread over the arrival period, the other process is a bunch of trains in the last 25 percentage of the arrival period. The arrival period is from 17.30h until 2:20h.

This results in differences in median and sigma compared to the total dataset. The results can be found in Table 1. Since the median and variance of the filtered samples differ from the median and variance of the total database, the arrival process is probably an interesting source.

Table 1: Median and variance results of the filtered database based on arrival processes

	Database	**Gradually arrival processes - Filtered**	**Arrival processes in the last 25% of the time - Filtered**
Median (m)	42	46	38
Variance (σ^2)	40,54	113,88	108,97

For the departure process could be done the same thing. However, it is chosen to leave the departure process out of the research scope since the arrival process and the departure process are closely linked. The total time that train units are present at service locations is dependent from both the arrival time and the departure time. This source will have more influence by presence of tasks theoretically, since the longer a train unit is present, the easier a plan can be made. On the other hand, sometimes it is preferred to get rid of a train unit, for example when a train unit is in the way for other train units. Since this has too many aspects, it is chosen to assume that in this research the arrival process and the departure process are of a similar nature.

3.2.2. Shunting process

The shunting process is an isolated problem. This cannot be investigated by the behandelcalculator since the behandelcalculator tries to find a solution for a problem. Shunting movements are necessary to execute a plan, so they are part of the solution. To improve the shunting process, the algorithm of the behandelcalculator needs to be improved.

3.2.3. Tasks

To choose whether the tasks are investigated in the remainder of this research the dataset used in chapter 2 is used. This dataset is filtered. In a part of those instances tasks were assigned to the train units and in a part of those instances no tasks were assigned. The database of chapter 2 is split up into two dataset, namely a dataset with the instances where tasks are assigned to the train units and a dataset in which no tasks were assigned. Those two different datasets resulted in the survival functions shown in Figure 6.

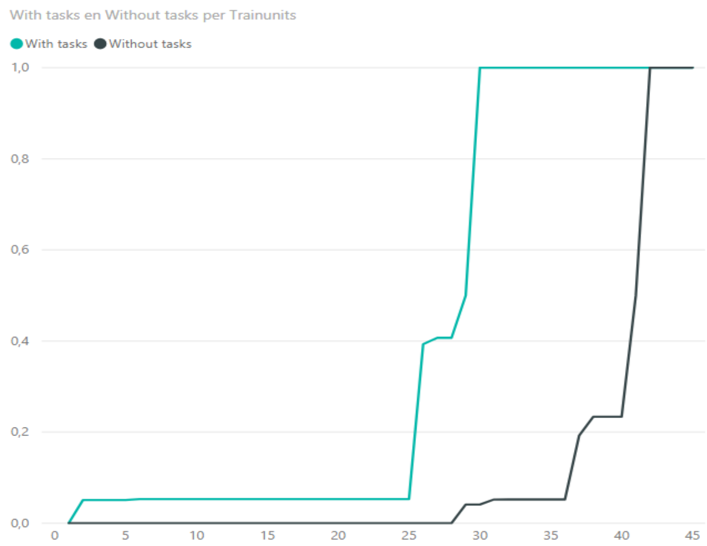

Figure 6: Survival functions with and without tasks (filtered data)

Those survival functions show that there is a large difference between the dataset where tasks were assigned to the train units and where they were not. Variation in capacity median and variance will also be very large due to the large differences in values. This large difference shows a clear influence. However, it is better to investigate other sources while not assigning tasks in this research since the other sources do not fluctuate due to the presence of tasks. In this way, a better picture of the other sources can be outlined.

3.2.4. Coupling and decoupling

Matching becomes more complicated when train compositions of arriving trains differ from the compositions of departing trains. Because more shunting movements will take place when coupling and decoupling is accepted, it is harder to find a solution for the algorithm. However, in practice, it happens a lot that this configurations differ from each other. This probably means that the capacity is dependent of this. Not allowing coupling and decoupling will be easier, so this will cause differences capacity compared to allowing coupling and decoupling. Differences between the influence will especially be interesting for the different locations. Therefore, this source will be investigated.

3.2.4. Train types

The different ratios in train types that arrive at a service location can be interesting, especially when coupling and decoupling is allowed. However, not all sources can be investigated, due to the available time period for this research. Therefore it is chosen that investigating the train types will not be in the scope of this research.

During this research only the standard ratio is train types will be used, since as little as possible factors at once will be set in the scenarios. Otherwise, the other factors can affect the capacity and differences can be caused by various sources.

3.2.5. Location layout

A little part of instances from the database used in chapter 2 are based on the topology of the Kleine Binckhorst and the other instances are based on the topology of the Grote Binckhorst. The Kleine Binckhorst has a carousel layout and the Grote Binckhorst has a shuffleboard layout. The dataset is split into two samples, one sample with the carousel topology and one sample with the shuffleboard topology. From these instances, survival functions are made. Since the Grote Binckhorst is roughly two times as large as the Kleine Binckhorst, capacity values, given in number of train units, of the Kleine Binckhorst are doubled. The survival functions are shown in Figure 7.

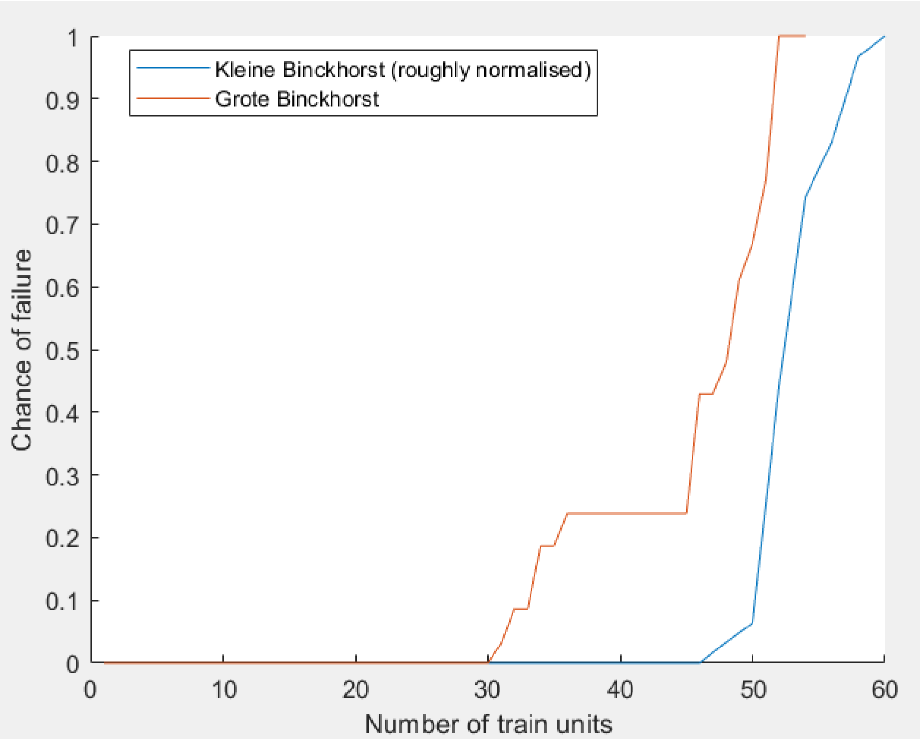

Figure 7: Survival functions Kleine Binckhorst and Grote Binckhorst (filtered data)

This figure shows that it is very likely that the location layout has a lot of influence on the capacity. In the table below, it also becomes clear that variation is probably caused by the location layout.

Table 2: Median and variance results of the filtered database based on topology

	Database	**Kleine Binckhorst (roughly normalised) - Filtered**	**Grote Binckhorst - Filtered**
Median (m)	42	52	46
Variance (σ^2)	40,54	8,26	46,13

Differences between the two types of topologies, shuffleboard and carousel, are expected especially in variance. The layout of a location can also interact with the arrival process or allowing coupling and decoupling. Therefore this source, and the interactions of this source with the other sources, will be investigated in this research.

Chapter 4
Influence of sources

In this chapter, the influence of the sources selected in the previous chapter will be investigated. Those sources are the arrival process, the tasks, the acceptance of coupling and decoupling and the location layout. These sources, except from the location layout itself, are going to be investigated at the Grote Binckhorst, which has a shuffleboard layout.

4.1. Methodology

To investigate the influence of the different sources, the settings of the behandelcalculator are modified for one variable while the remaining of the variables are kept constant. These settings generate data for corresponding scenarios. For the scenarios, survival functions are made, by making use of the same method as used in chapter 2.

The survival functions differ from each other, but the question is whether they differ significantly. To test whether the differences are significantly, a Kolmogorov-Smirnov test will be executed. This method tests, based on the survival distributions whether the samples can come from the same 'population' or not.

In case significant differences don't exists between survival functions, and thus the scenarios can be seen as the same, further investigation of those scenarios will not take place since the scenarios could have been the same due to the results. However, in case the survival functions differ significantly, capacity variation is caused by this source.

A source can influence the certainty of the capacity, which can be quantified by the standard σ or the variance deviation σ^2, but it can also influence the expected value of the capacity. In this research will be worked with the median as an indication for the expected value. To check how the source influences the capacity, the median value and the variance of the scenarios will be calculated. The median and variance of the capacity can be calculated from the instances that satisfy the conditions of the specific scenario. This will be compared with the median and the variance of the database which is used in chapter 2, for which the values are already calculated in chapter 2. To investigate whether the median and the variance are changed significantly, statistical tests will be executed. The Wilcoxon signed rank test can be used to test significant differences in the median and the chi-squared test can be used to test significant differences in variance and thus in standard deviation.

In case the variances of the samples from the scenarios differ significantly, this means that research on the certainty about the capacity will be interesting. To visualise this, boxplots will be made. In this way, the certainty about the capacity value in the scenarios will become more clear.

4.2. Statistical tests

Different statistical tests are chosen to execute to investigate the data on significant differences. The tests that are chosen are the Kolmogorov-Smirnov test to test whether survival functions are significantly different, the chi-squared test to test variance on significant differences and the Wilcoxon rank sum test to test sample medians on significant differences. The reason these tests are chosen to use in this research is that they are all non-parametric tests. For non-parametric tests, it is not necessary to approximate a sample with a known distribution.

4.2.1. Kolmogorov-Smirnov test

To test whether the two survival functions are significantly different, different statistical tests can be used to check whether two independent samples are drawn from the same population or not. A non-parametric test with these abilities, is the Kolmogorov-Smirnov test. For this test, two cumulative distribution function are compared which can be calculated by 1 – Survival function. From these cumulative distributions, the largest difference $D_{n,n'}$ between the values needs to be calculated. The larger the difference between the cumulative percentages of both distributions, the larger the probability that both distributions are not the same. Therefore, the KS-test needs to be executed.

The null hypotheses in the Kolmogorov-Smirnov test states that both distributions are drawn from the same population distribution, so they do not differ significantly from each other.

$$H_0\text{: Sample A and sample B are drawn from the same distribution.}$$

To test whether the null hypothesis can be accepted or needs to be rejected, the KS-statistic $D_{n,n',\alpha}$ can be calculated.

$$D_{n,n',\alpha} = c(\alpha)\sqrt{\frac{n+n'}{nn'}} \quad \text{where} \quad c(\alpha) = \sqrt{-\frac{1}{2}\ln\left(\frac{\alpha}{2}\right)}$$

In this formula, n represents the number of first sample and n' represents the number of the second sample. The parameter α is the level of significance. The outcome of $D_{n,n',\alpha}$ is the probability that D is coincidental.

The null hypothesis must be rejected at significance level α if the following formula holds.

$$D_{n,n'} > c(\alpha)\sqrt{\frac{n+n'}{nn'}}$$

Otherwise the null hypothesis can be accepted at significance level α.
(Non-parametric tests and regression, 2018)

Unfortunately, this method is not directly applicable in this research, since only the values of the survival function are known due to the censored data. There is not just a sample with measures available while this is needed. However, this sample can be generated by using a random generator. A script for making a sample with measures based on the survival function is made Matlab and can be found in appendix B. This will be used for the χ^2-test and the Wilcoxon signed rank test as well. The sample size used in this research will be 100.000 measures.

4.2.2. χ^2-test

A χ^2-test is used to test whether the variances of two samples are equal. This will be the null hypothesis.

$$H_0\text{: The variances of sample A and B are equal.}$$

This test can be a two-tailed test, which tests against the alternative hypothesis that the variances are not equal. The test can also be lower or upper one-tailed. This will test against the alternative hypothesis that the variance is significantly lower or higher respectively. In this research, only two-tailed tests will be performed. This means that the following will be the alternative hypothesis.

$$H_1: \sigma_1{}^2 \neq \sigma_2{}^2$$

To test whether the null hypothesis needs to be rejected or accepted, the test-statistic χ^2_0 needs to be calculated by using the following formula.

$$\chi^2{}_0 = \frac{(n-1)\sigma_1{}^2}{\sigma_2{}^2} \qquad \text{where n is the number of degrees of freedom}$$

The null hypothesis needs to be rejected if the value χ^2_0 is in the rejection region. The rejection region is described below. The critical values need to be read out the table given in appendix D. They are dependent from the number of degrees of freedom n and the level of significance α.

$$\chi^2{}_0 \geq \chi^2{}_{n-1;\alpha/2} \text{ or } \chi^2{}_0 \leq \chi^2{}_{n-1;1-\alpha/2}$$

(van Berkum & Di Bucchianico, 2007)

4.2.3. Wilcoxon signed rank test

The Wilcoxon signed rank test is a nonparametric alternative for the paired student's t-test. The paired student's t-test compares the means of two paired samples, however it assumes that the distribution of the samples are normally distributed. The Wilcoxon signed rank does not assume a distribution for the sample since it is a nonparametric test. This means that the Wilcoxon signed rank test is more robust than the paired student's t-test. On the other hand, the Wilcoxon signed rank test does not compare the means, but instead the medians of the samples. However, this this will still give a good indication.

The null hypothesis of the Wilcoxon signed rank test states that the medians of both samples are equal.

$$H_0\text{: Sample A and sample B do have the same median.}$$

To test whether this statement is true or whether it has to be rejected, the test-statistic W needs to be calculated. To calculate the test-statistic W, the absolute differences between the samples have to be calculated first. There are N pairs of which the absolute differences are calculated. From this set, the ones where the absolute difference is zero are excluded . The remaining N_r pairs are ranked from low to high. These ranks R_i are used to calculate the test-statistic W, which is the sum of the signed ranks. The sign function

$$W = \sum_{i=1}^{N_r} [\operatorname{sgn}(x_{2,i} - x_{1,i}) \cdot R_i]$$

In the case that $N_r < 20$ the absolute value of the test-statistic W can be compared to a critical value from a reference table. Since in this research $N_r \geq 20$, the z-score can be used. This can be calculated by the following formula.

$$x = \frac{W}{\sigma_W} \quad \text{where} \quad \sigma_W = \sqrt{\frac{N_r(N_r+1)(2N_r+1)}{6}}$$

Only two sided tests will be performed in this research. To perform a two-sided test the following holds.

$$\text{Reject } H_0 \text{ if } |z| > z_{critical}.$$

(van Berkum & Di Bucchianico, 2007)

4.3. Arrival process

The first source that will be investigated, is the arrival process. There will probably be differences between evenly distributed arriving trains and a bunch of trains which arrive simultaneously or shortly after each other. Many trains at once will probably end up in a lower capacity than evenly distributed arriving trains throughout the day.

4.3.1. Scenario description

To investigate differences capacity caused by the arrival process, four scenarios are set and investigated. In the behandelcalculator, the train units have to arrive between 17.30 and 2:20. The scenarios differ in the percentages of the arrival period in which the train units arrive. The scenarios are all based on two runs of the behandelcalculator, once for the Kleine Binckhorst and once for the Grote Binckhorst. An overview of the scenarios is given in Table 3.

Table 3: Composition of the scenarios set to investigate the arrival process

Scenario's	Runs	Assumptions
Train units arrive spread over the last 25% of the arrival period	Topology of the Kleine Binckhorst is used.	Coupling and decoupling is not allowed and no tasks need to be executed.
	Topology of the Grote Binckhorst is used.	Coupling and decoupling is not allowed and no tasks need to be executed.
Train units arrive spread over the last 50% of the arrival period	Topology of the Kleine Binckhorst is used.	Coupling and decoupling is not allowed and no tasks need to be executed.
	Topology of the Grote Binckhorst is used.	Coupling and decoupling is not allowed and no tasks need to be executed.
Train units arrive spread over the last 75% of the arrival period	Topology of the Kleine Binckhorst	Coupling and decoupling is not allowed and no tasks need to be executed.
	Topology of the Grote Binckhorst is used.	Coupling and decoupling is not allowed and no tasks need to be executed.
Train units arrive spread over 100% of the arrival period	Topology of the Kleine Binckhorst is used.	Coupling and decoupling is not allowed and no tasks need to be executed.
	Topology of the Grote Binckhorst is used.	Coupling and decoupling is not allowed and no tasks need to be executed.

4.3.2. Qualitative analysis

From the data that is obtained by running the behandelcalculator for the settings described in Table 3, four scenarios are made. The scenarios exists of two analyses of the behandelcalculator. To compose the survival function of two analyses together, the results are added together. This is done by making use of the Matlab script given in appendix E. The obtained survival functions are shown in Figure 8.

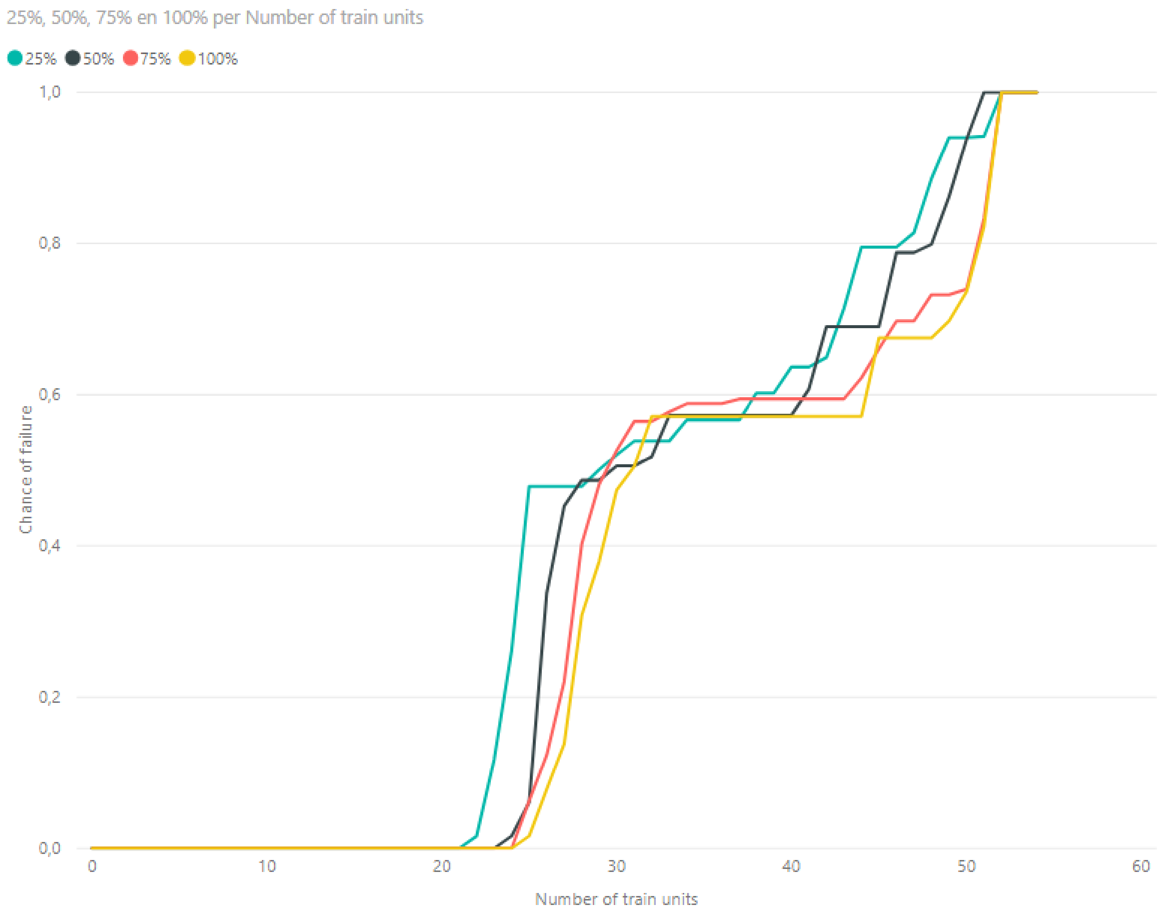

Figure 8: Survival functions of the arrival processes

At first, it is striking that all four survival functions have roughly the same shape. The first part of the survival functions are very steep, while the second part is much more divided. The number of train units are not extremely different for the four scenarios. Significant difference are not clearly determined. However, the survival functions achieve slightly higher values for capacity in its entirety by increasing the arrival period. This becomes clear when an estimate of the medians is made. The median is the central measure of a sample, which strokes with the capacity value where the chance of failure is 0,5. Estimations of the medians are shown in Table 4. By increasing the arrival period, the median becomes a little bit higher.

The certainty of capacity does not seem to be different for the different scenarios. The number of train unit values of the first failure and the last success of al scenarios, which can be found in Table 4, are pretty close together. This would imply that the arrival process does not cause variation in capacity.

The first parts of the survival functions are very steep. By increasing with only a few train units, the chance of failure increases from zero to 0,5. This means that the certainty on capacity is very high for the instances for which in at least 50% of the cases a solution will be found. When the chance of failure becomes even larger, the certainty on capacity the capacity value becomes much smaller. This becomes very clear in the sizes of the intervals of the values with a failure chance between 0,1 and 0,5 compared to the sizes of the intervals of the values with a failure chance between 0,5 and 0,9. Estimates of those intervals can be found in Table 4.

Table 4: Estimates based on the survival functions with different arrival processes

Scenario	**Median**	**First failure**	**No successes**	**Interval of the values between 0,1 and 0,5 chance of failure**	**Interval of the values between 0,5 and 0,9chance of failure**
Arrivals spread over the last 25% of the arrival period	24 train units	22 train units	52 train units	[23, 25] train units	[25, 48] train units
Arrivals spread over the last 50% of the arrival period	27 train units	24 train units	51 train units	[25, 28] train units	[28, 49] train units
Arrivals spread over the last 75% of the arrival period	29 train units	24 train units	52 train units	[25, 29] train units	[29, 51] train units
Arrivals spread over 100% of the arrival period	31 train units	25 train units	52 train units	[26, 30] train units	[30, 51] train units

4.3.3. Quantitative analyses

To execute the quantitative analysis, the statistical tests will be executed between the scenarios. The certainty on the capacity will also be investigated by studying the boxplots.

4.3.3.1. Executing the Kolmogorov-Smirnov test

At first, it is important to test whether the significant differences exists between the four scenarios. Therefore, Kolmogorov-Smirnov tests will be executed between all scenarios. This test will be executed by using the Matlab scripts given in appendix A to get the values of the survival functions and appendix F to test whether two scenarios differ significantly in capacity. In the qualitative analysis in paragraph 4.3.2 is already mentioned that the four scenarios do not differ obviously. Therefore, this test will be interesting. The null hypothesis that will be tested with a significance level of α = 0,01 is stated below.

$$\text{H}_0\text{: The two scenarios, in which the train units arrive spread over different percentages of the arrival period, do have the same capacity.}$$

Table 5: Results of the Kolmogorov-Smirnov test between the scenarios

	Arrivals spread over the last 25% of the arrival period	**Arrivals spread over the last 50% of the arrival period**	**Arrivals spread over the last 75% of the arrival period**	**Arrivals spread over 100% of the arrival period**
Arrivals spread over the last 25% of the arrival period	-	Significant differences exist	Significant differences exist	Significant differences exist
Arrivals spread over the last 50% of the arrival period		-	Significant differences exist	Significant differences exist
Arrivals spread over the last 75% of the arrival period			-	Significant differences exist

By executing the test, it appeared that between all scenarios, null hypothesis needs to be rejected. That means that the alternative hypothesis, stated below, will be accepted.

$$\mathrm{H_1}: \text{The two scenarios, in which the train units arrive spread over different percentages of the arrival period, do have significantly different capacities.}$$

Since significantly differences exist between the two scenarios, investigation of the nature of these differences is interesting.

4.3.3.2. Testing significant differences in median and variance

The sources can influence the capacity by mean value, for which in this research the median m is used as an indication, but it can also influence the certainty of the capacity. The certainty of the capacity is higher if the standard deviation σ is low, and thus if the variance σ^2, is low. The right estimate of the capacity becomes less certain when the standard deviation σ, and thus the variance σ^2, are higher.

In the table below, the values for the median and the variance of the two scenarios are given. Those values are calculated based on the sample generated by the random generator in appendix B.

Table 6: Results of the median and variance for the different arrival processes

	Median (m)	**Variance (σ^2)**
Arrivals gradually spread over the last 25% of the arrival period	28	112,59
Arrivals gradually spread over the last 50% of the arrival period	29	102,70
Arrivals gradually spread over the last 75% of the arrival period	29	117,74
Arrivals gradually spread over 100% of the arrival period	30	114,34

In this table can be seen that the median increases while the period in which the trains arrive increases. To check whether the differences in medians of the samples are significantly or not, Wilcoxon signed rank tests will be executed with a significance level of α = 0,01. The null hypothesis is stated below.

$$H_0: m_{scenarioA} = m_{scenarioB}$$

The test is performed using the Matlab script in appendix F. It appeared that the null hypothesis does not have to be rejected for all scenarios. Only the scenario in which the trains arrive in the last 25% of the arrival period compared to the scenario in which the trains arrive gradually spread over 100% of the arrival period, a significant difference exists. This means that the alternative hypotheses will be accepted for only those scenarios. The alternative hypothesis that will be accepted is stated below.

$$H_1: m_{25\%} \neq m_{100\%}$$

It can be stated that the different arrival processes cause variation in capacity median only when there is a large difference in the percentage of the arrival period the train units arrive in.

The values for the variances differ per scenario. With χ^2-tests will be tested whether the variances of the sample differ significantly. Those tests will be performed with a level of significance of α = 0,01. The null hypothesis is stated below.

$$H_0: {\sigma_{scenarioA}}^2 = {\sigma_{scenarioB}}^2$$

The tests are executed by making use of the Matlab script in appendix F. It appeared, as expected, that the null hypothesis has to be rejected. This means that the alternative hypothesis will be accepted. The alternative hypothesis is stated below.

$$H_1: {\sigma_{scenarioA}}^2 \neq {\sigma_{scenarioB}}^2$$

This means that the variances σ^2 of the all scenarios differ significantly. Based on the values in Table 6, not a really interesting relation can be expected with the percentage of the arrival period in which the train units arrive.

4.3.3.3. Certainty on capacity

Since the variances differ significantly, it is interesting to investigate the degree of uncertainty in capacity. This can be visualised by making a boxplot from the samples of the scenarios. The boxplots of the four scenarios in which the arrival processes differ are shown in Figure 9.

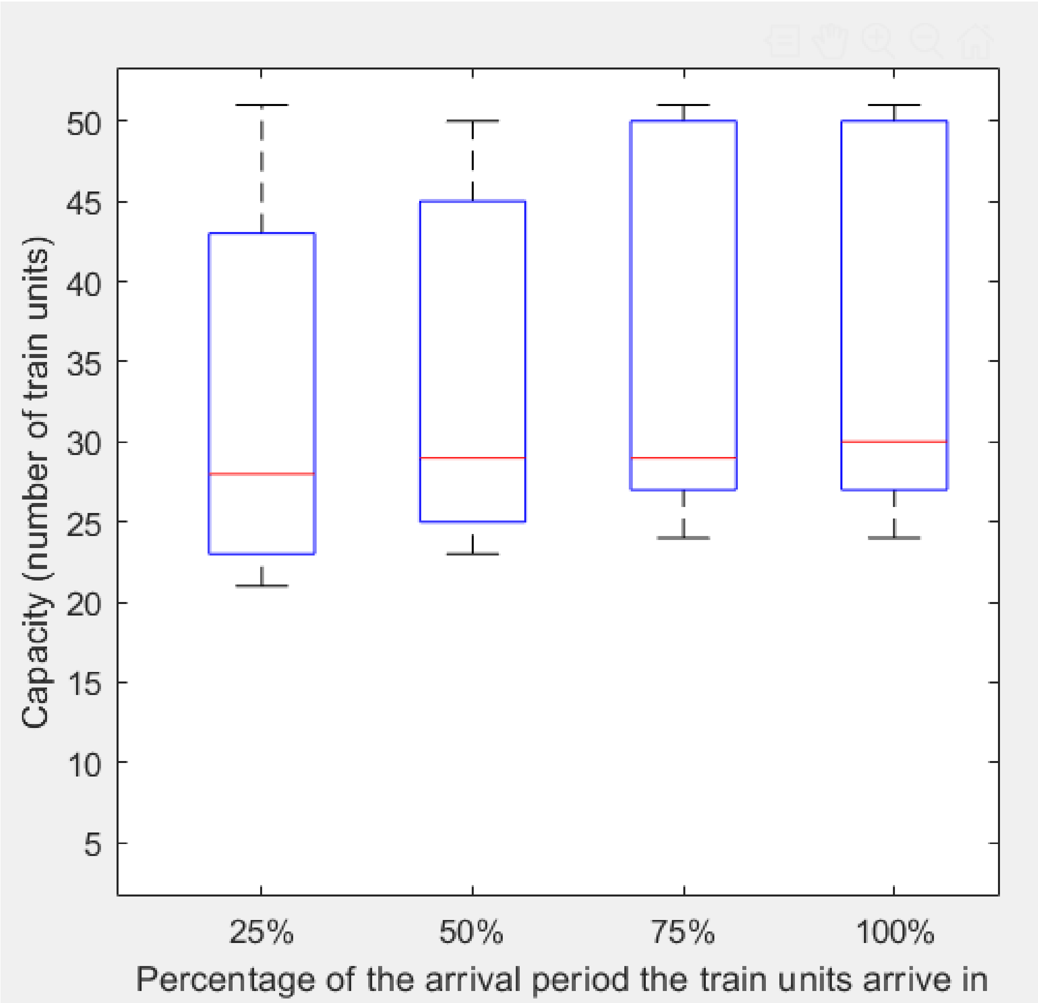

Figure 9: Boxplots of the scenarios with different arrival process

In this figure is visualised that the intervals in which values occur, become slightly smaller when the arrival period increases. However, these values are pretty close together. The intervals are equal for 50%, 75% and 100% of the time, but it is larger for 25%.

Besides that, it can be seen that the values increase, together with their intervals, by increasing the percentages of the arrival period in which the train units arrive. The largest differences are between 25% and 100%. This strokes with the results from the Wilcoxon singed rank test executed in paragraph 4.3.3.2.

4.3.4. Interim conclusion – arrival process

The conclusion that can be drawn from the investigation of the arrival process, is that the median values do not differ significantly for different percentages in which the train units arrive, only for the two extreme scenarios. This means that this source causes only little bit variation in capacity median. Increasing the time period in which the train units arrive, the capacity median increases.

The χ^2-test has shown that significant differences exist between the variances. The differences in sample variances imply that increasing the time period in which the train units arrive causes variation in capacity variance.

Compared to other sources, the variance causes a lot of variation in capacity variance. This becomes clear while comparing the variances σ^2 calculated in Table 6 with the variance of the database, calculated in chapter 2. It is not clear why the large variances has occurred. The certainty about the capacity is very small, since the capacity values occur over an interval of 27 train unit values. It is desirable to be more certain of the capacity.

4.4. Coupling and decoupling

The first source that will be investigated, is the allowing coupling and decoupling of trains. In the behandelcalculator can be chosen whether coupling and decoupling is allowed. If coupling and decoupling is not allowed, the trains will depart in the same configuration as they arrived in. If coupling and decoupling is allowed, the trains can depart in different configurations as they arrived in. This can make it harder to find solutions for the work packages.

4.4.1. Scenario description

To investigate allowing coupling and decoupling, two scenarios are composed. The scenarios are both based on two runs of the behandelcalculator. The compositions of the scenario's is clearly displayed in

Table 7 below.

Table 7: Composition of the scenarios set to investigate allowing coupling and decoupling

Scenario's	Runs	Assumptions
Allowing coupling and decoupling	Kleine Binckhorst at which coupling and decoupling is allowed.	The arrival process of the trains is set as the standard 'normal capacity' and no tasks need to be executed.
	Grote Binckhorst at which coupling and decoupling is allowed.	The arrival process of the trains is set as the standard 'normal capacity' and no tasks need to be executed.
Not allowing coupling and decoupling	Kleine Binckhorst at which coupling and decoupling is not allowed.	The arrival process of the trains is set as the standard 'normal capacity' and no tasks need to be executed.
	Grote Binckhorst at which coupling and decoupling is not allowed.	The arrival process of the trains is set as the standard 'normal capacity' and no tasks need to be executed.

4.4.2. Qualitative analysis

From the data that is obtained by running the behandelcalculator as described in

Table 7, two scenarios are made. To compose the survival functions of two analyses of the behandelcalculator together, the results are added together. This is done by making use of the Matlab script given in Appendix A. The obtained survival functions are shown in Figure 10.

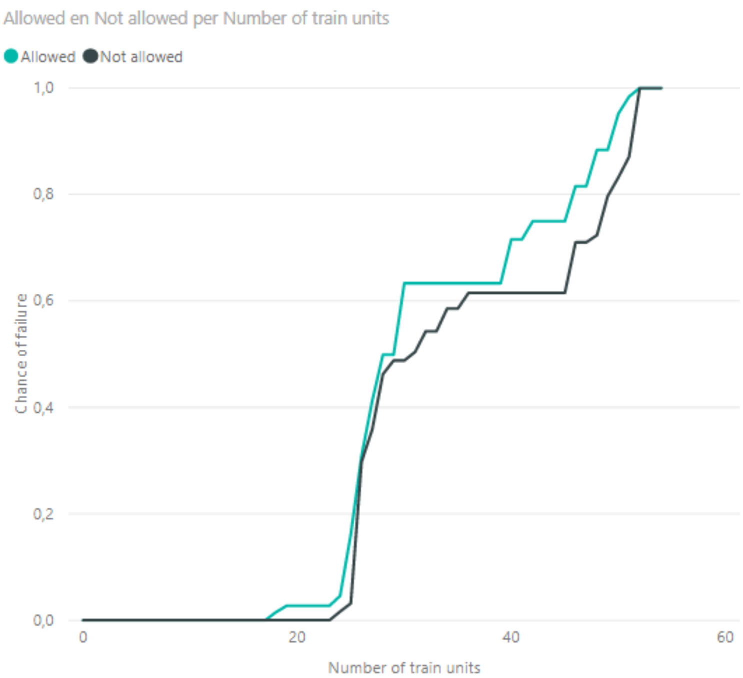

Figure 10: Survival functions allowing and not allowing coupling and decoupling

In this survival functions can be seen that they have roughly the same shape. In addition, the values for the capacity do not differ a lot. However, the survival function of the scenario where coupling and decoupling was not allowed, achieves slightly higher values for capacity in its entirety. This implies that the capacity when coupling and decoupling is not allowed, will be slightly higher. A higher capacity value was expected, since finding a solution becomes harder when the trains can depart in different configurations than they arrived in. Since there is not a large difference between the survival functions, the question is whether this will yield significant differences.

However, in the figure can also be seen that in case coupling and decoupling is allowed, the first failures occur at instances of 18 train units. The first failures in case coupling and decoupling is not allowed only occur at instances with 24 train units. On the other hand, the chance of failure becomes 1 at a train unit value of 52 in both scenarios.

The median is the central measure of a sample, which strokes with the capacity value where the chance of failure is 0,5. This means that for the scenario where coupling and decoupling is allowed the value for the median is estimated at 30 train units and for the scenario in which coupling and decoupling is not allowed this value is also estimated at 30 train units.

The certainty about the capacity is not clearly different in the two scenarios. When coupling and decoupling is allowed, the capacity values which have a chance of failure between 0,1 and 0,9 lie in the interval of 24 to 49 train units. When coupling and decoupling is not allowed, this interval is from 25 to 51 train units. This does not differ a lot relative to each other. However, generally seen, more certainty on the capacity is desirable. The standard deviations of both survival functions are probably large.

4.4.3. Quantitative analysis

To execute the quantitative analysis, the statistical tests will be executed. The certainty on the capacity will also be investigated by studying the boxplots.

4.4.3.1. Executing the Kolmogorov-Smirnov test

At first, it is important to test whether the two scenarios differ significantly from each other. This will be tested by using the Matlab scripts given in appendix A to get the values of the survival functions and appendix F to test whether the scenario in which coupling and decoupling is allowed and in which it is not allowed differ significantly in capacity. In the qualitative analysis in paragraph 4.4.2 is already mentioned that the two scenarios do not differ obviously. Therefore, this test will be interesting. The null hypothesis that will be tested with a significance level of α = 0,01 is stated below.

$$\text{H}_0\text{: The two scenarios, allowing the possibility of couping and decoupling and not allowing the possibility of coupling and decoupling, do have the same capacity.}$$

By executing the test, it appeared that the null hypothesis needs to be rejected. That means that the alternative hypothesis, stated below, will be accepted.

$$\text{H}_1\text{: The two scenarios, allowing the possibility of coupling and decoupling and not allowing the possibility of coupling and decoupling, do have significantly different capacities.}$$

Since significantly differences exist between the two scenarios, investigation of the nature of these differences might be useful.

4.4.3.2. Testing significant differences in median and variance

The sources can influence the capacity by mean value, for which in this research the median m is used as an indication, but it can also influence the certainty of the capacity. The certainty of the capacity is higher if the standard deviation σ is low, and thus if the variance σ^2, is low. The right estimate of the capacity becomes less certain when the standard deviation σ, and thus the variance σ^2, are higher.

In the table below, the values for the median and the variance of the two scenarios are given. Those values are calculated based on the sample generated by the random generator in appendix B.

Table 8: Results of the median and variance when coupling an decoupling is or is not allowed

	Coupling and decoupling not allowed	**Coupling and decoupling allowed**
Median (m)	30	29
Variance (σ^2)	117,28	98,08

The results in Table 8 show that the medians of the samples are about the same, but the variance of the samples differ a lot. To check whether the difference in medians of the samples is significantly or not, a Wilcoxon signed rank test will be executed with a significance level of α = 0,01. The null hypothesis is stated below.

$$\text{H}_0\text{: m}_{\text{allowed}} = \text{m}_{\text{not allowed}}$$

The test is performed using the Matlab script in appendix F. It appeared that the null hypothesis cannot be rejected. This means that the difference in medians is not large enough to state that allowing or not allowing coupling and decoupling causes variation in capacity median. The medians of the two scenarios can be considered equal.

To test whether the variances of the sample differ significantly, a χ^2-test will be performed with a level of significance of α = 0,01. The null hypothesis is stated below.

$$H_0: \sigma_{\text{allowed}}{}^2 = \sigma_{\text{not allowed}}{}^2$$

The test is performed using the Matlab script in appendix F. It appeared, as expected, that the null hypothesis has to be rejected. This means that the alternative hypothesis will be accepted. The alternative hypothesis is stated below.

$$H_1: \sigma_{\text{allowed}}{}^2 \neq \sigma_{\text{not allowed}}{}^2$$

This means the variances σ^2 of the two scenarios differ significantly.

4.4.3.3. Certainty on capacity

Since the variances differ significantly, it is interesting to investigate the degree of uncertainty in capacity. This can be visualised by making a boxplot from the samples of the scenarios. The boxplots of the two scenarios, one in which coupling and decoupling is not allowed and one in which coupling and decoupling is allowed, are shown in Figure 11.

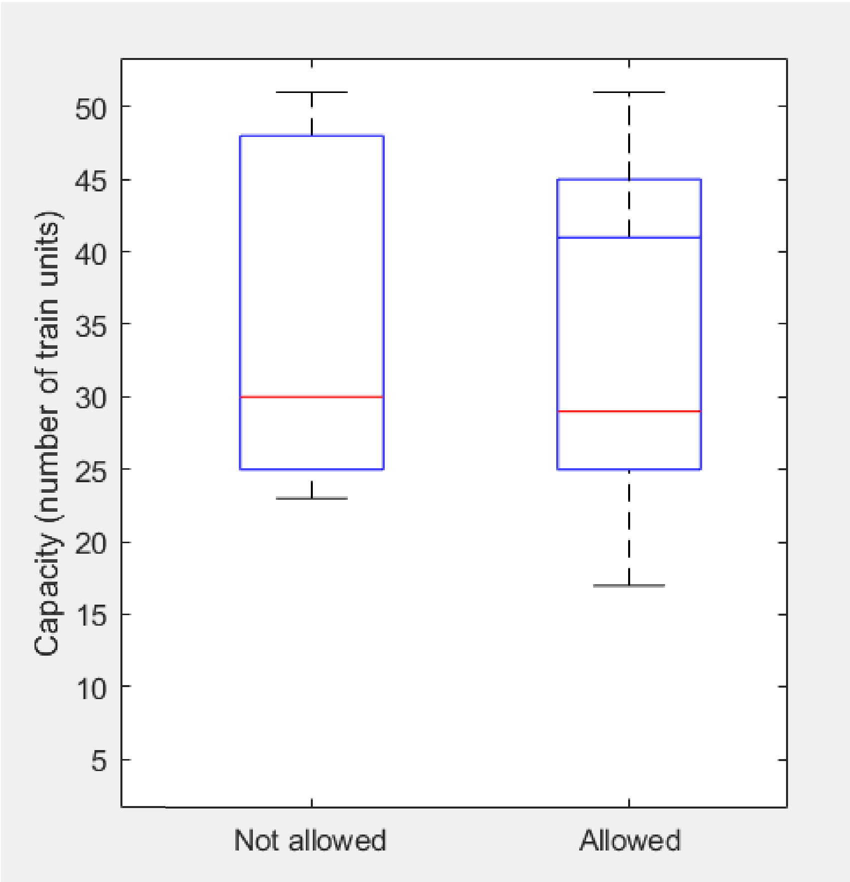

Figure 11: Boxplots of the capacity values of samples where coupling and decoupling is allowed and where it is not

In this figure is visualised that the interval in which values occur is larger when allowing coupling and decoupling in contrast to not allowing coupling and decoupling. However, the most values occur between 25 and 48 for when coupling an decoupling is not allowed, and between 25 and 45 when coupling is allowed. This is indicated by the blue area.

Based on this blue area, the interval of the scenario where coupling and decoupling is allowed, is smaller than the interval of the scenario where coupling and decoupling is not allowed. This means that while allowing coupling and decoupling most values will be found in a smaller area than while not allowing coupling and decoupling. This comes down to more certainty on capacity and strokes with the lower variance value, shown in Table 8. However, outliers are more extreme for allowing coupling and decoupling compared to the scenario in which coupling and decoupling was not allowed.

4.4.4. Interim conclusion – Coupling and decoupling

The conclusion that can be drawn from the investigation of allowing coupling and decoupling, based on the two scenarios described in paragraph 4.4.1, is that the median values do not differ for allowing or not allowing coupling and decoupling. It was expected that this should differ, since it becomes harder to find a solution when trains depart in different configurations than they arrived in. This was not the case. An explanation for this can be that calculation time, which was set as 120 seconds per instance, is amply sufficient in both scenarios.

The differences in sample variances imply that allowing coupling and decoupling is a source of variation in capacity variance and certainty on capacity. However, the certainty about the capacity is very low. It is desirable to decrease the interval in which values occur. The medians of the two scenarios are not significantly different, coupling and decoupling is not a source of variation in capacity median.

4.5. Topology

The topology is the last source that is going to be investigated in this research. The layout of a service location is important since different topologies can cause differences in capacity due to a better or worse logistic railway system. In this research, there are two different service locations with two different topologies are compared. The 'Kleine Binckhorst' and the 'Grote Binckhorst', which have a carousel layout and a shuffleboard layout respectively. However, these location differ also in size which makes that the two locations are to directly comparable. To compare the different locations a normalisation measure will be set. After that, the two locations will be compared by keeping all other settings constant.

4.5.1. Normalisation measure

To compare two locations with each other, the results for the capacity are not directly comparable since the sizes of the locations differ. If one location is larger than the other one, it is evident that that larger location end with a higher capacity. This is not necessarily caused by the location layout. To compare the topologies of the locations, two potential approaches come up. Comparison based on the maximum amount of trains that can be placed at a location, or comparison based on the length of railway available at the locations.

Determining a normalisation measure based on the maximum amount of train units that can be placed at every location would be a good approach. In this way, the cutting loss will be taken into account. The cutting loss is the remaining railway that cannot be used because no train can fit there. However, this maximum amount of train units cannot easily be established. Different train types come at the service locations and those train types differ in length. This makes that the maximum amount of train units that can be placed at a location cannot just be determined by hand.

An indication could be given by making use of the behandelcalculator. By executing a so called 'upper bound calculation', the behandelcalculator only checks the constraints which includes the constraint that the railway length does not exposes the length of train available at ones. It is not possible that some train units depart before other train units arrive, since departure process starts when the arrival process ends. This causes that, if the instance passed the constraint checker, the number of train units in an instance strokes with number of train units present at a service location at once in that particular instance. By making instances of every number of train units from 1 till 100 for example, it can be seen for which instances a solution is found, an thus how many train units fit at that service location. However, the constraint checker checks the instance for other things also. This causes that if an instance does not pass the constraint checker, it is not necessarily because they do not fit. It is possible that the constraint checker fails due to other causes. This ensures that this method gives no certainty about the maximum amount of train units that can be placed at the location. Therefore it is chosen to use the railway length available at a location as the normalisation measure.

The railway length available at a location gives a good indication of the size of a service location. The railway length at a location can be found in the behandelcalculator which contains a table with all track lengths. Since not every track can be used for positioning train unit, the tracks that do not allow this are excluded in this calculation. These are tracks which are meant for executing particular tasks or which are meant for sawing.

By making this calculation for the Kleine Binckhorst and the Grote Binckhorst the following results have come up.

Table 9: Useful track length per location

Location	**Useful track length**
Kleine Binckhorst	4687m
Grote Binckhorst	9088m

This means that the Grote Binckhorst is 9088/4687 = 1,94 times as large as the Kleine Binckhorst. The normalisation measure for the Kleine Binckhorst will be factor 1,94.

4.5.2. Scenario description

In this part of the research, two scenarios are compared. In those scenarios no tasks are assigned to the train units in the instances and the normal capacity is used for the arrival process. All other settings are standard settings, except that another topology is used. In the first scenario the Kleine Binckhorst is used and in the second scenario the Grote Binckhorst is used. The Kleine Binckhorst is a carousel layout and the Grote Binckhorst is a shuffleboard layout.

4.5.3. Qualitative analysis

From the data that is obtained by running the behandelcalculator with the two scenarios as described in paragraph 4.5.2 the survival functions shown in Figure 12 are gained.

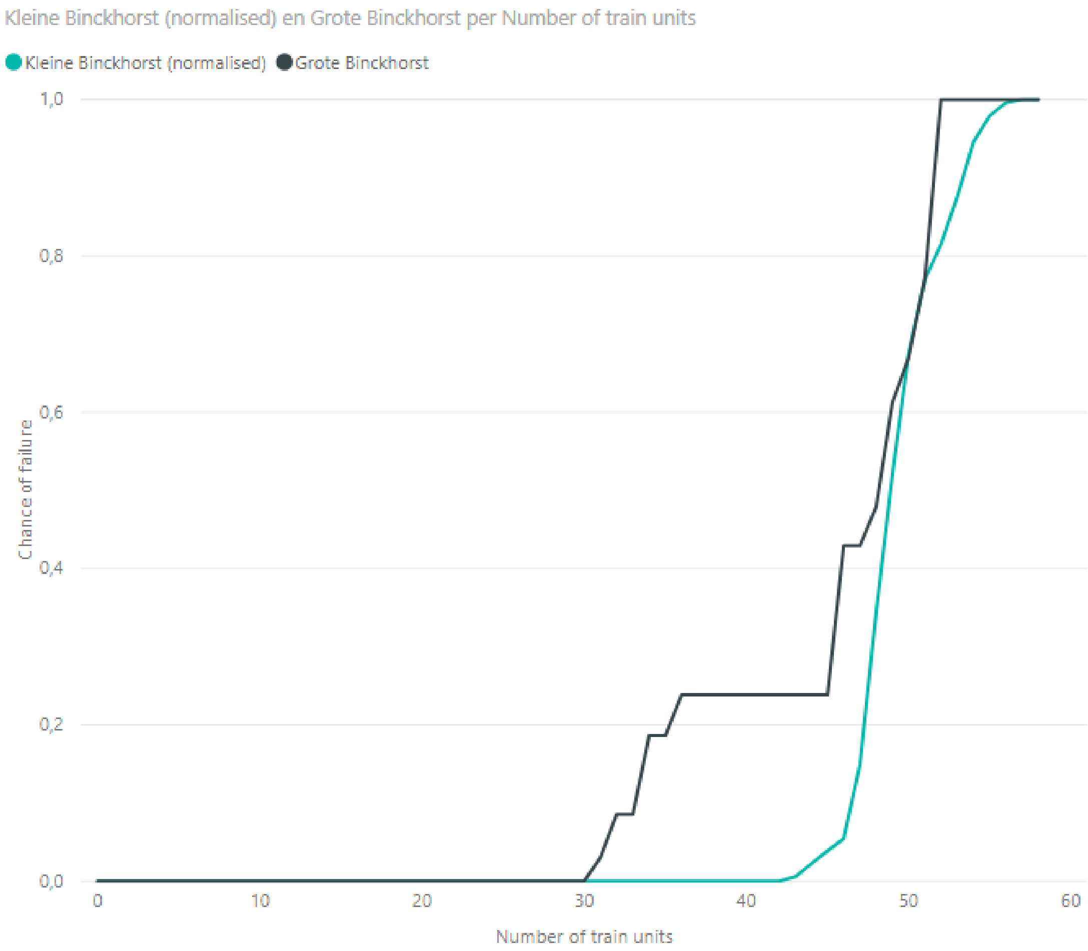

Figure 12: Survival functions of the Kleine Binckhorst (normalised) and the Grote Binckhorst

In the figure can be seen that the capacity at the Grote Binckhorst is about 15 train units smaller when the first instances start to fail. At the Grote Binckhorst the chance of failure is 0 for 30 train units and at the normalised value for the capacity at the Kleine Binckhorst is 45 train units. Those values come closer together while the chance of failure increases. When the chance that no solution will be found is 1, the Grote Binckhorst has a capacity value of 52 and the normalised Kleine Binckhorst has a capacity value of 58. The normalised Kleine Binckhorst scores overall a higher capacity value. This implies that the carousel layout works better than the shuffleboard layout which can be explained by the fact that in a carousel layout the train units can enter tracks from both sides, also mentioned in chapter 3. In a carousel topology there are many more shunting possibilities. It is not obligated to stick to the first-in-first-out rule anymore.

The median is the central measure of a sample, which strokes with the capacity value where the chance of failure is 0,5. This means that the median value of the sample can be identified in this way. For the Grote Binckhorst this value is estimated at 46 train units and for the normalised Kleine Binckhorst this value is estimated at 45 train units. These values are actually pretty close together compared the general differences between those two survival functions.

Another notable fact that strikes when looking at those survival functions is that the survival function of the Grote Binckhorst is less certain about the capacity. At the Grote Binckhorst the capacity values, which have a chance of failure between 0,1 and 0,9, is somewhere between the 33 and the 52. This is a much smaller interval for the Kleine Binckhorst, for which this is between the 49 and the 55. This means that the interval of potential values for the capacity is much larger for the Grote Binckhorst than for the Kleine Binckhorst. This implies that the standard deviation of the capacity at the Grote Binckhorst will be higher than the one at the Kleine Binckhorst.

In addition, the survival function of the normalised Kleine Binckhorst is much cleaner than the survival function of the Grote Binckhorst.

4.5.4. Quantitative analysis

To execute the quantitative analysis, the statistical tests will be executed. The certainty on the capacity will also be investigated by studying the boxplots.

4.5.4.1. Executing the Kolmogorov-Smirnov test

At first, it is important to test whether the two topology scenarios differ significantly from each other. This will be tested by using the Matlab scripts given in appendix A to get the values of the survival functions and appendix F to test whether the Kleine Binckhorst and the Grote Binckhorst differ significantly in capacity. The null hypothesis that will be tested with a significance level of α = 0,01 is stated below.

$$\text{H}_0\text{: The two topologies, the Kleine Binckhorst and the Grote Binckhorst, have the same capacity.}$$

By executing the test, it appeared that the null hypothesis needs to be rejected. That means that the alternative hypothesis, stated below, will be accepted.

$$\text{H}_1\text{: The two topologies, the Kleine Binckhorst and the Grote Binckhorst, do have significantly different capacities.}$$

Since significantly differences exist between the two scenarios, investigation of the nature of these differences might be useful.

4.5.4.2. Testing significant differences in median and variance

The sources can influence the capacity by mean value, for which in this research the median m is used as an indication, but it can also influence the certainty of the capacity. The certainty of the capacity is high if the standard deviation σ is low, and thus if the variance σ^2, is low. The right estimate of the capacity becomes less certain when the standard deviation σ, and thus the variance σ^2, are higher.

In the table below, the values for the median and the variance of the two scenarios are given. Those values are calculated based on the sample generated by the random generator in appendix B.

Table 10: Results of the median and variance at the Grote Binckhorst and the Kleine Binckhorst

	Grote Binckhorst	**Kleine Binckhorst (normalised)**
Median (m)	48	48
Variance (σ^2)	50,29	6,90

The results in Table 10 show that the medians of the samples are equal but the variance of the samples differ a lot. To check that the medians of the samples do not differ significantly a Wilcoxon signed rank test will be executed with a significance level of α = 0,01. The null hypothesis is stated below.

$$H_0: m_{\text{Kleine Binckhorst}} = m_{\text{Grote Binckhorst}}$$

The test is performed using the Matlab script in appendix F. It appeared, as expected, that the null hypothesis cannot be rejected. This means that the medians can be considered equal.

To test whether the variances of the sample differ significantly, a χ^2-test will be performed with a level of significance of α = 0,01. The null hypothesis is stated below.

$$H_0: {\sigma_{\text{Kleine Binckhorst}}}^2 = {\sigma_{\text{Grote Binckhorst}}}^2$$

The test is performed using the Matlab script in appendix F. It appeared, as expected, that the null hypothesis has to be rejected. This means that the alternative hypothesis has to be accepted. The alternative hypothesis is stated below.

$$H_1: {\sigma_{\text{Kleine Binckhorst}}}^2 \neq {\sigma_{\text{Grote Binckhorst}}}^2$$

This means the variances σ^2 of the two scenarios differ significantly.

4.5.4.3. Visualising the certainty on capacity

Since the variances differ significantly, it is interesting to investigate the degree of uncertainty in capacity. This can be visualised by making a boxplot from the samples of the scenarios. The boxplots of the topology, will in contrast to the other sources not be fed back to the database described in chapter 2, which exists of a set of instances based on random characteristic settings. This will not be done for the topology since the values in the scenario of the Kleine Binckhorst are multiplied with the normalisation factor to make comparison with the Grote Binckhorst possible. Due to this, the sample of the Kleine Binckhorst consists of higher numbers of train units than the values in the database since the values for the Kleine Binckhorst are not normalised in the database. The boxplots of the Grote Binckhorst and the normalised Kleine Binckhorst are shown in Figure 13.

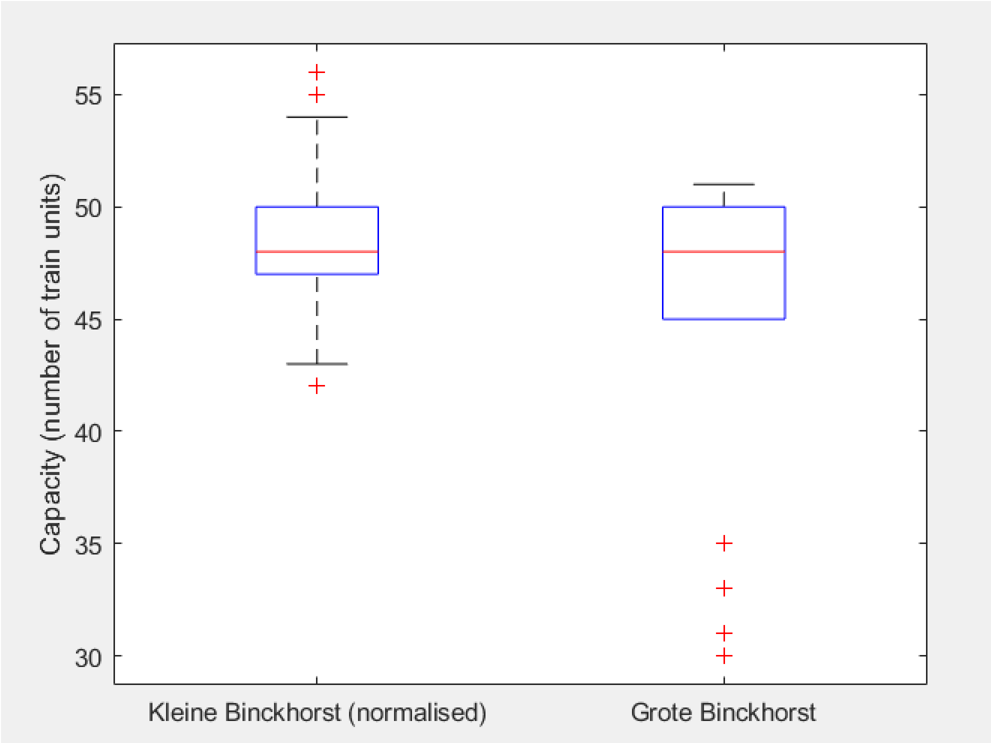

Figure 13: Boxplots of the capacity values of samples of the Kleine Binckhorst (normalised) and the Grote Binckhorst

In this figure is visualised that the interval at which values occur is much larger for the Grote Binckhorst than for the Kleine Binckhorst. This means that a carousel layout causes more certainty about the capacity value. The medians, represented by the red dash, are equal for the samples.

4.5.5. Interim conclusion – location layout

The conclusion that can be drawn from the research on the layout, based on a scenario at which the normalised Kleine Binckhorst is used as a topology and a scenario at which the Grote Binckhorst is used as a scenario, is that the median value does not differ for a carousel layout or a shuffleboard layout. This is interesting, since it should be logically that the carousel layout ends a higher capacity value.

In this problem, it could be that the median has not been a good indication for the mean, since the mean values do differ for both samples in contrast to the medians. In Table 11, the mean value for the sample of the normalised Kleine Binckhorst and the Grote Binckhorst are added to the values given in Table 10.

Table 11: Results of the median and variance at the Grote Binckhorst and the Kleine Binckhorst (normalised) expanded with the mean value

	Grote Binckhorst	**Kleine Binckhorst (normalised)**
Median (m)	48	48
Variance (σ^2)	50,29	6,90
Mean (μ)	44,68	48,80

The differences in the means imply that the location layout is a source of variation in capacity mean. Besides this expectation has revealed that the topology of the service location is a source of variation in capacity variance and certainty on capacity.

4.6. Influence of the topology on the capacity

In paragraph 4.5 the topology is investigated based on two scenarios: the Kleine Binckhorst and the Grote Binckhorst, where coupling and decoupling was not allowed and the standard setting for the arrival process, the 'normal capacity' was used. In this paragraphs appeared that the location layouts cause differences in capacity with standard settings.

In paragraph 4.3 the arrival process and in paragraph 4.4 allowing coupling and decoupling was investigated. In the research of those two sources, the results were based on scenarios in which instances were based on both the Kleine Binckhorst and the Grote Binckhorst. However, these results can differ, or at least be specified, per location layout. The specification of the sources per topology will be done in this paragraph.

4.6.1. Layout vs arrival process

The arrival process which is investigated in in paragraph 4.3 will be specified for the location topologies in this paragraph.

4.6.1.1. Scenario description

The scenarios which are used to investigate the arrival process will be used again in this specification. However, from now on the topologies also differ for the scenarios. This means that there are four scenarios per location layout. The total eight scenarios are showed in the table below.

Table 12: Composition of the scenarios set to investigate the arrival process per location

Scenario's		**Assumptions**
Kleine Binckhorst (normalised)	**Train units arrive spread over the last 25% of the arrival period at the Kleine Binckhorst**	Coupling and decoupling is not allowed and no tasks need to be executed.
	Train units arrive spread over the last 50% of the arrival period at the Kleine Binckhorst	Coupling and decoupling is not allowed and no tasks need to be executed.
	Train units arrive spread over the last 75% of the arrival period at the Kleine Binckhorst	Coupling and decoupling is not allowed and no tasks need to be executed.
	Train units arrive spread over 100% of the arrival period at the Kleine Binckhorst	Coupling and decoupling is not allowed and no tasks need to be executed.
Grote Binckhorst	**Train units arrive spread over the last 25% of the arrival period at the Grote Binckhorst**	Coupling and decoupling is not allowed and no tasks need to be executed.
	Train units arrive spread over the last 50% of the arrival period at the Grote Binckhorst	Coupling and decoupling is not allowed and no tasks need to be executed.

	Train units arrive spread over the last 75% of the arrival period at the Grote Binckhorst	Coupling and decoupling is not allowed and no tasks need to be executed.
	Train units arrive spread over 100% of the arrival period at the Grote Binckhorst	Coupling and decoupling is not allowed and no tasks need to be executed.

4.6.1.2. Qualitative analyses

From all scenarios survival functions are made. Those survival functions are displayed Figure 15 for the Kleine Binckhorst and Figure 14 for the Grote Binckhorst.

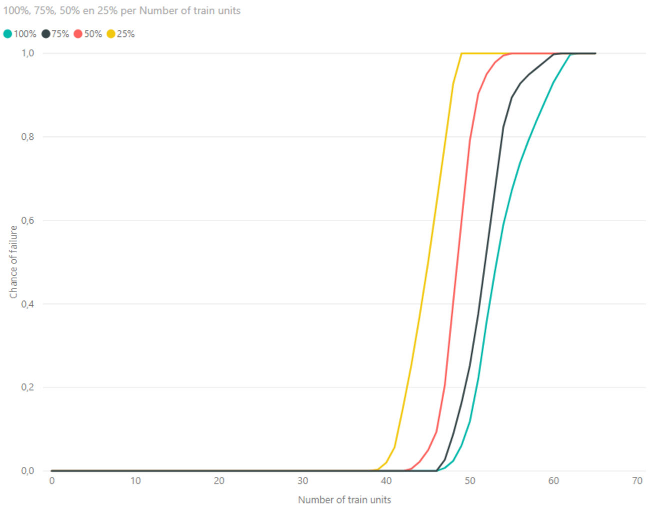

Figure 15: Arrival processes at the Kleine Binckhorst (normalised)

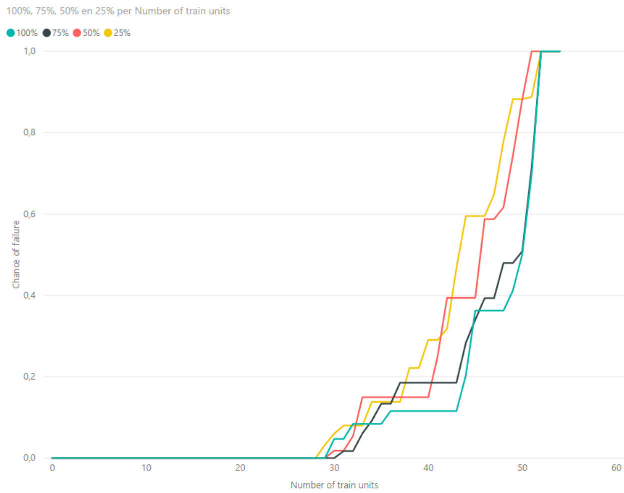

Figure 14: Arrival processes at the Grote Binckhorst

In these figures can be seen that the survival functions of the Kleine Binckhorst are much more smooth than the survival functions of the Grote Binckhorst. This was already encountered in paragraph 4.5.

In paragraph 4.3.2, the survival functions of the arrival processes were shown without specifying the location layout. Compared to those survival functions it can be concluded that the steep first parts of those survival functions is caused by values of the Kleine Binckhorst and the uncertainty of the second part of the survival functions is probably caused by the measures of the Grote Binckhorst.

In both Figure 16 and Figure 17 can be seen that the survival functions are increasingly higher while increasing the percentage of the arrival period of the train units. This implies that the values of the capacity could be significantly higher. In this research, this will be tested with the median m. The median is the central measure and an estimation of the central measure is the value at a chance of failure of 0,5. For the Kleine Binckhorst the medians can be estimated from the figure for 25%, 50%, 75% and 100% of the time. Those values for the medians are about 43, 46, 50 and 53 respectively. For the Grote Binckhorst, an estimate can also be made. For the survival functions in which the trains arrive in 25%, 50%, 75% and 100% of the arrival period, the median values can be estimated at 43, 45, 48 and 50 respectively. The medians for the same arrival process at the other location however differ not clearly. This would stroke with the experiences in discussed in paragraph 4.5.4.2.

The survival functions of the Kleine Binckhorst are certain about the value of the capacity. The values for the capacity are in intervals of about 10 train units. On the other hand, the survival functions of the Grote Binckhorst are very uncertain about the capacity value. Those values are lying in intervals of about 20 train units. This is a large difference, which will imply that the variance of the Grote Binckhorst is much larger than the variance of the Kleine Binckhorst.

4.6.1.3. Quantitative analyses

To execute the quantitative analysis for the specification of the arrival process per topology, statistical tests will be performed. The certainty on the capacity will also be compared for the two locations.

Executing the Kolmogorov-Smirnov test

At first, it is important to test whether the survival functions of two different scenarios differ significantly from each other. In case no significantly differences between two samples exist, there is no use in doing further research.

In this paragraph, differences between the locations will be investigated. Therefore two types of scenarios will be compared to each other. The first type is comparing the scenarios with the same arrival processes but with different locations with each other. The second type is investigating the survival functions with the same topologies between themselves.

Therefore, Kolmogorov-Smirnov tests will be executed. This will be tested by using the Matlab scripts given in appendix A to get the values of the survival functions and appendix F to test whether the scenarios differ significantly in capacity. The null hypothesis that will be tested with a significance level of α = 0,01 is stated below.

$$\text{H}_0\text{: The two scenarios do have the same capacity.}$$

Table 13: Results KS-test arrival process per topology

	25%	**50%**	**75%**	**100%**
25%	1	KB: 1 GB: 1	KB: 1 GB: 1	KB: 1 GB: 1
50%		1	KB: 1 GB: 1	KB: 1 GB: 1
75%			1	KB: 1 GB: 1
100%				1

(blue)	The blue cells show the results of the tests of different arrival processes of the same topology.
(yellow)	The yellow cells show the results of tests of the same arrival processes with the different topologies.

The result of the test for the scenarios are shown in Table 13. By executing the test, it appeared that the null hypothesis needs to be rejected in all cases. That means that the alternative hypothesis, stated below, will be accepted.

$$\text{H}_1\text{: The two scenarios do have significantly different capacities.}$$

This means that all scenarios investigated differ significantly from each other. Since significantly differences exist between all scenarios, investigation of the nature of these differences might be useful.

Testing significant differences in median and variance

The sources can influence the capacity by mean value, for which in this research the median m is used as an indication, but it can also influence the certainty of the capacity. The certainty of the capacity is higher if the standard deviation σ is low, and thus if the variance σ^2, is low. The right estimate of the capacity becomes less certain when the standard deviation σ, and thus the variance σ^2, are higher.

In the table below, the values for the median and the variance of the eight scenarios are given. Those values are calculated based on the sample generated by the random generator in appendix B.

Table 14: Results of the median and variance for the specified scenarios

Scenario's		Median (m)	Variance (σ^2)
Kleine Binckhorst (normalised)	**Train units arrive spread over the last 25% of the arrival period.**	45	4,19
	Train units arrive spread over the last 50% of the arrival period.	48	5,59
	Train units arrive spread over the last 75% of the arrival period.	51	7,70
	Train units arrive spread over 100% of the arrival period.	53	12,65
Grote Binckhorst	**Train units arrive spread over the last 25% of the arrival period.**	43	38,40
	Train units arrive spread over the last 50% of the arrival period.	45	38,56
	Train units arrive spread over the last 75% of the arrival period.	49	39,36
	Train units arrive spread over 100% of the arrival period.	49	38,50

The results in Table 14 show that the medians of the samples are increasing when increasing the percentage of the arrival period in which the train units arrive. This holds for both topologies. In the analysis of the arrival process was executed, but it was not specified for the location layouts. In this analyses also appeared that the values were increasing while increasing the percentage of the arrival period in which the train units arrive. In these results, larger differences can be seen, but this might have to do with the fact that the normalisation measure of the Kleine Binckhorst was not yet applied to the data.

In Table 14 can also be seen that the variance of increases when increasing the percentage of the arrival period in which the train units arrive at the Kleine Binckhorst. However, at the Grote Binckhorst the values for the variance are about the same.

To check whether the difference in medians of the samples is significant or not, a Wilcoxon signed rank test will be executed with a significance level of α = 0,01. To check whether the differences in variance of the samples are significant or not, χ^2-tests are executed with a significance level of α = 0,01. The tests are performed using the Matlab scripts in appendix F.

It appeared that the null hypothesis of the Wilcoxon rank test is rejected in almost all cases. Only differences in median between the scenario of 75% and 100% at the Kleine Binckhorst are not significant, which is obvious since the medians of those scenarios turned out to be equal. The results are clearly shown in Table 15, together with the results of the χ^2-test. The results of the χ^2-test stroke with the expectation. Significantly differences exist at the Kleine Binckhorst and don't exists at the Grote Binckhorst.

Table 15: Results Wilcoxon signed rank test and χ^2-test arrival process per topology

	25%	50%	75%	100%
25%	m: 1 σ^2 : 1	KB: m: 1 σ^2 : 1 GB: m: 1 σ^2 : 0	KB: m: 1 σ^2 : 1 GB: m: 1 σ^2 : 0	KB: m: 1 σ^2 : 1 GB: m: 1 σ^2 : 0
50%		m: 1 σ^2 : 1	KB: m: 1 σ^2 : 1 GB: m: 1 σ^2 : 0	KB: m: 1 σ^2 : 1 GB: m: 1 σ^2 : 0
75%			m: 1 σ^2 : 1	KB: m: 1 σ^2 : 1 GB: m: 0 σ^2 : 0
100%				m: 1 σ^2 : 1

(blue)	The blue cells show the results of the tests of different arrival processes of the same topology.
(yellow)	The yellow cells show the results of tests of the same arrival processes with the different topologies.

All differences in median, except for one at the Grote Binckhorst, are significant. Therefore, it can be concluded that the values for the capacity increase by increasing the percentage of the arrival period in which the train units arrive. This means that the arrival process causes variation in median at both locations.

It can also be concluded that the arrival process causes variation in variance at the Kleine Binckhorst, but it does not cause variation at the Grote Binckhorst, since the variances of the scenarios at the Grote Binckhorst do not show significant differences, but at the Kleine Binckhorst they do.

4.6.2. Layout vs coupling and decoupling

The possibility of allowing coupling and decoupling which is investigated in in paragraph 4.4 will be specified for the location topologies in this paragraph.

4.6.2.1. Scenario description

The scenarios which were used to investigate the allowing coupling and decoupling will be used again in this specification. However, from now on the topologies also differ for the scenarios. This means that there are two scenarios per location layout. The total four scenarios are showed in the table below.

Table 16: Composition of the scenarios set to investigate allowing coupling and decoupling per location

Scenario's		Assumptions
Kleine Binckhorst	**Coupling and decoupling allowed at the Kleine Binckhorst**	Coupling and decoupling is not allowed and no tasks need to be executed.
	Coupling and decoupling not allowed at the Kleine Binckhorst	Coupling and decoupling is not allowed and no tasks need to be executed.

Grote Binckhorst	**Coupling and decoupling allowed at the Grote Binckhorst**	Coupling and decoupling is not allowed and no tasks need to be executed.
	Coupling and decoupling not allowed at the Grote Binckhorst	Coupling and decoupling is not allowed and no tasks need to be executed.

4.6.2.2. Qualitative analyses

From all scenarios described in Table 16, survival functions are made. Those survival functions are displayed Figure 16 for the Kleine Binckhorst and Figure 17 for the Grote Binckhorst.

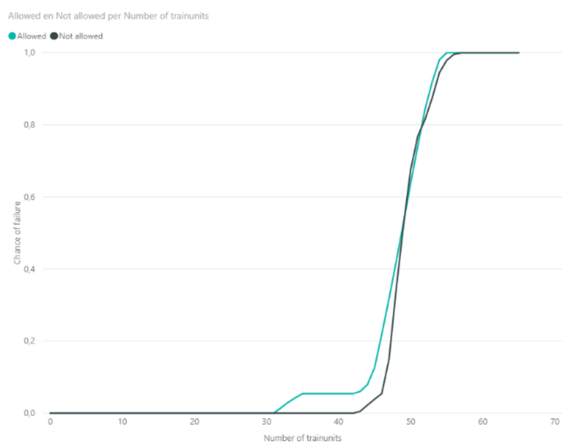

Figure 16: Coupling and decoupling at the Kleine Binckhorst

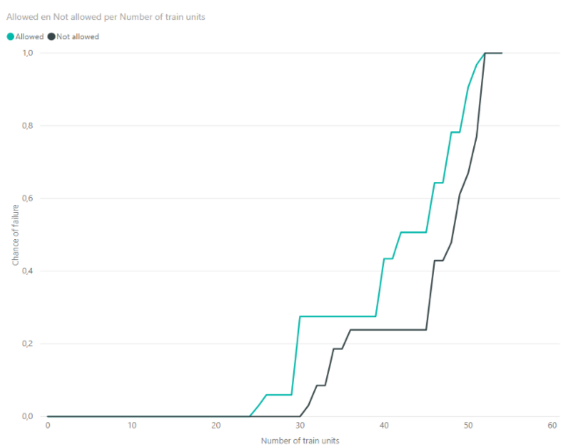

Figure 17: Coupling and decoupling at the Grote Binckhorst

In these figures can be seen that the survival functions of the Kleine Binckhorst are much more smooth than the survival functions of the Grote Binckhorst. This was already encountered in paragraph 4.5.

In paragraph 4.4.2, the survival functions of allowing and not allowing coupling and decoupling were shown without specifying the location layout. Compared to those survival functions it can be concluded that the steep first parts of those survival functions is caused by values of the Kleine Binckhorst and the uncertainty of the second part of the survival functions is probably caused by the measures of the Grote Binckhorst.

In Figure 16 can be seen that allowing coupling and decoupling at the Kleine Binckhorst does not have a lot influence overall. Only at the start of the survival function clear differences exist. The medians, which is the central measure, can be estimated at 50 for both allowing and not allowing coupling and decoupling at the Kleine Binckhorst. The variance could be different for the two scenarios at the Kleine Binckhorst since the first failure while not allowing coupling and decoupling is at 42 train units and this value is 32 while allowing coupling and decoupling. This difference disappears very soon, but it causes that the interval in which values appear is much smaller when not allowing coupling and decoupling.

In Figure 17 more differences between the values can be seen. The whole survival function scores higher capacity while not allowing coupling and decoupling. This can be indicated by the median values of both survival functions. At the Grote Binckhorst, the survival function of allowing coupling and decoupling has a median value of about 42 train units, while not allowing coupling and decoupling had a median value of about 47 train units. The interval in which the values appear is also smaller while not allowing coupling and decoupling at the Grote Binckhorst. This means that there is more uncertainty when allowing coupling and decoupling.

4.6.2.3. Quantitative analysis

To execute the quantitative analysis for the specification of the arrival process per topology, statistical tests will be performed. The certainty on the capacity will also be compared for the two locations.

Executing the Kolmogorov-Smirnov test

At first, it is important to test whether the survival functions of two different scenarios differ significantly from each other. In case no significantly differences between two samples exist, there is no use in doing further research.

In this paragraph, differences between the locations will be investigated. Therefore two types of scenarios will be compared to each other. The first type is comparing the scenarios with the same status for coupling and decoupling but with different locations with each other. The second type is investigating the survival functions with the same topologies between themselves.

Therefore, Kolmogorov-Smirnov tests will be executed. This will be tested by using the Matlab scripts given in appendix A to get the values of the survival functions and appendix F to test whether the scenarios differ significantly in capacity. The null hypothesis that will be tested with a significance level of α = 0,01 is stated below.

$$\mathrm{H_0}: \text{The two scenarios do have the same capacity.}$$

Table 17: Results KS-test allowing coupling and decoupling per topology

	Coupling and decoupling allowed	**Coupling and decoupling not allowed**
Coupling and decoupling allowed	1	GB: 1 KB: 0
Coupling and decoupling not allowed		1

(blue)	The blue cells show the results of the tests of different statuses for allowing coupling and decoupling of the same topology.
(yellow)	The yellow cells show the results of the tests of the same statuses for allowing coupling and decoupling at different topologies.

The result of the test for the scenarios are shown in Table 13. By executing the test, it appeared that the null hypothesis has to be accepted for the two scenarios at the Kleine Binckhorst. This means that the survival functions can be considered equal and that there allowing coupling and decoupling has no influence at the capacity at the Kleine Binckhorst.

At the Grote Binckhorst, significantly differences do exist. This means that the influence that is encountered in paragraph 4.4 can be assigned to the Grote Binckhorst. In paragraph 4.4, the influence of coupling and decoupling is investigated without taking the topologies into account. In that analyses, only the Grote Binckhorst and the Kleine Binckhorst were part of the data. Since the

Kleine Binckhorst turned out to have no influence, the influence can be assigned to the Grote Binckhorst.

This means that coupling and decoupling at the Grote Binckhorst, which has a shuffleboard layout, causes variation in variance and thus certainty. The medians did not differ, which means that coupling and decoupling is not a source of variation in median at the Grote Binckhorst.

At the Kleine Binckhorst, no significant influence was caused by allowing or not allowing coupling and decoupling. This means that for the Kleine Binckhorst, which has a carousel layout, allowing or not allowing coupling and decoupling is not a source of variation.

Chapter 5
Conclusion and discussion

In this last chapter of the research the conclusion will be stated. Also a discussion of the results will be done. Finally, some recommendations will be made.

5.1. Conclusion

In this research the following research question is investigated:

How do different sources influence the variation in capacity of the NS service location the Binckhorst?

The sources that were investigated are the arrival process, allowing coupling and decoupling and the topology of the service location. Also the influence of the service location on the arrival process and allowing coupling and decoupling is researched.' This is researched by making survival functions of different scenarios. These survival functions are tested on significant differences by executing Kolmogorov Smirnov tests. The cause of the impact is specified testing for significant differences in median and in variance.

At first, the arrival process was investigated. The conclusion that could be drawn from the investigation of the arrival process, was that the median values do not differ significantly for different percentages in which the train units arrive, only for two extreme scenarios. This means that this source causes only a little bit variation in capacity median. Increasing the time period in which the train units arrive, the capacity median increases. The variance was also tested and it turned out that significant differences exist between the variances. The differences in sample variances imply that increasing the time period in which the train units arrive causes variation in capacity variance.

The second source that was researched is allowing coupling and decoupling. The conclusion that could be drawn from the investigation of allowing coupling and decoupling, based on the scenarios that are investigated in this research, is that the median values do not differ for allowing or not allowing coupling and decoupling. It was expected that this should differ, since it becomes harder to find a solution when trains depart in different configurations than they arrived in. This was not the case. An explanation for this can be that calculation time, which was set as 120 seconds per instance, is amply sufficient in both scenarios. The differences in sample variances imply that allowing coupling and decoupling is a source of variation in capacity variance and certainty on capacity.

However, since the variances are very high, the certainty about the capacity is very low for both the arrival process and allowing coupling and decoupling. It is desirable to decrease the interval in which values occur. An explanation for this low certainty on capacity is that the Kleine Binckhorst and the Grote Binckhorst differ in size. This will be investigated in the last source.

The last source of investigation is the topology of the service location. This was investigated by comparing the Kleine Binckhorst, which has a carousel layout, and the Grote Binckhorst, which has a shuffleboard layout. Those two topologies differ in size. Therefore a normalisation measure is determined, which resulted in smaller differences between the two locations. The conclusion that could be drawn from the research on the layout, is that the median value does not differ for a carousel layout or a shuffleboard layout. This is interesting, since it should be logically that the carousel layout ends a higher capacity value. However, the variance of both locations differs a lot.

This means that the location layout is a source of variation in variance. The variance of the Kleine Binckhorst is much lower than the variance of the Grote Binckhorst. Therefore, the certainty on capacity is much higher for the Kleine Binckhorst compared to the Grote Binckhorst. This can be an explanation of the high values for the variance in the research on the arrival process and on allowing coupling and decoupling. Therefore those two sources are also investigated when specified for the Kleine Binckhorst and the Grote Binckhorst.

Investigation of the arrival process specified per location layout gave the information that the medians of the scenarios of the Kleine Binckhorst differed significantly, as well as the medians of the scenarios of the Grote Binckhorst. This was not the case when the arrival process was not specified per location. The arrival process causes also variation in variance for the Kleine Binckhorst. For the Grote Binckhorst however there was no significant influence.

Investigation of allowing coupling and decoupling per location layout gave the information that allowing or not allowing coupling and decoupling has no influence on the Kleine Binckhorst. On the other hand, at the Grote Binckhorst it does have influence. This means that the results of the investigation allowing coupling and decoupling without specifying the topologies are applicable to the Grote Binckhorst. The medians did not differ, which means that coupling and decoupling is not a source of variation in median at the Grote Binckhorst, but the variances did differ significantly. This means that coupling and decoupling turned out to be a source of variation in variance at a location with a shuffleboard layout.

5.2. Discussion

In this research some assumptions has to be made. These assumptions can cause uncertainty or errors. These uncertainties are discussed in this paragraph.

An important assumption was that the survival functions which were based on the data generated gave a good indication of the capacity. The number of train units per work package was varied from 1 to 50 for a scenario. The data used for one survival function of one scenario was 30 instances per number of train units. It is assumed that this was sufficient.

Another assumption is that the model used to gain the data, the behandelcalculator, is completely in accordance with the reality. In the behandelcalculator the calculation time can be set. This is the time that the algorithm is searching for a solution for one instance. The longer this time, the higher the possibility that a solution will be found. It is assumed that the calculation time, which was set on 120 seconds, was sufficient for getting a good insight in the capacity.

To construct the survival functions, right and left censored data is taken into account. When no solution is found for an instance, this means that the work package had another limit which was for a smaller number of train units. This left censored data is taken into account by using the Turnbull algorithm. When a solution is found for an instance, this means that the work package had another limit which was for the same or a higher number of train units. This right censored data is taken into account by using the Product Limit method. These methods increase the certainty of the research since more information is obtained from the data.

In this research is worked with the Wilcoxon signed rank test to test significant differences in the median. The median is used since the distribution of the survival functions were not known. The statistical test for the median is more robust than for the mean, because the statistical test for testing means assumes a normal distribution. However, it is possible that the data was normally distributed. Especially the survival functions of the Kleine Binckhorst seemed to have the shape of a normal distribution. In that case, testing by means would have been more logical.

By investigating the location topologies, the Kleine Binckhorst and the Grote Binckhorst needed to be compared. However, these two locations differ in size. These differences in sizes are not taken into account when the layout was not specified. This has caused larger differences than necessary which might have turned out in more uncertainty. By comparing the two locations with each other a normalisation measure is determined based on the length of railway at the locations. In this method, some railway length where no train unit fits anymore (cutting loss) is calculated as useful track length, while it is not. This causes less certainty about the normalisation measure.

5.3. Recommendations

Based on the results, the following future steps are recommended.

In this research survival functions are constructed to get insight in the capacity. These survival functions are determined based on 30 measures per number of train units, which is varied from 1-50. This means that a survival function is based 1500 instances for which a solution is found or not. To increase the certainty, it is recommended to do more measures.

To improve the normalisation measure, which is used to compare the Kleine Binckhorst and the Grote Binckhorst with each other, only the useful track length needs to be used. In the current method, cutting loss is calculated as useful track length, while it is not. A normalisation measure based on the maximum amount of train units that can be placed at a location would be a better approach. In this way, the cutting loss will be taken into account.

In this research some choices had to be made about which sources were going to be investigated. This is done to limit the scope of the research. Therefore only three sources are investigated. However, it would be interesting to investigate more sources to get an even better picture. Sources that might be interesting are the influence of the departure process and the interaction of the arrival process with the departure process, the train types and the tasks that need to be carried out on the trains. The layout of the service locations can also be more specified. In this research only two types of layout are investigated, but the influence of adjusting one characteristic of the topology might also be interesting.

References

Apallius de Vos, J. I., & Van Dongen, L. A. (2015). Performance Centered Maintenance as a core policy in strategic maintenance control . *Elsevier*.

Beerthuizen, E. (2017). *Optimization train parking and shunting at NS service yards.* Delft.

Boxtel, R. v. (2018, juli 17). *Insite*. Retrieved from Roffel: NS onthult nieuwe toekomstvisie: https://nsdigitaal.sharepoint.com/Actueel/Paginas/Roffel-NS-onthult-nieuwe-toekomstvisie.aspx

Brilon, W., Geistefeldt, J., & Regler, M. (2005). Reliability of freeway traffic flow: A stochastic concept of capacity. *International Symposium of Transportation and Traffic Theory*, 125-144.

Busstra, M., & van Dongen, L. (2015). Creating value by integrating logistic trains services and maintenance activities. *Elsevier*.

Calvert, S., Taale, H., & Hoogendoorn, S. (2015). Quantification of motorway capacity variation: influence of day type specific variatio and capacity drop. *Department of Transport & Planning*.

Chakravarti IM, L. (2009). Handbook of methods of applied statistics. In J. W. Sons, *Handbook of emthods of applied statistics.* London.

Freling, R., Lenting, R., Kroon, L., & Huisman, D. (2005). Shunting of passenger train units in a railway station. *Transportation Science*.

Freling, R., Lentink, R., Kroon, L., & Huisman, D. (2002). *Shunting of Passenger Train Units in a Railway Station.* Rotterdam.

Haahr, J., Lusby, R., & Wagenaar, J. (2017). Optimization methods for the Train Unit Shunting Problem. *European Journal of Operational Research*.

Hoepel, F. (2017). *Decision support on capacity problems NS.* Eindhoven.

Huizingh, E. (2018). *Planning first-line services on NS service stations.* Enschede.

Jia A, W. (2010). Identification and Calibration of Site-Specific Stochastic Freeway Breakdown and Queue Discharge. *Transportation Research Record: Journal of the Transportation Research Board*, 148-155.

Klein, J. P., & Moeschberger, M. L. (2003). *Survival Analysis Techniques for Censored and Truncated Data.* New York: Springer.

Kucherov, S., Rogozov, Y., Lipko, J., & Elkin, D. (2018). The concept of the method for dynamic control of traffic flows on multi-lane roads based on configurable information systems. *Advances in Intelligent system computing.* Taganrog, Russian Federation.

Lentink, R. (2006). *Algorithmic Decision Support for Shunt Planning.* Rotterdam.

Mathew, T., & Krishna Rao, K. (2007). Capacity and Level of service. In T. V. Mathew, & K. V. Krishna Rao, *Introduction to Transport Engineering* (p. Chapter 35).

NedTrain. (2014). *Werkbeschrijving Servicebedrijven A- en B-controle.* Utrecht.

Non-parametric tests and regression. (2018, maart 15). *Empirical analysis for Transport and Planning*. Delft.

NS. (2018). *Herkennen treintypes.* Utrecht.

NS en ProRail: het spoor is bijna vol. (2018, augustus 20). Retrieved from Treinreiziger: https://www.treinreiziger.nl/ns-en-prorail-spoor-is-bijna/

NS Techniek. (2018). *Rekenmethode voor de opstel- en behandelcapaciteit van emplacementen of servicebedrijven.* Urecht.

Olba, X., Daamen, W., Vellinga, T., & Hoogendoorn, S. (2017). *Network Capacity Estimation of Vessel Traffic: An Approach for Port Planning.*

Özkan, E., Nas, S., & Nil, G. (2016). Capacity Analysis of Ro-Ro Terminals by Using Simulation Modeling Method. *The Asian Journal of Shipping and Logistics*, 139-147.

Rosales, L. (2015). *Analysing uncertainty and delays in aircragt heavy maintenance.* Manchester.

Ross SM. (2009). Introduction to probability and statistics for engineers and scientists. *Academic Press Elsevier*, London.

van Berkum, E., & Di Bucchianico, A. (2007). *Statistical Compendium.* Eindhoven.

van den Broek, R., Hoogeveen, H., & van den Akker, M. (2012). *How to Measure the Robustness of Shunting Plans.* Utrecht.

Wolfe. (2000). Biostat 675. In Wolfe, *Censoring Mechanisms.*

-

Appendices

A. Matlab script constructing a survival function

```
clear all
tic

%% Load data
[Data_previous] = xlsread('Analyse_B','AnalyseB'); %Column 1: Number_trainunits, 
Column 2: Number_instances, Column 3: Number_failures, Column 4: Number_successes, 
Column 5: S_estimate, Column 6: F_estimate



%% Prepare data
m_1 = 30; %number of instances per trainunitvalue
m_2 = 54; %number of trainunitvalues

Data = [1:m_2]';
Data(:,2) = m_1;

for i = 1 : m_2;
    k = 0;
    for j = 1 : m_1;
        if Data_previous(j,2) == 0
            k = k + 1
        end
    end
    
    Data(i,3) = k;
    Data_previous(1:m_1,:) = [];
    
end

Data(:,4) = Data(:,2) - Data(:,3);
Data(:,5) = Data(:,4)./Data(:,2);

for i = 2:m_2;
    if Data(i,5) > Data(i-1,5);
        Data(i,5) = Data(i-1,5);
    end
end


%% Start situation
Sk = [Data(:,1),Data(:,5)]; % Column 1: Number_trainunits, Column 2: S, Column 3: 
Di
Finished = 0;
Marge = 0.001;
count = 0;
n = length(Data(:,1));

CDF = 1 %1 if CDF, 0 if PDF
trainunits = Sk(:,1);
startsurvival = 1 - Sk(:,2);
Boxplotvalue = 100;

%% While loop

while Finished == 0;
    
    %% Pij waarden berekenen
    Pij = zeros(n-1);
    for i = 2:n;
        
        for j = 2:n;
```

```
        if Sk(i,2) == 1;
            Pij(j-1,i-1) = 0;
        elseif j <= i;
            Pij(j-1,i-1) = (Sk(j-1,2) - Sk(j,2))/(1-(Sk(i,2)));
        else
            Pij(j-1,i-1) = 0;
        end
    end
end

%% Di waarden berekenen

for i = 2 : n;
    Di = 0;
    for j = 2 : n;
        x = Pij(i-1,j-1) .* Data(j,3);
        Di = Di + x;
    end
    Sk(i,3) = Di;
end


%% Kaplan-Meier tabel maken
KM = zeros(n,6);

%Number of trainunits
KM(:,1) = Data(:,1);

%Number of events
KM(:,2) = Sk(:,3);

%Number of censored data (=successes at ti-1)
for i = 2:n;
    KM(i,3) = Data(i-1,4);
end

%Number of individuals
A_1 = [KM(:,2) + KM(:,3)];
for i = 1:n;
    A_2 = sum(A_1);
    A_1(i) = 0;
    KM(i,4) = A_2;
end

%Factor
for i = 1:n;
    if KM(i,4) == 0;
        KM(i,5) = 0;
    else
        KM(i,5) = 1-(KM(i,2)./KM(i,4));
    end
end


%% S(k plus 1) waarden berekenen


KM(:,6) = 1;
for i = 2:n;
    KM(i,6) = KM(i-1,6) * KM(i,5);
end


Som = 0;
for i = 1:length(KM(:,6));
    x = abs(KM(i,6)-Sk(i,2));
    Som = Som + x;
end
```

```
    %% End
    %Som = 0.001
    if Som <= Marge
        Finished = 1;
    else
        Sk(:,2) = KM(:,6);
        Sk(:,3) = 0;
        KM(:,5) = 1;
        KM(:,6) = 1;
    end

    count = count + 1;


end

if CDF == 1;
    figure(1), clear clf, hold on;
    %plot(trainunits,startsurvival)
    %plot(trainunits,1-KM(:,6)) %standaard
    %plot((trainunits.*2),1-KM(:,6)) %kleine binckhorst genormaliseerd
    xlabel('Number of train units')
    ylabel('Chance of failure')
    legend('Kleine Binckhorst (roughly normalised)', 'Grote Binckhorst')
    hold off;
elseif CDF == 0;
    PDF = zeros(n,3);
    PDF(:,1) = 1-KM(:,6); %failure function
    for i = 1:n-1;
        PDF(i+1,2) = PDF(i,1);
    end
    PDF(:,3) = PDF(:,1) - PDF(:,2);

    figure(1), clear clf, hold on;
    plot(trainunits,PDF(:,3));
    hold off

    %% Boxplot
    for i = 2:m_2;
        PDF(1,4) = PDF(1,3);
        PDF(i,4)= PDF(i-1,4) + PDF(i,3);
        PDF(i,5) = PDF(i,4) .* Boxplotvalue;
    end

    Cum = 0;
    for i = 1:Boxplotvalue;
        z = 0;
        for j = 1:m_2;
            if z == 0;
                if Cum < PDF(j,5);
                    X(i) = j
                    z = 1;
                end
            end
        end
        Cum = Cum + 1;
    end

    figure(2), clear clf, hold on
    boxplot(X)
    hold off

end

KM(:,7) = 1 - KM(:,6);
toc
```

B. Matlab script random generator

```
clear all, clc

%add zero as a first value
SF =
[0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0.0322934724650821;0.216020475748690;0
.445951721874444;0.717822227947671;1;1;1;1;1];
S = 100000;
v = 1;

    V = zeros(1,S);
    n = length(SF)-1;
    N = zeros(n+1,2);
    N(:,1) = [0:n];
    SF = 1-SF;
    N(:,2) = SF;

    for i = 1:S

        R_1 = rand; %random value between 0 and 1

        x = 0;
        for j = 1:n+1
            if x == 0
                if R_1 > N(j,2)
                    value = N(j,1)-1;
                    x = 1;
                end
            end
        end

        V(v) = value;
        v = v+1;


    end
```

C. Matlab script filter

```
clear all
tic

%% Load data
[DataSec] = xlsread('Data_Filteren_SF_AP','AP_1'); %AP
[DataInst] = xlsread('Data_Filteren_SF_AP_2','TU_1'); %TU

%% Gegevens instellen
n_start = 1; %lowest instance_id
n_eind = 20; %highest instance_id
x = 1; %1 = gradually, 2 = last 25%


%% Loop

z = 0;
count = 1;
C = [];


for i = n_start:n_eind;
    %for j = 1:60 %maximum number of train units in an instance
    j = 1;
    k = 2;
    K = [0];

    %vector K
    while DataSec(j,1) == i; %as long as instance_id is a certain number
        if DataSec(j,2) > 0
            if DataSec(j,3) == 1
                K(k) = DataSec(j,2);
                k = k + 1;
            elseif DataSec(j,3) == 2
                K(k) = DataSec(j,2);
                K(k+1) = DataSec(j,2);
                k = k + 2;
            elseif DataSec(j,3) == 3
                K(k) = DataSec(j,2);
                K(k+1) = DataSec(j,2);
                K(k+2) = DataSec(j,2);
                k = k + 3;
            end
        end

        j = j + 1;

    end


    % Reference survival distribution
    if x == 1 %gradually
        SFRef = [0:(1/(j-1)):1];
    elseif x == 2 %last 25% of the time
        SFRef = [0,0.75:(0.25/(j-1)):1];
    end


            % Complete Survival function
            K(k) = 1;
            K = sort(K);


            if i == DataSec(1,1);
            z = z + 1;
```

```
        %Delete rows
        DataSec(1:(DataInst(z,3)),:) = [];

        %KS test
        ks_value = kstest2(SFRef,K);
        DataInst(z,4) = ks_value;
        end

        count = count + 1;
        C(count) = count

end

%% Plots
figure(2), clear clf, hold on
plot(K)
hold off


%% End

toc
```

D. Critical values χ²-test

10.3 χ^2-distribution

Example: $P(\chi_3^2 \geq 6.25) = 0.1$, thus $\chi_{3;0.1}^2 = 6.25$.

	0.005	0.01	0.025	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.975	0.99	0.995
1	7.88	6.63	5.02	3.84	2.71	1.32	.455	.102	.016	.004	.000	.000	.000
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39	.575	.211	.103	.051	.020	.010
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37	1.21	.584	.352	.216	.115	.072
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36	1.92	1.06	.711	.484	.297	.207
5	16.7	15.1	12.8	11.1	9.24	6.63	4.35	2.67	1.61	1.15	.831	0.55	.412
6	18.5	16.8	14.4	12.6	10.6	7.84	5.35	3.45	2.20	1.64	1.24	0.87	.676
7	20.3	18.5	16.0	14.1	12.0	9.04	6.35	4.25	2.83	2.17	1.69	1.24	.989
8	22.0	20.1	17.5	15.5	13.4	10.2	7.34	5.07	3.49	2.73	2.18	1.65	1.34
9	23.6	21.7	19.0	16.9	14.7	11.4	8.34	5.90	4.17	3.33	2.70	2.09	1.73
10	25.2	23.2	20.5	18.3	16.0	12.5	9.34	6.74	4.87	3.94	3.25	2.56	2.16
11	26.8	24.7	21.9	19.7	17.3	13.7	10.3	7.58	5.58	4.57	3.82	3.05	2.60
12	28.3	26.2	23.3	21.0	18.5	14.8	11.3	8.44	6.30	5.23	4.40	3.57	3.07
13	29.8	27.7	24.7	22.4	19.8	16.0	12.3	9.30	7.04	5.89	5.01	4.11	3.57
14	31.3	29.1	26.1	23.7	21.1	17.1	13.3	10.2	7.79	6.57	5.63	4.66	4.07
15	32.8	30.6	27.5	25.0	22.3	18.2	14.3	11.0	8.55	7.26	6.26	5.23	4.60
16	34.3	32.0	28.8	26.3	23.5	19.4	15.3	11.9	9.31	7.96	6.91	5.81	5.14
17	35.7	33.4	30.2	27.6	24.8	20.5	16.3	12.8	10.1	8.67	7.56	6.41	5.70
18	37.2	34.8	31.5	28.9	26.0	21.6	17.3	13.7	10.9	9.39	8.23	7.01	6.26
19	38.6	36.2	32.9	30.1	27.2	22.7	18.3	14.6	11.7	10.1	8.91	7.63	6.84
20	40.0	37.6	34.2	31.4	28.4	23.8	19.3	15.5	12.4	10.9	9.59	8.26	7.43
21	41.4	38.9	35.5	32.7	29.6	24.9	20.3	16.3	13.2	11.6	10.3	8.90	8.03
22	42.8	40.3	36.8	33.9	30.8	26.0	21.3	17.2	14.0	12.3	11.0	9.54	8.64
23	44.2	41.6	38.1	35.2	32.0	27.1	22.3	18.1	14.8	13.1	11.7	10.2	9.26
24	45.6	43.0	39.4	36.4	33.2	28.2	23.3	19.0	15.7	13.8	12.4	10.9	9.89
25	46.9	44.3	40.6	37.7	34.4	29.3	24.3	19.9	16.5	14.6	13.1	11.5	10.5
26	48.3	45.6	41.9	38.9	35.6	30.4	25.3	20.8	17.3	15.4	13.8	12.2	11.2
27	49.6	47.0	43.2	40.1	36.7	31.5	26.3	21.7	18.1	16.2	14.6	12.9	11.8
28	51.0	48.3	44.5	41.3	37.9	32.6	27.3	22.7	18.9	16.9	15.3	13.6	12.5
29	52.3	49.6	45.7	42.6	39.1	33.7	28.3	23.6	19.8	17.7	16.0	14.3	13.1
30	53.7	50.9	47.0	43.8	40.3	34.8	29.3	24.5	20.6	18.5	16.8	15.0	13.8
31	55.0	52.2	48.2	45.0	41.4	35.9	30.3	25.4	21.4	19.3	17.5	15.7	14.5
32	56.3	53.5	49.5	46.2	42.6	37.0	31.3	26.3	22.3	20.1	18.3	16.4	15.1
33	57.6	54.8	50.7	47.4	43.7	38.1	32.3	27.2	23.1	20.9	19.0	17.1	15.8
34	59.0	56.1	52.0	48.6	44.9	39.1	33.3	28.1	24.0	21.7	19.8	17.8	16.5
35	60.3	57.3	53.2	49.8	46.1	40.2	34.3	29.1	24.8	22.5	20.6	18.5	17.2
36	61.6	58.6	54.4	51.0	47.2	41.3	35.3	30.0	25.6	23.3	21.3	19.2	17.9
37	62.9	59.9	55.7	52.2	48.4	42.4	36.3	30.9	26.5	24.1	22.1	20.0	18.6
38	64.2	61.2	56.9	53.4	49.5	43.5	37.3	31.8	27.3	24.9	22.9	20.7	19.3
39	65.5	62.4	58.1	54.6	50.7	44.5	38.3	32.7	28.2	25.7	23.7	21.4	20.0
40	66.8	63.7	59.3	55.8	51.8	45.6	39.3	33.7	29.1	26.5	24.4	22.2	20.7
u	2.58	2.33	1.96	1.64	1.28	0.67	0.00	−0.67	−1.28	−1.64	−1.96	−2.33	−2.58

For values not in the table one can use the approximation of Wilson and Hilferty (see [6, p. 176]) to obtain the critical value: $\chi_{\nu,\alpha}^2 = \nu \left(u\sqrt{\frac{2}{9\nu}} + 1 - \frac{2}{9\nu} \right)^3$, where u is given in the bottom line of the table.

Figure 18: Critical values χ² test *(van Berkum & Di Bucchianico, 2007)*

E. Matlab script adding two analyses

```
clear all
tic

%% Turnbulls algorithm 3 wordt gebruikt om 2 analyses bij elkaar op te tellen

%% Load data
[Data_previous1] = xlsread('Analyse_B','AnalyseB'); %Column 1: Number_trainunits,
Column 2: Number_instances, Column 3: Number_failures, Column 4: Number_successes,
Column 5: S_estimate, Column 6: F_estimate
[Data_previous2] = xlsread('Analyse_K','AnalyseK');


%% Prepare data
m_1 = 30; %number of instances per trainunitvalue
m_2 = 54; %number of trainunitvalues Data_previous1
m_3 = 30; %number of trainunitvalues Data_previous2

Data1 = [1:m_2]';
Data1(:,2) = m_1;

for i = 1 : m_2;
    k = 0;
    for j = 1 : m_1;
        if Data_previous1(j,2) == 0
            k = k + 1
        end
    end

    Data1(i,3) = k;
    Data_previous1(1:m_1,:) = [];

end

Data1(:,4) = Data1(:,2) - Data1(:,3);

Data2 = [1:m_2]';
Data2(:,2) = m_1;

for i = 1 : m_3;
    k = 0;
    for j = 1 : m_3;
        if Data_previous2(j,2) == 0
            k = k + 1
        end
    end

    Data2(i,3) = k;
    Data_previous2(1:m_1,:) = [];

end

Data2(:,4) = Data2(:,2) - Data2(:,3);


if m_2 < m_3
    for i = (m_2+1):m_3
        Data1(i,1) = i;
        Data1(i,2) = m_1;
        Data1(i,3) = m_1;
    end
    m = m_3;

elseif m_3 < m_2
    for i = (m_3+1):m_2
        Data2(i,1) = i;
```

```
        Data2(i,2) = m_1;
        Data2(i,3) = m_1;
    end
    m = m_2;
end

Data = [1:m]';
Data(:,2) = Data1(:,2) + Data2(:,2);
Data(:,3) = Data1(:,3) + Data2(:,3);
Data(:,4) = Data(:,2) - Data(:,3);
Data(:,5) = Data(:,4)./Data(:,2);

for i = 2:m;
    if Data(i,5) > Data(i-1,5);
        Data(i,5) = Data(i-1,5);
    end
end


%% Start situation
Sk = [Data(:,1),Data(:,5)]; % Column 1: Number_trainunits, Column 2: S, Column 3:
Di
Finished = 0;
Marge = 0.001;
count = 0;
n = length(Data(:,1));

CDF = 1 %1 if CDF, 0 if PDF
trainunits = Sk(:,1);
startsurvival = 1 - Sk(:,2);
Boxplotvalue = 100;

%% While loop

while Finished == 0;

    %% Pij waarden berekenen
    Pij = zeros(n-1);
    for i = 2:n;

        for j = 2:n;
            if Sk(i,2) == 1;
                Pij(j-1,i-1) = 0;
            elseif j <= i;
                Pij(j-1,i-1) = (Sk(j-1,2) - Sk(j,2))/(1-(Sk(i,2)));
            else
                Pij(j-1,i-1) = 0;
            end
        end
    end

    %% Di waarden berekenen

    for i = 2 : n;
        Di = 0;
        for j = 2 : n;
            x = Pij(i-1,j-1) .* Data(j,3);
            Di = Di + x;
        end
        Sk(i,3) = Di;
    end


    %% Kaplan-Meier tabel maken
    KM = zeros(n,6);

    %Number of trainunits
    KM(:,1) = Data(:,1);
```

```
    %Number of events
    KM(:,2) = Sk(:,3);


    %Number of censored data (=successes at ti-1)
    for i = 2:n;
        KM(i,3) = Data(i-1,4);
    end

    %Number of individuals
    A_1 = [KM(:,2) + KM(:,3)];
    for i = 1:n;
        A_2 = sum(A_1);
        A_1(i) = 0;
        KM(i,4) = A_2;
    end

    %Factor
    for i = 1:n;
        if KM(i,4) == 0;
            KM(i,5) = 0;
        else
            KM(i,5) = 1-(KM(i,2)./KM(i,4));
        end
    end


    %% S(k plus 1) waarden berekenen


    KM(:,6) = 1;
    for i = 2:n;
        KM(i,6) = KM(i-1,6) * KM(i,5);
    end


    Som = 0;
    for i = 1:length(KM(:,6));
        x = abs(KM(i,6)-Sk(i,2));
        Som = Som + x;
    end


    %% End
    %Som = 0.001
    if Som <= Marge
        Finished = 1;
    else
        Sk(:,2) = KM(:,6);
        Sk(:,3) = 0;
        KM(:,5) = 1;
        KM(:,6) = 1;
    end

    count = count + 1;


end

if CDF == 1;
    figure(1), clear clf, hold on;
    %plot(trainunits,startsurvival)
    plot(trainunits,1-KM(:,6))
    hold off;
elseif CDF == 0;
    PDF = zeros(n,3);
```

```
    PDF(:,1) = 1-KM(:,6); %failure function
    for i = 1:n-1;
        PDF(i+1,2) = PDF(i,1);
    end
    PDF(:,3) = PDF(:,1) - PDF(:,2);

    figure(1), clear clf, hold on;
    plot(trainunits,PDF(:,3));
    hold off

    %% Boxplot
    for i = 2:m_2;
        PDF(1,4) = PDF(1,3);
        PDF(i,4)= PDF(i-1,4) + PDF(i,3);
        PDF(i,5) = PDF(i,4) .* Boxplotvalue;
    end

    Cum = 0;
    for i = 1:Boxplotvalue;
        z = 0;
        for j = 1:m_2;
            if z == 0;
                if Cum < PDF(j,5);
                    X(i) = j
                    z = 1;
                end
            end
        end
        Cum = Cum + 1;
    end

    figure(2), clear clf, hold on
    boxplot(X)
    hold off

end

KM(:,7) = 1 - KM(:,6);

toc
```

F. Matlab script statistical tests

```
clear all

Topology_test = 0 %test between Kleine Binickhorst and Grote Binckhorst = 1, test
between the same locations = 0.

% Topology_test = 1
KB =
[0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0.0298020652821975;0.0542759760183325;0.05427597
60183325;0.0542759760183325;0.0542759760183325;0.0542759760183325;0.090881918165366
2;0.271230967155945;0.473019876970530;0.685506364806328;0.881981194107615;1;1];
% add zero at the beginning of the GB vector
GB =
[0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0.0298020652821975;0.0542759760183325;0.05427597
60183325;0.0542759760183325;0.0542759760183325;0.0542759760183325;0.090881918165366
2;0.271230967155945;0.473019876970530;0.685506364806328;0.881981194107615;1;1];

% Topology_test = 0
% add zero at the beginning of the GB vector
SF_1 =
[0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0.0298020652821975;0.0542759760183325;0.054275
9760183325;0.0542759760183325;0.0542759760183325;0.0542759760183325;0.0908819181653
662;0.271230967155945;0.473019876970530;0.685506364806328;0.881981194107615;1;1];
% add zero at the beginning of the GB vector
SF_2
=[0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0.0328410087146179;0.063439634124
4782;0.441864459922199;0.743714901647711;0.832129141745544;0.967622968877717;1];
%% Initial values
S = 100000; %Samplesize


%% Continu maken normalisatie



if Topology_test == 1

    Factor = 1.94;

    Interval_L = 0;
    Interval_R = Factor;


    KB = 1-KB;

    n = floor(length(KB)*Factor);

    T(:,1) = [0:length(KB)]';
    KB(length(KB)+1) = 0
    T(:,2) = KB;
    T(:,3) = T(:,1) .* Factor;

    N = zeros(n+1,2);
    N(:,1) = [0:n]';

    j = 0;
    for i = 0:n
        if i > Interval_R
            Interval_L = Interval_L + Factor;
            Interval_R = Interval_R + Factor;
            j = j + 1;

            N(i+1,2) = T(j+1,2) - (((i - Factor.*j)./Factor).* (T(j+1,2)-T(j+2,2)))

        else
```

```
        N(i+1,2) = T(j+1,2) - (((i - Factor.*j)./Factor).* (T(j+1,2)-T(j+2,2)))

    end
end


%% Random generator

%% Kleine Binckhorst

v = 1;
VKB = zeros(1,S);
for i = 1:S

    R = rand; %random value between 0 and 1

    x = 0;
    for j = 1:n+1
        if x == 0
            if R > N(j,2)
                value = N(j,1)-1;
                x = 1;
            end
        end
    end

    VKB(v) = value;
    v = v+1;

end


%% Grote Binckhorst
v = 1;

VGB = zeros(1,S);
m = length(GB);
M = zeros(m+1,2);
M(:,1) = [0:length(GB)];
GB(length(GB)+1) = 1;
GB = 1-GB;
M(:,2) = GB;

for i = 1:S

    K = rand; %random value between 0 and 1

    x = 0;
    for j = 1:n+1
        if x == 0
            if K > M(j,2)
                value = M(j,1)-1;
                x = 1;
            end
        end
    end

    VGB(v) = value;
    v = v+1;

end

%% Tests
%Medians
Median_KB = median(VKB);
Median_GB = median(VGB);
[p,SignificantInfluenceM] = signrank(VKB,VGB,'alpha',0.01)
```

```
    Mean_KB = mean(VKB)
    Mean_GB = mean(VGB)

    %Variances
    Variance_KB= var(VKB);
    Variance_GB = var(VGB);
    SignificantInfluenceVar_1 = vartest(VKB,Variance_GB,'alpha',0.01)
    SignificantInfluenceVar_2 = vartest(VGB,Variance_KB,'alpha',0.01)

    %KS
    SignificantInfluence = kstest2(VKB,VGB,'alpha',0.01)

    %% Boxplot figuurtje
    figure(3), clear clf, hold on
    boxplot([VKB',VGB'],'labels',{'Kleine Binckhorst (normalised)','Grote
Binckhorst'})
    ylabel('Capacity (number of train units)')
    hold off


elseif Topology_test == 0


    v = 1;

    V_1 = zeros(1,S);
    n = length(SF_1);
    N = zeros(n+1,2);
    N(:,1) = [0:length(SF_1)];
    SF_1(length(SF_1)+1) = 1;
    SF_1 = 1-SF_1;
    N(:,2) = SF_1;

    for i = 1:S

        R_1 = rand; %random value between 0 and 1

        x = 0;
        for j = 1:n+1
            if x == 0
                if R_1 > N(j,2)
                    value = N(j,1)-1;
                    x = 1;
                end
            end
        end

        V_1(v) = value;
        v = v+1;

    end



    v = 1;
    V_2 = zeros(1,S);
    m = length(SF_2);
    M = zeros(m+1,2);
    M(:,1) = [0:length(SF_2)];
    SF_2(length(SF_2)+1) = 1;
    SF_2 = 1-SF_2;
    M(:,2) = SF_2;

    for i = 1:S

        R_2 = rand; %random value between 0 and 1
```

```
        x = 0;
        for j = 1:n+1
            if x == 0
                if R_2 > M(j,2)
                    value = M(j,1)-1;
                    x = 1;
                end
            end
        end

        V_2(v) = value;
        v = v+1;

    end

    %% Test
    %Medians
    Median_1 = median(V_1);
    Median_2 = median(V_2);
    [p,SignificantInfluenceM] = signrank(V_1,V_2,'alpha',0.01)

    Mean_1 = mean(V_1);
    Mean_2 = mean(V_2);

    %Variances
    Variance_1 = var(V_1);
    Variance_2 = var(V_2);
    SignificantInfluenceV_1 = vartest(V_1,var(V_2),'alpha',0.01);
    SignificantInfluenceV_2 = vartest(V_2,var(V_1),'alpha',0.01);

    %KS Test
    SignificantInfluence = kstest2(V_1,V_2,'alpha',0.01)

    % Boxplot figuurtje
    figure(3), clear clf, hold on
    boxplot([V_1',V_2'],'labels',{'Not allowed','Allowed'})
    ylabel('Capacity (number of train units)')
    hold off
end
```