S.E.B. Janssens

Empirical analysis of Service Locations at NedTrain

A closer look at the stabling and handling capacity

TUDelft

Cover photo: (NS, Beeldbank, 2016)

Empirical analysis of Service Locations at NedTrain

A closer look at the stabling and handling capacity

Steef Edu Bernard Janssens
1504428

in partial fulfilment of the requirements for the degree of

Master of Science
in Transport, Infrastructure & Logistics

at Delft University of Technology

to be defended publicly on 10 February 2017

Committee:

Prof. dr. ir. S.P. Hoogendoorn	TU Delft - CEG	Transport and Planning
Dr. R.M.P. Goverde	TU Delft - CEG	Transport and Planning
Dr. ir. H.P.M. Veeke	TU Delft - 3ME	Transport, Engineering and Logistics
Ir. B. Huisman	NedTrain	Fleet Services

TU Delft

Preface

The thesis in front of you is the final step for the degree of Master of Science in Transport, Infrastructure and Logistics at Delft University of Technology. I worked on the study from 1 May 2016 until 10 February 2017. During this period, I was an intern at the department Maintenance Development at NedTrain in Utrecht, part of NS. My time at NedTrain also gave me the opportunity to look around, to talk with people and to visit various interesting locations.

Finally, I am very proud to present this thesis. Studying has not always been easy or obviously to me: I have had hard times, with various challenges, but of course also great times. I have shown, especially to myself, that going on and being perseverant finally brings you what you want and where you want to be. I have learned a lot: about the study topics, of course, but also about myself and the hurdles one can face. This graduation project was in that sense very representative for my time as a student.

I would like to thank all people directly involved in this thesis project: my graduation committee. Serge Hoogendoorn, for being positive and constructive and giving me the opportunity to do it my way. Rob Goverde, for always being available, giving relevant and practical comments and showing insight and interest in the topic. Hans Veeke, for being so flexible to enter the committee during the project. Bob Huisman, for welcoming me at NS and giving the freedom to develop my own project. It has been a pleasure.

Apart from the professional supervisors, I could always count on the people close to me. Marlie has been and still is a wonderful girlfriend to me. She has always been supportive and has always been there for me. And my parents, although critical, always have stood behind me. Thank you! My friends and fellow students, especially at NedTrain, have been great company during the days we have spent.

From now on, it is time for the next step. I take the experiences from this project with me and hope to become a proud TU Delft graduate. Enjoy reading this thesis!

Steef Janssens
Rotterdam, January 2017

Executive summary

NedTrain is the train service and maintenance subsidiary of the Dutch railways. The head office is in Utrecht and there are more than 40 locations spread over the Netherlands. The NedTrain department responsible for cleaning, washing and straightforward maintenance is Maintenance and Services, of which the Service Locations are part of. There are in total 33 Service Locations, all varying in size and equipment.

NS has ordered several hundred new train units for delivery up to 2024. By that time, more rolling stock will be in service. Therefore, the existing capacity of the Service Locations is not sufficient anymore. Within NedTrain, many studies are done about how to extend the capacity in the most efficient way. In this study, the empirical approach is chosen, to analyse what the actual capacity and characteristics of the Service Locations and the related stabling yards are. The capacity of a Service Locations is the number of coaches from train units that can be handled during a time frame, mainly the normative night shift. The limiting factor is when the capacity of one of the tasks is reached. In order to define the capacity, a generalised insight of the production of today at the Service Locations is considered. The scope of the study is on infrastructure, rolling stock and process variance.

The study consists of qualitative and quantitative analysis, both from an empirical perspective. It is done by evaluating the following research question:

How can the capacity of the Service Locations and stabling yards, given the characteristics of that location, be estimated, as a result of empirical analysis?

Qualitative and quantitative analysis

The tasks at a Service Location can be divided into five categories: checks, repairs, cleaning, washing and shunting. For this study, all Service Locations are considered. The data of realised work orders for twelve months (August 2015 – July 2016) for all locations is prepared. The analysis of both combined results in a detailed insight in the performance of the Service Locations. All rolling stock used for domestic services is included in the study.

Checks are the safety (B-check) and comfort (A-check) inspections. The interval of a B-check is in general every 2 days, the A-check every 12 days. The A-check takes up to three times as much time as a B-check. A train does not have to be moved for a check. All checks together are about half of all work orders at all Service Locations. The checks are a standardised task and therefore the same at all Service Locations, so directly comparable. However, realised average lead times of the checks vary much between the locations. The average lead times also vary per rolling stock type: some checks take twice as long for types compared to others.

Repairs are subdivided in five categories, from preventive repairs to direct repairs. There is a very large variety in size and duration of the repairs. Therefore, they are not directly comparable to each other. Some repairs require dedicated equipment, for example to work at roofs. This has considerable influence on other tasks, since trains has to be moved over the location. Repairs are about a third of all work orders. There are locations that do almost no repairs, especially the smaller Service Locations.

Washing is the cleaning of the exterior of the train. Only 12 Service Locations have a washing machine. There are soap washings and oxalic washings, at 5 out of 12 washing machines the oxalic washings can be done. Oxalic washings have a longer interval of 63 days, compared to 7 days for soap washings. There also is a large variation within productivity of the washing machines at the various locations. The performance of washing compared to standards is poor: only 45% of the planned washings are done in time. Washings are not planned for each task on beforehand, resulting in this poor result.

Cleaning is outsourced to a third company and therefore not considered in the study. Shunting is a shared task with NS Reizigers. The main logistic task is to get all trains ready and at the right locations to start the next service. Especially during the night, this happens a lot. Also for work at the Service Location shunting is needed sometimes. Some tasks that need equipment require shunting, like washing. This takes a lot of capacity, since the infrastructure is occupied for longer time, limiting the shunting possibilities for other tasks during that time.

The Service Locations all have different characteristics. The infrastructure layout and related process is divided in two options: the carousel and the shuffleboard layout. With the shuffleboard, trains are first parked, mostly at dead end tracks, and serviced. The carousel is a more rotating process, where trains are moved while they are serviced, to visit for example a cleaning platform or a washing machine. Most locations have a shuffleboard layout. There is also a hybrid layout type, without clear characteristics of one of the two others, which is called the station layout.

Eleven rolling stock types exist within the NS Reizigers fleet, with in total 18 compositions, that are all considered in the study. All train units consist of coaches: some are two coaches, others eight, including two locomotives. Rolling stock of different types have general characteristics: some types have considerably more work orders than others. Also the distribution of work orders types is different for all train types. The ratio of repair work orders compared to check work orders is mostly leading in that situation.

The variation in work is high. The number of realised work orders at the same location during the same night of the week can be double the work of another day. Also the number of train units worked on fluctuates in the same way. This is a result of the random assignment of specific train units to Service Locations. The work on specific units is always different than the work on other units and can never be predicted on beforehand. There are busy days and quiet days at the Service Locations.

Data analysis

The data set with realised work orders is used to find relevant links between characteristics of the Service Locations. These links are correlations, as well as other relationships. A correlation is only feasible if it is statistically relevant. It is assumed that all links have a linear appearance, since the fit is in all cases the best for this assumption. The correlations are studied per layout type.

The stabling capacity and the number of tracks available at a location have a positive correlation with the number of realised work orders. The spread of the data points is considerable, but for each of the layout types, a good fit is indicated. Another positive link is between the presence of a washing machine and the production. A washing machine itself generates work orders, but Service Locations with a washing machine also have more work orders of others type compared to other locations.

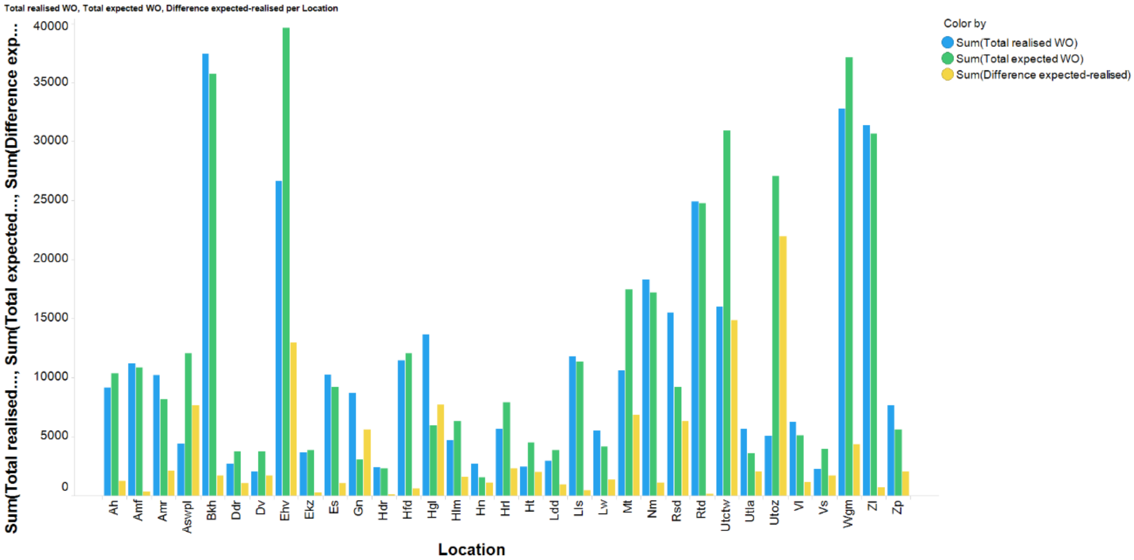
For each train unit worked on at a Service Locations during a shift, another train unit is present at that location which is not worked on. That train unit is only subject to shunting and is stabled at a Service Location.

Model

From all topics studied, it is proved that the factors of influence on the production of a Service Location are the layout type, the stabling capacity, the number of tracks available, a washing machine available, the position within the network and the rolling stock type. These factors are all included in a model, that describe the predicted production of a location. The model is based on the linear regression model from the relationship between number of tracks and number of realised work orders. The model is expanded with factors that describe the influence of the other parameters. The predicted production of a Service Location is used to estimate the number of train units stabled at a location. This is finally translated into the number of coaches that can be stabled at a location. The stabling capacity is assumed to be related to the maximum handling capacity of today at a location. The modelling approach is the following:

1. Select input variables: layout type, number of tracks, presence of a washing machine, washing machine type (soap or oxalic), at the end of the network or not
2. Calculate expected number of work orders using the model
3. Select mix of train types: number of train units of each train type
4. Calculate predicted number of work orders for mix of train types
5. Check: predicted number of work orders has to be smaller than expected number of work orders
6. Calculate predicted number of coaches stabled based on work orders
7. Calculate total predicted number of coaches stabled

The model results are compared to the existing data of the real Service Locations. Also a test data set of work orders from previous months is used to verify the model results. The results are visualised in the illustration below, with the number of work orders for each Service Location. For 18 out of 33 locations, an estimation is made within 30% of the realised work orders. This study cannot declare the differences for individual locations that could not be predicted accurately.



Conclusions and recommendations

The complex and dynamic process at the Service Locations is simplified to a few factors of influence in this study. Based on that, a reliable estimation of the expected production of a Service Location, including coaches to be stabled, can be made. Other factors that were expected to have influence or could describe a location, such as size of a location, could not be validated. The variation within the process could not be estimated properly, as a result of a lack of data. Due to that, there is an uncertainty within the results for the expected number of coaches stabled. The results gathered give good insight in the average performance of the Service Locations. The approach of the study cannot declare all phenomena at all locations.

It is recommended to improve the registration of train units present at Service Locations, in order to further improve study to the occupation of the infrastructure. A more integrated planning of tasks will result in less missed tasks, especially regarding washings. An integral approach for the interests of both NedTrain and NS Reizigers will also result in a better and more reliable availability of train units.

Table of content

1 Introduction

This chapter gives an introduction to the company, the research problem, the purpose of the study and the methodology used. After the problem statement, the research questions are illustrated and the method to answer the questions and to solve the problem is given. The literature references are included in this chapter as well.

1.1 NedTrain

NedTrain is a Dutch-based service provider for train maintenance. It is a subsidiary of NS, the Dutch national railway company, and is currently being fully integrated into NS. NedTrain offers full-service train maintenance, from line maintenance and failure control to complete overhauls. NedTrain is well-known from its overhaul workshop in Haarlem (Refurbishment & Overhaul), but there are many more locations. The organisation is visualised in a simplified way in Figure 1, with its main branches and the departments of interest for this study being given. The NedTrain department which is responsible for the cleaning, washing and straightforward maintenance is called Maintenance and Services, with the head office in Utrecht and over 40 locations spread over the Netherlands. The more strategic and research driven department is called Fleet Services, of which Maintenance Development is a subsidiary, being the client for this study, also based in Utrecht.

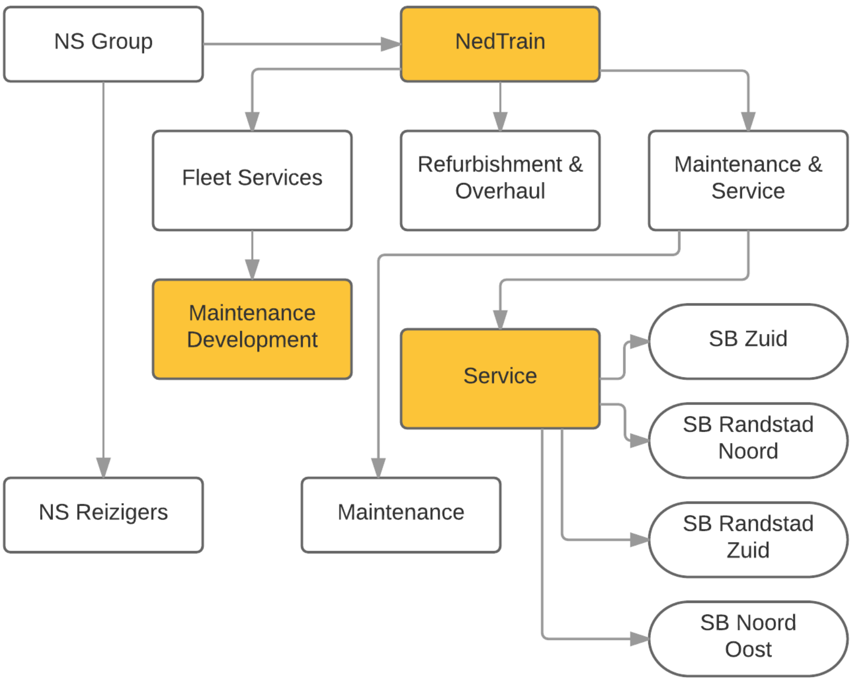


Figure 1: NedTrain organisation chart

Maintenance and Services is responsible for the Service Locations (in Dutch: Servicebedrijf (SB)), where the trains are treated mainly during nights. At those locations, of which 33 exist spread over the Netherlands (see Figure 2), the trains are cleaned inside, washed outside, are subject to technical checks and small repairs are performed. The cleaning is outsourced to a third party, although this company is controlled by NedTrain. The repairs and checks are performed by NedTrain itself and the washing is done in close cooperation with NS Reizigers. For all types of work, a regular schedule of frequencies is available, based on which the work is planned. A more detailed description of what is done at the Service Locations is given in chapter 2.

For the upcoming years, NS has ordered several hundred new train units (NS.nl, 2016), which only partly replace old trains. This will result in many more trains in operation and thus also more trains to be handled over the years at the Service Locations. As a result of a first analysis from NedTrain, it seems that the capacity at the Service Locations that is currently available could not be sufficient in order to handle all new trains up to the year 2024. Due to this statement, a strategic solution has to be made about how to extend this capacity. In order to solve the problem, in this study the realised work nowadays is considered. That is done to get an insight in what the capabilities of the Service Locations exactly are and what can be learned from that.

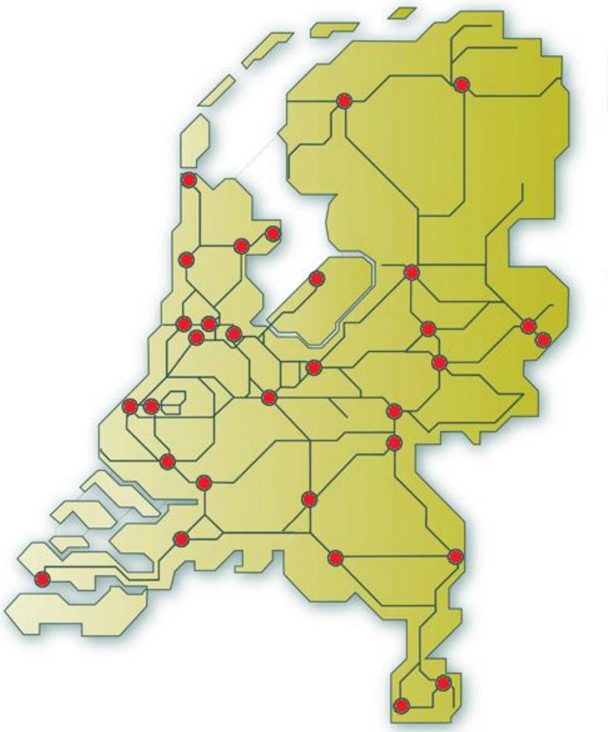


Figure 2: Service Locations (based on (NedTrain, Overzicht van de NedTrain locaties, 2015))

1.2 Problem and purpose of the study

As mentioned in paragraph 1.1, the capacity currently available may not be sufficient to handle all trains in the correct way in the upcoming years. Based on the information gathered, insight in the capacity of the Service Locations is given from an empirical point of view.

1.2.1 Problem statement

Due to a higher demand for service in the years up to 2024, the capacity at the Service Locations might not be sufficient.

Since capacity is the general KPI in this study, the definition of it is of key importance. At NedTrain, the capacity of a Service Location is expressed as the number of coaches which can be handled during a night (which is the normative work shift). This implies that, given the variation within the process, this is the probability that a number of coaches can be handled given a planned schedule, with a certainty of 95%. The production, which depends on the capacity, is the number of coaches out of all coaches which are serviced in the correct way. This production can be expressed in percentages for all tasks, indicating the performance of a Service Location.

By choosing the empirical starting point, as is done in this study, a new approach will be added to the existing knowledge at NedTrain. Research so far has been done to the topic based on the theory and there is also a team working on the topic, but with the aim to suggest process improvements. Besides that, several theoretical studies take place, also considering the capacity, but based on standards and with the goal to optimise the process. In this study, practice will always be leading, while planning, which is theory, is compared to what is realised to give insight in the difference between the two. This is relevant for NedTrain, since there is no proper information about this yet.

As mentioned earlier, the capacity currently available is not sufficient to handle all trains in the correct way in the upcoming years. In this study, the service process as used nowadays is considered as given. A redesign of this process is not in the scope of the study, but research about the process is necessary to get insight in the problem. The planning problem for a Service Location is not the same as the approach chosen in this study. Planning will help to optimise the performance of a location. It is a solution for the problem, while in this study the process is only shown and declared.

1.2.2 Purpose of the study

Given the research problem, the purpose of the study is to give a generalised insight in what the performance of the Service Locations is, by indicating the capacity of a Service Location. In order to do this, the capacity of a location, given its characteristics, has to be predicted, based on the information gathered about all Service Locations. A schematic representation of the study is given in Figure 3. This will in the end lead to a by now called 'black box', the model, which will give the expected performance of a general Service Location, given its characteristics, such as the infrastructure layout or the rolling stock to be serviced. By using this model, the expected work to be done by a Service Location, existing or non-existing, and thus the capacity of that location can be predicted.

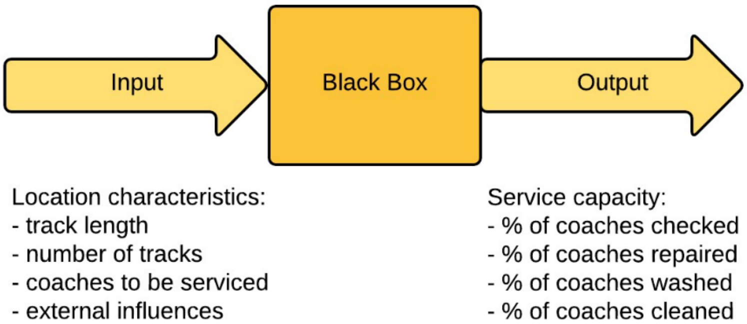


Figure 3: Study goal

In order to understand the process, all the work done at the Service Locations, this process is described in a detailed way. The process description is a study purpose on itself. By using this process description, the bottlenecks can be pointed out, which will be used to focus the study on.

The model has to be able to predict the capacity of a Service Location. By giving its characteristics, the model will state what the expected capacity is. This can be used by further development of Service Locations. By taking the model and the work that is expected at a location, it can be possible to predict whether the desired location will match the requirements. By changing characteristics when necessary, the model can be used to optimise a design for a Service Location.

1.2.3 Scientific relevance

The scientific contribution of the study is diverse. Empirical analysis of processes is applied widely, but not for the specific train maintenance or service process, in the Netherlands or elsewhere. The capacity of a service facility is also not yet defined based on empirical analysis. Most studies are based on mathematical models and make use of theoretical data. This study uses a data set from practice, making the results both realistic and actual. Other studies are more focussed on process improvements or solving the planning problem. This study combines the existing knowledge and experience from practice by taking a different approach. Statistical tests of the data support the conclusions.

1.3 Research questions

In order to solve the problem stated in the previous paragraph, research questions are formulated to guide the study.

1.3.1 Main research question

The exact capacity of the Service Locations and the stabling yard linked to this location has to be defined based on the realised work at all locations. With this result, there will be insight in the capacity from an empirical point of view. Based on that, a model can be defined, which will give a generic idea of what the performance of a location is, given the characteristics of that location and the input of the location. The input, the work to be done and the process itself, so how the work is done, is considered as fixed. This information results in the following research question:

How can the capacity of the Service Locations and stabling yards, given the characteristics of that location, be estimated, as a result of empirical analysis?

The planning for the process is also input, but it is not part of the main question of this research project. This planning is used as a reference for the performance measurement.

1.3.2 Sub questions

Numerous sub questions are posed, in order to come to a comprehensive answer to the main research question. The questions have to be answered in order to give a complete overview of the problem stated and to give insight in what is important. Those sub questions are all treated in the study, are answered explicitly in the conclusion of the report (chapter 6).

1. What data is needed to get insight in the production of the Service Locations?
2. What is the variance in the input of the service process?
3. What are the characteristics of the different layout types of a Service Location?
4. What is the influence of the infrastructure and layout of a Service Location?
5. What is the influence of the different trains types treated at the Service Locations?
6. What are the exact functions to be defined within the process?
7. What is the difference between planned work and realised work in number of coaches for each task?
8. What correlations exist within production between various characteristics, if they exist?
9. Where are the bottlenecks, within the stabling or the servicing and in which step within the process, and what are those bottlenecks?
10. What is the difference between theory and practice in terms of the process steps?

1.4 Methodology

For this research project, various research methods are used. The main method is data analysis: descriptive as well as predictive analysis. Literature study will be done throughout the whole period of study.

1.4.1 Research design

Various research methods are used. In the correct order, literature review, interviews, process description, empirical data analysis and modelling will be used.
Many of the qualitative results will come from own research, as this will be based on visits and interviews executed at the NedTrain locations. As the characteristics differ a lot between the Service

Locations, many visits will be needed in order to get a complete view of the situation. During the interviews, a fixed schedule will be followed.

1.4.2 Data review

An important section of the study is about data review. The data used is the planned and the realised work for each Service Location over a period of one year. This data is used to determine the performance of each location. That is done by comparing this data to the characteristics of the locations. These characteristics consist of the appearance of the location, the equipment available and the work that is done at that location.

In order to make use of the outcome from the data analysis, various hypotheses are drafted and tested. By doing this, the desired characteristics of a location will become clear. To make the conclusions more feasible, they are supported by statistical results from the data analysis.

1.5 Report structure

This report consists of six chapters, extended with multiple appendices. A list of definitions is given in appendix 1 and a list of abbreviations is given in appendix 2. All chapters start with an introduction and end with a concise conclusion.

This introduction chapter is followed by a qualitative and a quantitative analysis, that are linked to each other. The qualitative analysis is described in chapter 2. That chapter consists of a description of all tasks, of the layout of the locations and of the rolling stock. It roughly shows the expected production, based on standards. The chapter is concluded by a detailed description of all processes. Chapter 3 is the quantitative analysis. That chapter shows the performance of the Service Locations from various relevant perspectives. It shows the relevant tasks and the distribution of work orders. The differences between performance at Service Locations of various rolling stock types are indicated. It also gives an overview of the variation within the work at the Service Locations.

Based on the information gathered from the qualitative and quantitative analysis, a data analysis is made in chapter 4. In that chapter, statistics are used to find relevant links between characteristics of Service Locations. By doing that, correlations are found and distinction between layout types is indicated. Also characteristics that do not have a relation to production are indicated. The correlations from chapter 4 are used in chapter 5, where a model is made. The model is using all information previously gathered and describes the production of a Service Location. This model is based on the characteristics with a proved positive relation to production. Conclusions and recommendations about the study and further research are given in chapter 6. That chapter also concisely answers all research questions.

1.6 Literature references

The literature study is used to gather knowledge from previous studies in a broad sense. Various aspects of the problem are looked after, with a summary of the results given below.

1.6.1 Service processes

A general introduction to the processes at NedTrain can be obtained by various studies (Lentink, 2006) or conference papers (Busstra & Dongen, 2015; Wilson, Roos, Huisman, & Witteveen, 2011). Although these studies mainly focus on process improvements, while this study focuses on process analysis, major parts are still useful. Information about location considerations (Busstra & Dongen, 2015) give insight in the available workload at locations.

Considering the servicing process and turn-around times of trains, a lot can be learned from the airline business (Wu & Caves, 2002). Service processes of aircraft are planned carefully and exact, with a fixed ordering within the tasks. Similarities between aircraft and trains are that utilisation during the day is high and that a lot of money is involved when the units are not available for service in time. With aircraft turn-around, a lot of processes have to be terminated on time, while trains can run less constrained throughout a day. In order to denote the process steps, an aircraft rotation model is developed, being useful for the analysis of Service Locations as well. Another process that can be considered is when a vessel arrives at a port. More about this topic is described in (Olba, Daamen, Vellinga, & Hoogendoorn, 2014). A vessel usually is a subject to multiple processes when it is at a port: unloading, loading, provisioning or maintenance. To determine the capacity of the links coupled to these processes, some methods are presented. It is done by identifying and defining the main constraints and bottlenecks of the separate processes. The analogy between the study to port service processes and train service processes is evident and therefore, the method description used in this paper is useful for this study.

1.6.2 Infrastructure capacity and planning

A lot of research is done about open line railway capacity, also under disturbed conditions (Goverde, Corman, & D'Ariano, 2013; Quaglietta, Corman, & Goverde, 2013). Although, the calculations from these studies cannot be translated directly into capacity calculations for shunting yards and thus for Service Locations. The capacity of shunting in stations and shunting yards and the coherence between the two is described as well (Freling, Lentink, Kroon, & Huisman, 2005). Especially time and the minimum headway between two shunting movements are a complicating factor in shunt planning. Also the railway infrastructure is complicating, since train units are restricted in their movements. The general approach to establish a feasible railway timetable is described in (Hansen & Pachl, 2014). Two methods to determine the capacity of rail infrastructure during shunting works are described in (Broek & Kroon, 2007). One method is based on an approach with fixed routing and the other with variable routing. This study however only considers the influence of those shunting processes on the timetable by looking at the shunting movements on the main tracks, so not at the shunting yards themselves. A full description of only the shunting process at shunting yards and interactions between shunting trains could not be found in the literature.

For the shunt planning, considering the infrastructure capacity, an algorithm is developed by (Lentink, 2006). This thesis, combined with the study by (Broek & Kroon, 2007), gives good general reference points for the shunt approach in this study. A planning methodology is developed, based on the actual procedures for shunting. Especially the infrastructure occupation in the case of saw movements is pointed out. Saw movements are needed when a shunting train cannot reach the desired track in one forward movement. In order to give the driver the opportunity to walk to the other end of the train, extra time is planned, resulting in a long track occupation.

Capacity calculations for roads are often done with free-flow traffic conditions. In reality, the capacity is influenced by congestion. This principle is described (Leclercq, Knoop, Marczak, & Hoogendoorn, 2014; Daamen, Loot, & Hoogendoorn, 2010) and a framework for the interaction between upstream congestion and downstream consequences is given. Although the study is developed for freeway merges, it is also applicable for railway yards, since the principles are the same, but with different capacity restrictions. Also the infrastructure characteristics are different, since for rail they are more restricted than for freeway flows. Following the study, a more congested road results in less capacity, so the capacity is not linear, but decreases from a certain point.

1.6.3 Process description and quantification

One of the primarily goals of the study is to make a process description of the service process at the Service Locations. A proper process description tool is given in (Veeke, Lodewijks, & Ottjes, Conceptual design of industrial systems: an approach to support collaboration, 2006), that is a base for the method described as the Delft Systems Approach (Veeke, Ottjes, & Lodewijks, The Delft Systems Approach, 2008). Both refer to the complex nature of many processes and stipulate the importance of dissection into main components. By doing this in the right order, a comprehensive insight in the problem may be gathered and, following the theory of the approach, the study will grow in some natural way, as steps follow each other in a logical order. In order to give a proper decomposition of the process, the CATWOE principle may be used, which defines the customers, actors, transformation, world view, owners and environment of the situation considered. This method is described in more detail in (Bergvall-Kareborn, Mirijamdotter, & Basden, 2004), although the focus in this study is on the human activity and not the system only. This is called the soft systems methodology. The method described also defines an intermediate step between analysis and modelling of the system, what is useful further on in the study.

Important for the analysis of the performance of the Service Locations is the quantification of the process steps. The lean approach is a hands-on approach to the problem (Lean Enterprise Institute, 2003) (George, 2003). Lean focuses on the value adding steps within the full process by identifying waste, which is time not being used for production. The waste can for example be identified by using the six sigma approach, which is used for optimisation of processes. A very useful and evident tool is the value stream map, which shows visually the value adding time apart from the non-value adding time. The identification and distinction of process steps to make use of this tool is also applicable in case of this study, where the process steps have to be valued and quantified. Although lean is mainly developed for process optimisation and improvement, the basic idea behind it is applicable to the study done here. Also the distinction of different states within the process is made, which is a relevant approach for further use in this study.

2 Defining NedTrain Service Locations

In this chapter, a qualitative description of the NedTrain Service Locations is given. This description is based on location visits, interviews and literature review. The chapter starts with a theoretical description of the attributes of the study. After that, a detailed process description is given. Consequently the qualitative description will serve as input for the quantitative analysis.

2.1 Service Locations

The 33 Service Locations are strategically distributed over the country, most of the time at end stations and also close to the workshops. All Service Locations have different characteristics and thus different capabilities. For example, about one third of the locations has a washing machine, while almost all locations have cleaning facilities. Besides that, the scale and physical appearance of the locations varies, as all locations are different. The infrastructure layout strongly influences the work methods at a location. Also the performance and the quality delivered by the different locations varies. Most locations have a 24 hour staff occupation, in three shifts. The work intensity is the highest during the night shift, since most rolling stock at this time does not have to be available for running trains and thus can be serviced.

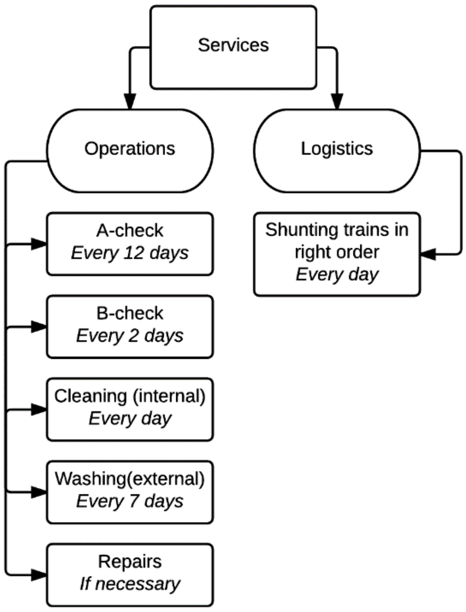


Figure 4: Processes at Service Location

At the Service Locations, a number of trains arrive each night based on a planning schedule. The specific train sets and the composition of the work to be done can be different every day. The two main activities to be done on a train set are operational and logistics, as can be seen in Figure 4. The main tasks at a Service Location can be seen in that figure as well. The actual process is planned locally every day, based on the input the planner receives from the circulation of the trains at that day, provided by NSR. Each train set has its specific work to be done and also typical malfunctions, based on the actual state of the train set at that moment. This implies that the work executed at a location is different every single day, due to the fact that different train sets arrive at a location every other day. The capacity of a Service Location is expressed in the number of coaches that are serviced per night.

The physical appearance of the Service Locations is strongly related to the process type. It can roughly be divided into two types: the carousel and the shuffleboard. The difference between the two is a rotating train set within a carousel and an up and down going train in the shuffleboard alternative. The distinction for one of the two types is made based on history or on the geographical possibilities of a certain locations. Both types have their advantages and disadvantages in terms of capacity and reliability of the process. The appearance or layout of a Service Location is adjusted to the service process type within that location. The influence of this appearance is one of the main topics of this study, since it will become clear which type of a location performs better.

2.2 Characteristics of tasks

As mentioned, different tasks are executed at Service Locations. A rough insight can be obtained from Figure 4. The majority of tasks can be executed at all locations, but not all locations can manage all tasks. The more complicated maintenance, for example on top or at the carriage of a train, can only be done at a few locations. Also, only twelve washing machines are installed, distributed over the country.

An essential distinction can be made within the service processes, as there are tasks related to the cleaning of the trains on the one hand and technical tasks on the other hand.

2.2.1 Checks

Two types of checks exist: the A-check and the B-check. The B-check has to be done every second day and takes about 20 minutes for a mechanic to be fulfilled. An A-check is more extensive and takes an hour to be done, while it has to be done every twelve days. A B-check contains visual inspection of all major components, such as parts of the braking system or the condition of the pantograph. With an A-check, the functionality of certain components is checked, on top of the things checked with a B-check. An example of this extra work involved in an A-check is the check of the condition of the interior of the train unit.

Some trains have aberrant check intervals. For trains running at the high speed line, a special term exists, the 24 hour-check, which implies that it has to be done every day. A 24 hour-check is almost the same as an A-check. Besides that, the A-check interval of DM90 is once every three days, while SLT-train sets are subject to a B-check every day.

All checks have to be performed on time, otherwise the train will not be released for revenue service, which will lead to a shortage in available rolling stock. It may result in short trains or even abolished trains. This is why planning the checks is a major task for the Service Locations.

2.2.2 Cleaning

Cleaning consists of the internal tidying and cleaning of all compartments of the train, from the seats to the toilets. The process is outsourced to a third company, which plans and executes the work fully independent from NedTrain. Cleaning needs to be done every day for all trains. It can be done at many locations, as only a minor equipment level is needed. Trains which are not cleaned can be excluded from revenue service, but this is not regular policy.

2.2.3 Washing

Washing is the external cleaning of trains. This can only be done at a washing machine (TWI), of which twelve exist around the country. The work in the washing machines is executed by NedTrain, in direct cooperation with NSR. Washing is not included in the general NedTrain planning at Service Locations and is only executed if possible, regarding occupation of a location and logistic purpose. Due to this lack of planning, it strongly influences the continuity of the process, as it has a big impact on the logistic task and track occupation at a location. All washing machines are different: some are fast (Enschede), others are very busy (Binckhorst). In most machines, the train has to run on continuous low speed through the machine, while in Rotterdam the train is parked inside the machine.

The trains have to be washed with soap every 7 days, and with oxalic every 63 days. The oxalic is a stronger cleaning solution, which also treats the dust from the overhead wires. The washing intervals are guidelines: a train will not be excluded from service in case it is not washed on time, unless it is a really extreme case. Washing is in general the process which has the lowest priority and therefore it is not always done in time. Besides that, washing trains might become more necessary during the winter than in summer due to the weather, as more dirt exists in winter time.

A special branch of washing is the graffiti removal, that is executed as fast as possible. It is done by the cleaning company. The process requires special equipment, that is only present at some locations, and therefore it influences the logistic process.

2.2.4 Repairs

A broad variety of repairs is done at the Service Locations. Locations have various equipment levels: some have a Technical Centre, with full equipment, others have a platform or a pit, but most locations do not have any special equipment. It depends on the location what can be repaired, regarding complexity and equipment. The repairs can be split into two types: the necessary and the desired work. Necessary work has to be done to keep a train safe and thus available for service, for example a change of the shuttle of a pantograph. Desired work is not obligatory to run the train safely, but is comfort related, for example the change of a seat when it's damaged.

2.2.5 Shunting

The main task at the Service Locations is shunting the trains units in the correct order, as the majority of the trains running during the day start their service from a NedTrain-location. Mainly during night, this has to be planned carefully, to make sure all trains are ready for service the next day. Planning is based on the demand from NSR for the availability of specific train units at specific moments. This demand is used to create a basic shunt plan for all train units present at a location, with all movements between arrival and departure. The logistic process itself, all shunting movements, can be quite intense, as it depends on infrastructure availability and the restrictions that some train units have a dedicated circulation on a specific day, for example to make sure they will end their service near a workshop. This implies that all train units are planned to specific trains and may therefore not be interchanged with another train unit of the same type. All basic movements within a shift are planned, to make sure there is a feasible plan. A complication is that movements corresponding to washing are not planned, as well as necessary movements for checks or unplanned repairs. Those movements have to be done in between the planned tasks and therefore they influence the capacity.

When the logistic task cannot be completed correctly, it will result in too short trains, trains with a wrong composition or even cancelled trains. The shunting is performed by a mixture of NSR drivers and dedicated NedTrain drivers, which work in cooperation and are available throughout all shifts. It depends on the location whether dedicated shunt drivers are available. While the mainline tracks are all centrally controlled by ProRail controllers, the majority of the tracks at Service Locations is under control of local NedTrain controllers. Their cooperation in the logistic task is limited, although they know exactly where each train unit is at each point in time. In theory, the controllers only perform planned tasks, but in case of disruptions, they have strong influence on the activities at a Service Location and the allocation of train units to available tracks.

2.3 Characteristics of locations

Due to the different tasks to be done at the locations, the appearance of the locations is not directly comparable to any other location. This is because of a variation in layout, equipment or size. The characteristics that can be of influence on the capacity are discussed in this paragraph.

2.3.1 Layout

At the Service Locations, two main types of layout occur. The first one is the shuffleboard, which originates from a central switch complex and from there on, various dead end track branches exist, so-called stacks. In this configuration, trains have to leave the location in the opposite direction as they originate from. This also implies that trains can be parked behind each other and that the last arriving unit has to leave first, as train units can be blocked by others. That characteristic can make the logistic task more complex, as the sequence is crucial. The servicing facilities can be at the dead end tracks or at other tracks, with staying and servicing separated, what makes moves within the process necessary.

The second layout type is the carousel, in which trains can pass continuously, as there are switch complexes at both ends of the location. Tracks can be used in both directions in this situation. At locations with this type of layout, the service facilities are located at the main routes within the location. The stabling location is situated in between the facilities. This layout type has a proper logistic advantage, as it offers more flexibility and the order of the trains at arrival is less crucial, as the risk of locking is lower. On the other hand, within this location there also exist switches which are used very intensive. Both layout types are illustrated in Figure 5 and Figure 6. The layout type of each Service Location is given in Appendix 3, including track length and the specific equipment for each location. In case no specific layout type could be derived, it is called a station.

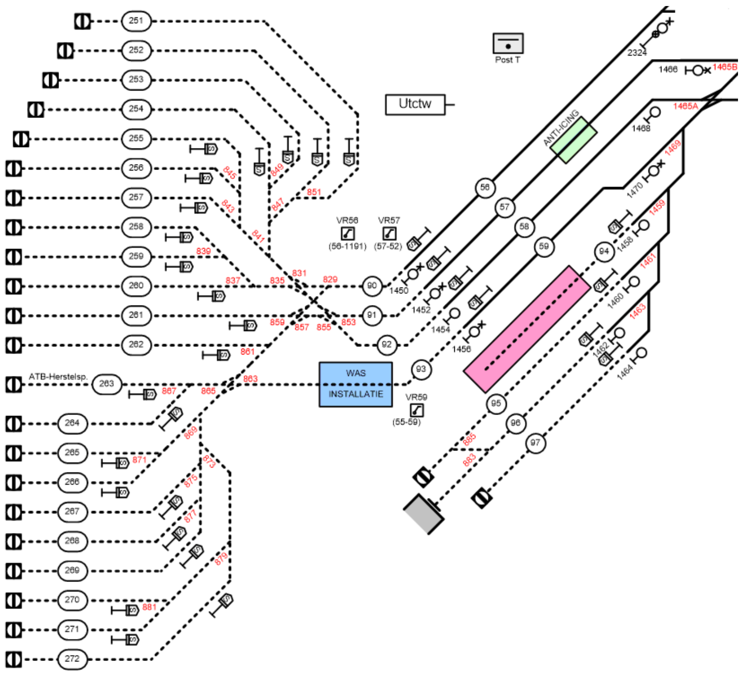


Figure 5: Service Location with shuffleboard layout (Utrecht Cartesiusweg) (Zeegers, 2016)

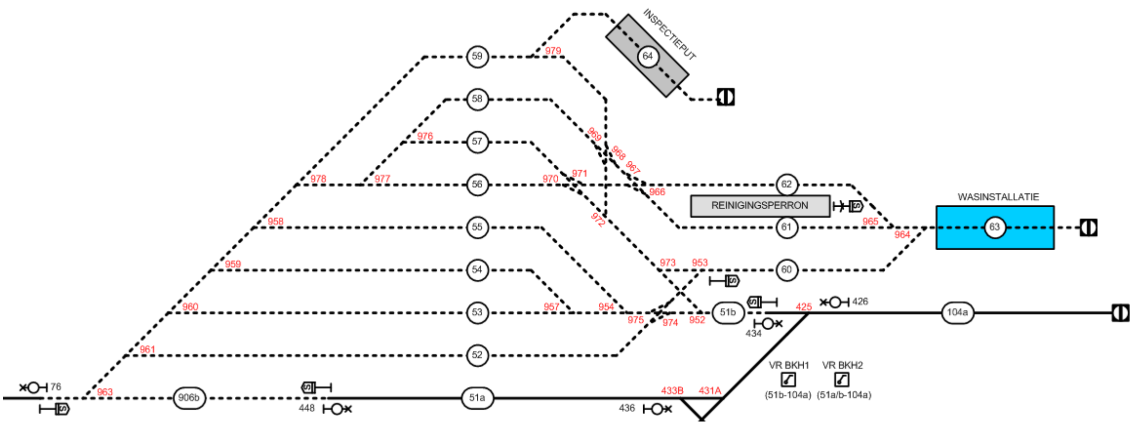


Figure 6: Service Location with carousel layout (Den Haag Kleine Binckhorst) (Zeegers, 2016)

2.3.2 Equipment

The equipment available at a location strongly depends on the tasks to be done at a location, described in paragraph 2.2. An overview of the equipment available at all Service Locations is given in appendix 3.2. This appendix shows that twelve locations have a washing machine. Other equipment is more generic, but there are differences in the appearance of this equipment. A general assumption throughout the study is that catenary is available at all tracks considered, to make sure all train units can be handled at all tracks.

Cleaning facilities are available at all locations. Those can be divided in two alternatives: a cleaning platform and mobile stairs to enter the train. The platform is the rarer alternative, as it has the downside that inspections cannot be done parallel, as the carriage of the train is not attainable. The benefit of a platform is that the dumpsters can be placed next to the train, which saves transport time. At some locations with a cleaning platform, there is only one track with such a platform, which creates the need to move within the location. A cleaning stairs is more flexible, as it can be built in between tracks and it can be moved easily to the desired location. Moving the garbage out of the train costs a lot of effort for the cleaners, but the benefits of those stairs make this the most widely spread solution to enter the trains.

About half of the locations have an aerial platform, which is used for maintenance at the top of the train, for example at the pantograph. More and more systems have been moved to the roof on modern trains, so this equipment becomes more important. To work at wheels or systems under the train, also working shafts are widely spread. Some locations have also dedicated tracks to work on for example ATB, the train protection system, with the equipment to test the installation after repair. For all this equipment it applies that locations have at most one of it.

There are five Service Locations that have a so called Technical Centre, in which various equipment is combined under a roof. At these facilities, the more complex repairs can be done. These Technical Centres are formally part of the Service Locations, although the typical service work is extended with work which was formerly done at Maintenance Facilities. Despite this, work orders from the Technical Centres are taken into account in this study, as also many checks and regular work is done there and the staff is the same as at the Service Location.

The washing machine appearance is various, since all washing machines are different. In most washing machines, the train has to run on low speed and has to stop to wash the heads of the train. There is also a washing machine where the train has to stop while the machine moves along the train. The last variant is the high-speed washing machine, which is not capable to wash the heads, but which is up to four times as fast as a standard washing machine. All washing machines are capable to wash with soap, five are also equipped for oxalic washings. At all Service Locations, the washing machine has an isolated location, to make sure the influence on other tasks is minimised.

2.4 Characteristics of rolling stock

The input and output of the Service Locations is the rolling stock of NSR, which are the passenger trains for operation on the Dutch main railway network. A general characteristic of the Dutch trains is that they appear at almost all locations, as they are used throughout the country. This implies that all Service Locations are capable to handle all rolling stock types. An exception is made for the diesel motor units DM'90, which only appear at dedicated parts of the network in the eastern part of The Netherlands, and the TRAXX-ICR compositions for operation on the high speed line, which are only serviced in Watergraafsmeer (Amsterdam) and Rotterdam. The other exceptions are the high-speed trains, Thalys and ICE, which are serviced at Service Location Watergraafsmeer. Those train units are not considered in the study, since they are an outsider among the other Dutch trains and their number is low. Tracks at Watergraafsmeer that are reserved for international trains are excluded from the study, minimising the impact of those train units on the remaining service process.

A general characteristic of the rolling stock of NSR is that all trains are train units, which means that they have driver cabins at both ends of the train. This is a relevant characteristic at Service Locations, since it improves possibilities for efficient shunting. All types of rolling stock currently used by NSR and handled at the Service Locations are listed in Table 1. Some types do not appear in the data during the whole period considered (see chapter 3.3), as they are phased out (MAT64, since April 2016) or reintroduced (DDM, from June 2016), but they all influence the data. All train types are electric train units, except for DM'90 which is a diesel train unit. All other train types, not mentioned in Table 1 and used on the Dutch railway network, are out of scope of the study. As can be seen in the table, the number of units per train type varies, which already is an indication of the number of Service Locations where a train type occurs. In total, 689 train units have been in service over the year considered.

Type	Number of coaches	Length (m)	Length per coach (m)	Number of units
SGM-II	2	52,2	26,1	30
SGM-III	3	79,0	26,3	60
SLT-IV	4	69,4	17,4	69
SLT-VI	6	100,5	16,8	62
MAT64	2	52,1	26,0	37
DM'90	2	52,3	26,1	22
DDAR	4 (incl. locomotive)	97,3	24,3	18
DDM	5 (incl. locomotive)	123,7	24,7	11
ICM-III	3	80,6	26,9	87
ICM-IV	4	107,1	26,8	50
VIRM-IV	4	108,6	27,2	98
VIRM-VI	6	162,1	27,0	78
DDZ-IV	4	101,1	25,3	30
DDZ-VI	6	154,0	25,7	20
TRAXX+ICR	8 (incl. 2 locomotives)	196,2	24,5	17

Table 1: Characteristics of rolling stock

More detailed information regarding standards for maintenance of all rolling stock types at the Service Locations is given in Appendix 4.

2.5 Characteristics of the process variability

The process at the Service Locations is a continuous process. It does not result in a stable process, as every shift during every day is different, regarding the work to be done. This is caused by an ongoing deviation from the planning, due to various causes. These causes can be, amongst others, delays on the network due infrastructural causes or collisions, staff problems or problems with passengers.

Multiple causes exist to vary and to deviate from the planning. For this study, the following cases are defined. With respect to what was planned for an arriving train unit at a Service Location:

- The train does not arrive
- An extra train arrives
- The train arrives earlier than planned
- The train arrives later than planned
- The order of the train units is different than planned
- The train type arriving is not the same type as the train type planned
- The train unit has other defects than planned

The situation of unplanned process disturbances occurs at all Service Locations, but some locations suffer more from it than others. This is caused by the importance of the location within the network and the variation in the workload. A Service Location near a more important station or a station which is more sensitive for disruptions, has a less stable process.

2.6 Reality check

Having defined the exact input and boundary conditions for the service process, the basic feasibility of the topic has to be stated. By doing this, some general insight in the relevance and need for the study can be gathered. Also the desired process state is described roughly, in order to give a reference point for further analysis. After this, the process description is given, which indicates the exact method at the Service Locations.

2.6.1 Theoretical throughput

The theoretical throughput is used as a reference for the performance of the Service Locations. By investigating the basic characteristics and testing these based on the standards, the theoretical capabilities of the Service Locations become clear.

By combining the information from Table 1 and Appendix 3, the basic stabling capacity can be determined if the full fleet has to be stabled at a Service Location. This results (see Table 2) in a daily occupation of 72,5% of all available stabling tracks at Service Locations. Since the train units cannot be split, it is often not possible to use the full length of a track. For example, on a track with a capacity of 7 coaches, a train unit of 6 coaches leaves one coach length of track unused. Another remark is that, besides being at the Service Locations, trains can also be stacked on regular tracks during the night or they can be in service. Besides that, there is always some rolling stock in the Maintenance Workshops, about 6% of the total fleet, which does not have to be stacked at a Service Location. The daily track occupation is, in other words, a rough estimate.

Total rolling stock length (m)	69.027
Total available stable length (m)	95.206
Total unused stable length (m)	26.179
Occupation (%)	72,5

Table 2: Track occupation

By combining information from Table 1 and Appendix 4.1, the required characteristics for the checks can be determined. The results, which can be found in Table 3, show the total number of checks of all three types per day for all Service Locations. Besides that, the total time consumed following the standards, corrected by the 6% of the fleet which is in a Maintenance Location and therefore does not need a check. By averaging all checks over the locations, ignoring size, it can be concluded that each location has to perform 2,2 A-checks and 11,6 B-checks per day. The 24 hour-checks are more concentrated at Watergraafsmeer. The total time consumption per day for all checks together is 13.120 minutes. With an expected productivity of 6 hours per shift per check mechanic, there are 37 of those mechanics needed for the whole day over all Service Locations. This results in slightly more than one check mechanic per location on average per day, who is only working during one of the three shifts.

Totals	**Number/day**	**Time/day (min.)**
A-checks	76	3.653
B-checks	406	6.591
24 hour-checks	16	2.876

Table 3: Checks production

When considering information from Table 1 and Appendix 4.2, theoretical production of cleaning and washing can be established. The total time needed to clean the entire fleet every day is 19.255 minutes, what results in about 1,5 cleaners per location over the day, with a work shift of 6 hours per cleaner.

For washing, the results are given in Table 4. For the two types of washing, with an interval of 7 days for soap washing and 63 days for oxalic washing, the number of trains washed per day is given. This implies that every washing machine on average has to wash 8,2 trains per day and the oxalic washing machines have to wash on average 2,2 trains each on a daily base.

Totals	Number/day
Soap washing	98
Oxalic washing	11

Table 4: Washing production

The information generated above is used to reflect on, when the production is discussed later on in the study. That information is based on real values, while the numbers in this section come from standards. The calculations made above regarding required staff are theoretical, as most locations work with a continuous occupation of mechanics and cleaners.

2.6.2 Desired process state

An imaginary but desired process is established here. This is done based on the description of information need in a business information paper (Veld, 1971) and based on the principles of lean processes (Bokhoven, 2011). It is a concise description, based on reality but ignoring the boundary conditions from practice.

The key element of an ideal process is stability. Having a stable process implies that it is highly predictable, what will result in good manageability. Stability means that the production or throughput of a process is equal during a longer period. For the service process this results in a stable supply and demand over the days: all Mondays are the same as all other Mondays etcetera. This results in a supply of trains which is predictable: every Monday at point in time X, a train unit of type X arrives at track X at Service Location X. The work to be done on this train unit is also known on beforehand, as this is always the same for this particular train unit. By knowing this, the ideal desired track can be chosen, taking into account the task and the departure location.

Following this, it is clear that planning is key. In the desired situation, information is known beforehand and based on that, a planning is made. This planning is more detailed, covering all tasks to be done, including locations and parts needed and also including necessary shunting movements. By doing this, an integral planning is created, covering all tasks and not ignoring some tasks. In the ideal planning, this planning is directly followed, creating the ideal opportunities and circumstances to do all tasks properly and in time. This implies that all unplanned tasks have to be done separately, without disturbing the planned tasks. The unplanned work should be treated more flexible, by using opportunities like flex workers or separate parts of the Service Locations for those treatments. This approach should secure the good course of regular work. An important element is that the right people know the desired information in the correct appearance. If information is clear and available in time, there is less room for mistakes or missed tasks.

The main task for a Service Location is delivering trains to NSR, having the right quality. The quality, what is the sum of all other tasks, should be always leading. The efficient planning and consequence in practice should always have this main task as the central target. To make that possible, a correct delivery of information is necessary. This implicates that a strong leader has to take care of the service process, working in very close cooperation with the NSR controller or even working directly together with the NSR controller. That cooperation secures the flexibility needed within the process, resulting in the most efficient results for all parties.

2.7 Process description

The actual process within the Service Locations is determined based on location visits. The observations at the Service Locations and the interviews with staff explain the process (see Appendix 9). It is described carefully, in order to draw formulas for the process steps, which is used in the next chapters.

2.7.1 Planning

Although planning is not a study goal in itself, it is a key element in order to determine if there is any gap between what should be done and what is done during a shift in a Service Location. The process to generate a planning is given in Figure 7. The planning has the goal to make sure the matching, parking and routing of train units can be guaranteed, which is a complicated task, as described in a previous study (Lentink, 2006). For definitions of terms used, see appendix 1.

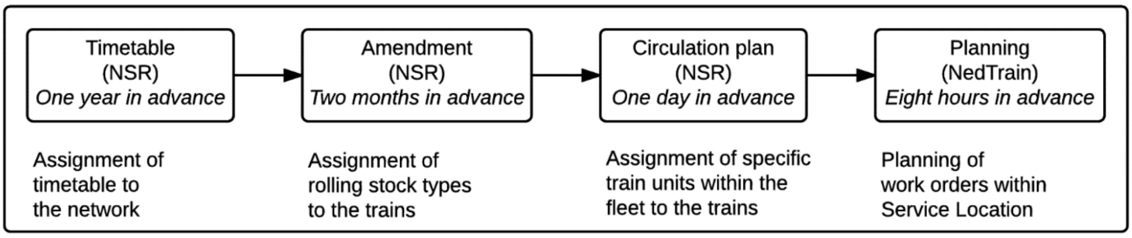


Figure 7: Rolling stock planning schedule

The workload for a Service Location is basically set a year on beforehand. At that time, the timetable for the upcoming year is made, in which NSR declares which trains will run. Also a conceptual rolling stock assignment is made, which is changed slightly every two months in an amendment. The precise assignment of specific train units to trains is made in the circulation plan, which is made for every single day of service. With this plan made, it becomes clear which train units will end their service on a day at which end stations. Based on that, they will very likely visit the Service Location coupled to that end station. All work up till this step is done by NSR. In the last step, NedTrain comes in, when planners at the various Service Locations make the planning for a specific shift. This is done based on the information in the management system, based on which the work at a Service Location is considered. Planners assign the required work on train units to the Service Location and create, based on that, an overview of the work for that location during a specific shift.

Planners of NedTrain are involved in the phase of the amendments, when long term planners already consider the work load for the various locations. That is based on the information of the NSR planners in that phase. The NedTrain planners make a rough planning of the situation during a shift, but this information in general is not used for the daily planning at the locations. This is due to the fact that the situation varies very much over the days. The planning from the long term planners is only used to determine the basic capacity requirements of a Service Location and to see whether the plan of NSR is feasible for NedTrain.

At each of all locations, there is some general input: every same day of the week, the same number of trains are coming in and have to be handled. However, the workload is not the same, as this strongly depends on the work that have to be done on every particular train unit coming from those trains. It can happen that on one day a lot of A-checks have to be performed, which are very time consuming, while the other day only some B-checks have to be done, which cost a third of the time per check compared to the A-check. This implies that a day or a shift can be very busy, while the next day or shift can be less busy. This instability in the process makes it harder to draw conclusions with respect to the performance of a Service Location. It can be even busier in the case of disruptions, when trains accumulate at a location, with a lot of unplanned and unexpected work as a result.

The leading task in the service process however is to arrange the train sets in the correct order. This logistic task is strongly depending on the infrastructure layout. It is executed in close cooperation with NSR. The shunting movements at each location are planned in a detailed way on beforehand by the work planner, in order to indicate track occupation and to determine to which extent the planning matches the realisation. All other tasks, such as the checks and the cleaning, are performed in between the logistic task. Therefore, those tasks require a different planning approach.

Planning is a major challenge for NedTrain and constantly studied. The largest difficulty for the planning of the work at the Service Locations is the difference in the approach of NedTrain versus the approach of NSR. The process at NedTrain has a stochastic character, since train units for revenue service are planned generic, while for maintenance, all train units are planned individually.

2.7.2 Work in practice

A working day is split up in three shifts of eight hours each. This results in a continuous occupation, at least at all major locations. At that location, there is always staff, which may drive to minor Service Locations if staff is needed there. For each shift, the work that has to be done is planned on beforehand, as explained in paragraph 2.7.1. An overview of all staff involved in the Service Locations is given in Table 5, including a short description of the tasks.

Name	Task
Long term planner	Makes rough capacity estimation for longer term
Manager	Manager of all Service Locations in one of the four regions
Planner	Assigns work to the shifts based on input from the management system
Team leader	Coordinating the mechanics during a shift, in direct contact with planner
Mechanic	Responsible for both checks and repairs
Check mechanic	Responsible for checks only
Rail traffic controller	Controlling the shunting movements at the NedTrain tracks
Shunting driver	NSR or NedTrain driver, dedicated for shunting movements
Cleaner	Cleaning the interior of the trains
NSR controller	Responsible for the shunt movements within a main station

Table 5: Staff directly involved in the Service Locations

The work is divided over the mechanics, with different persons being responsible for the checks and the repairs. Most mechanics are qualified to do checks and repairs, but check mechanics are trained to do only checks. The number of people working at a location during a shift depends on the time of day, which is related to the occupation of that location. During the daytime, most trains are running and thus not at a Service Location, so the number of mechanics is smaller. Consequently, the same rule about occupation applies to the cleaners.

The logistic process is controlled by a rail traffic controller, being responsible for the NedTrain service area (a non-centrally controlled area), or the ProRail train controller in the case of a centrally controlled area. This controller is also continuously available, but does not have direct influence on the shunting process. Driving the trains is the responsibility of NSR drivers. At most locations, some dedicated shunting drivers are available, who can be deployed for this type of work. They secure the planned shunting work, but are also available for unplanned work, when extra train sets arrive at or have to leave the Service Location.

The logistic process is, as mentioned before, always leading. This means that a feasible plan is always made on beforehand, with tracks to use at certain moments. This plan does not take into account the maintenance process: checks and repairs are planned real time when they fit within the logistic process. Mechanics receive an overview of what to do during a shift and they plan it themselves. When the aerial platform or trench is needed, the trains are driven to that location only when possible. As mentioned, the work load of mechanics really depends on the day, as it may be very busy during disruptions or extreme weather conditions.

2.7.3 Major process steps

In order to define different steps and functions within each process part, the five major process steps are described carefully and translated into formula form. The information is visualised in flow graphs, made by the author of this study, that represent a single work order. All data given in this section is based on the author's experience and supported by numerous interviews with people working at the Service Locations (see appendix 9). The time indication used in the figures is a generalised estimate, based on experience at the larger Service Locations. It is intended to give insight in everyday practice.

The parameters used in the formulas are defined as follows:

- $t = time\ in\ minutes$
- $x = train\ type$
- $n = number$

A-checks, B-checks, 24 hour-checks

A typical process flow of one of the checks is visualised in Figure 8 and described afterwards.

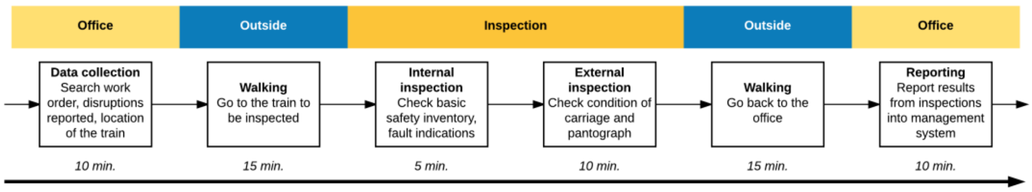


Figure 8: Typical B-check process flow

The basic check is a B-check. The major focus is safety, while comfort is not considered in this check. The mechanic starts at the office, where an iPad is used to collect relevant data and to consult relevant manuals. Depending on the location, walking time can be up to fifteen minutes to reach the train. At the train, functionality is not tested, only the diagnose screen in the driving cabin is consulted in order to check whether safety related irregularities are reported. From there on, the work is done outside the train. The carriage of the train is inspected, including the bogies and wheel sets, to establish whether everything is present and in good condition. Also the pantograph is considered, it has to be in good condition. After this, the B-check is complete and the mechanic can go back to the office or continue checks on other trains. At the office, the results of the check have to be logged in the management system and failures have to be reported, what may result in work orders for the repair mechanics.

The A-check and 24 hour-check also contain the safety test described for the B-check. This check is expanded with a functional test of the train and an inspection of the interior of the train. The functional test focusses on the brakes, the lights and the outside doors of the train. The interior check is meant to inspect the condition of the seats and folding tables, the infotainment system and the inner doors. If anything is missing or needs to be repaired, it is reported directly by the mechanic, while the train will pass its A-check. By failing the functional test, the train is not released for service, but will be repaired directly if possible. Otherwise, it is sent to a Maintenance Location for repair. After the inspection, the same holds as described for the B-check: the results have to be reported in the management system and the results of the inspection are reported to the planner, who makes work orders to repair what is needed.

The production time in formula form can be stated as follows:

$$t_{check,A} = t_{WO} + 2 * t_{walk} + t_{internal,x} * n_{coaches} + t_{test} + t_{external,x} * n_{coaches} + t_{report}$$

$$t_{check,B} = t_{WO} + 2 * t_{walk} + t_{internal,x} + t_{external,x} * n_{coaches} + t_{report}$$

$$t_{check,24h} = t_{WO} + 2 * t_{walk} + t_{internal,x} * n_{coaches} + t_{test} + t_{external,x} * n_{coaches} + t_{report}$$

Repairs
A typical process flow chart of a repair is given in Figure 9 and described afterwards.

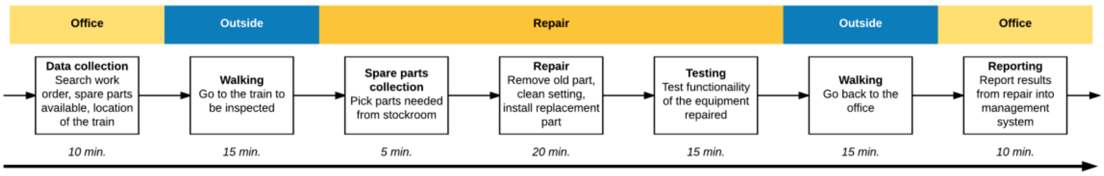


Figure 9: Typical repair process flow

For this study, four types of repair are defined, which all have different causes and priority. These repair types are, defined by their abbreviation:

- Corrective order (CO): crucial item is defect, has to be repaired before release for service
- Preventive order (PO): mainly filling water for toilet, horn check or automatic coupler cleaning
- Direct repair (DHST): small repairs done directly when mechanic ascertains
- Work from inspection (WUI): repairs following from A- or B-check

The planner gives every work shift an overview of the work of these different types. Every mechanic available receives a list of work that is assigned to him. With this list, he starts looking at the work instruction for the task assigned. Based on this, he considers spare parts needed for the repair. If they are available at the Service Location, the task can be continued, otherwise the part is ordered, which implies that the task is cancelled. Collection of the parts takes a while, as the stockroom is not very centrally located at most Service Locations.

The repair starts when the mechanic has reached the train. This walk can be as long as fifteen minutes, as most locations are quite extensive. A common repair starts with disassembling of the broken part and then assembling of the replacement part. After that, a functional test is carried out, to confirm that the repair has been successful. If the repair is successful, the mechanic goes back to the office to close the work order and to finally buy out the spare parts from the stock. From this point, the specific work order is closed and the mechanic has finished the task.

There exists a large variance within the range of the repair work orders. Some are relatively simple replacements, for example when a lamp has to be replaced, but some orders are more complex. It may even be necessary to move the train to a track with special equipment to work under or on the roof of the train. These shunt actions are executed by NSR drivers and not by NedTrain staff.

The production time in formula form can be stated as follows:

$$t_{repair} \quad = \quad t_{WO} + 2 * t_{walk} + t_{spareparts} + t_{repair} + t_{test} + t_{report}$$

Washing
A typical process flow of washing a train is given in Figure 10 and described afterwards.

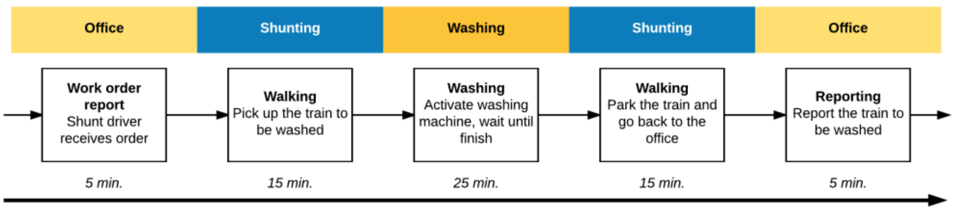


Figure 10: Typical washing process flow

The commission for washing a train comes from the NedTrain planning, but the work is actually done by NSR shunting drivers. The trains that need to be washed are listed and become part of the shunt plan, made by the local NSR controller. NSR drivers pick up the trains and drive them into or through the washing machine. In some cases, a NedTrain employee is responsible for controlling the washing machine, but most washing machines work automatically.

The important observation is that shunting trains within a location takes capacity, resulting in track occupation. Another important fact is that no other tasks on the train can be executed parallel while the train is being shunted and washed. The time needed to wash a train varies heavily, depending on the washing machine and the train type, but at most locations, only several trains are washed every day. When the washing is done, the train is shunted back towards a stabling track, from where other tasks can be started.

The production time in formula form can be stated as follows:

$$t_{wash} = t_{WO} + 2 * t_{walk} + t_{washing} + t_{report}$$

Cleaning

The cleaning process is a straightforward process: all trains are cleaned after each other, following the same approach. A basic cleaning is done, including emptying the trash, the toilets are cleaned and missing paper or soap is filled. When needed, cleaning the floor or seats is also included in the cleaning process.

Special teams exist for graffiti cleaning, internally as well as externally. During the graffiti cleaning, no other tasks can be done on the train, while regular cleaning can be done parallel to other tasks, except for washing.

Shunting

The shunting process is executed in direct cooperation between NSR and NedTrain. The NSR train driver receives an order via its own local controller. The commission to do this may come from various sources, for example from the NedTrain planner, when it is needed to reposition a train to make it attainable for the mechanics.

When the driver receives the order to move a train, he first has to walk to the train and make it ready to drive. After that, he has to call the rail traffic controller (NedTrain or ProRail) and wait for permission to carry out the desired movement. When he has the permission, the actual movement is done and afterwards, he has to declare the movement completed to the rail traffic controller. A single shunting movement can take up to 15 minutes, of which a lot is waiting. The actual movement takes one or two minutes, the rest is waiting. Coupling or decoupling trains also prolongs the time needed.

2.8 Conclusions qualitative analysis

In this chapter, the Service Locations are described based on location visits and literature review. This is done based on the three main pillars of the study: the location characteristics, rolling stock characteristics and the process variability.

There are five main tasks at a Service Location: checks, repairs, cleaning, shunting and washing. Washing can only be done at 12 out of 33 locations, since a washing machine is needed. The checks are important, since train units may not run if they are not checked in time. Shunting is also important, since it is directly related to the train service.

Two layout types exist: the carousel and the shuffleboard. The shuffleboard is the more static alternative: train units do not have to move within the location once they are arrived. The carousel type is more dynamic: train units move within the location to complete tasks. Also other infrastructural characteristics describe a Service Location, such as the available equipment. Some locations without a distinct layout type are called a station.

The fleet exists of 11 train types of in total 18 different compositions. There is a variety in the size of train series and also in the length of a train unit.

The variety in the process is caused by a inequality between planning and work in reality. Since many factors can influence the train service over the day, also the input of a Service Location varies. Examples of possible consequences are extra train units arriving at a Service Location or other train units than planned. A most common cause for this are disruptions in the train service.

Also the expected realised work is described, based on standards. This gives insight in what the extent of the problem is. In contrast to this, the work in practice is described, based on visits at various Service Locations. The role of the various persons involved in the process is described. Also detailed process descriptions are given for the main tasks within the process.

Summarised

- In total 33 Service Locations across the country
- 12 Service Locations have a washing machine
- Main tasks: checks, repairs, cleaning, washing, shunting
- Layout types: shuffleboard, carrousel, station
- Trains: 11 types in 18 compositions
- Every day is different due to process variance
- Tasks are decomposed into parts to describe them

3 Data preparation

In order to describe the Service Locations in a quantitative way, data is needed. In this chapter, it is stated what the current performance of the Service Locations is, based on the data provided and generated. Performance is the throughput of a location: the realised or closed work orders of that location. This is done for the overall performance, with a specific focus towards the three main pillars of the study: the rolling stock, the infrastructure and the process variance.

The data used in this chapter is gathered from the NedTrain management system, called Maximo (NS, OBIEE, 2016). For this study, the data from a full year is taken and prepared for analysis. This includes all closed work orders (WO) from August 2015 until July 2016, for all 33 Service Locations currently open. In total, there are over 368.000 work orders considered in the study. Only work orders on train types mentioned in chapter 2.4 are included in the data set, so only domestic trains. The data set was generated in August 2016, so the chosen time frame simply is the most actual year available at the time. It is a reliable data source with a significant volume of data available. A limitation is that the information is partly generated automatically, while the basic input is filled in by humans. That will mainly have an influence on realised times, since people responsible for the order closing can wait, with some badging as a result. Badging is when multiple work orders are collected and closed together. Within NS, a more reliable data source does not exist, so using Maximo data is the best solution to gather insight in Service Location production.

Data about the cleaning process is not available, since this task is outsourced to a third company. Because of that, the characteristics of cleaning are not described here. As assumed in chapter 2, cleaning is not the critical task. Therefore, and because of the lack of data for this process, cleaning is not in the scope of this study. This chapter is mainly build up based on the main pillars of the study: the infrastructure, rolling stock and the process variance. Those factors are further defined in this chapter.

3.1 General performance

This section describes the general performance of all Service Locations. This is done by drawing the production of each task for all locations. By doing this, a good insight in the scale and relevance of the topic can be gathered. In addition, a distinction within the Service Locations can be made.

In order to describe the general performance (and most other indicators in this study) the number of work orders is considered. One work order represents a single train unit. It is chosen to do so and not to describe the total time consumption of all work orders. On one hand, this is caused by the fact that not all data lines are coupled to the corresponding realised labour time. Besides that, for the general purpose of this study, detail level is sufficient by taking average values for the times consumption of the tasks within the process. By taking the number of work orders, it is made sure that all work orders are represented. This train unit and the task of that work order is directly linked to the infrastructure occupation, which is one of the main points of interest of this study. All other properties can be coupled the train type from the work order.

3.1.1 WO distribution

The total number of work orders for each work order type for all locations for the full year is given in Figure 11. In this graph, all seven work type categories that a work order can represent are given. Those work types are used throughout the study. The count on the vertical axis represents the total realised work orders of that type for all Service Locations for the full year considered.

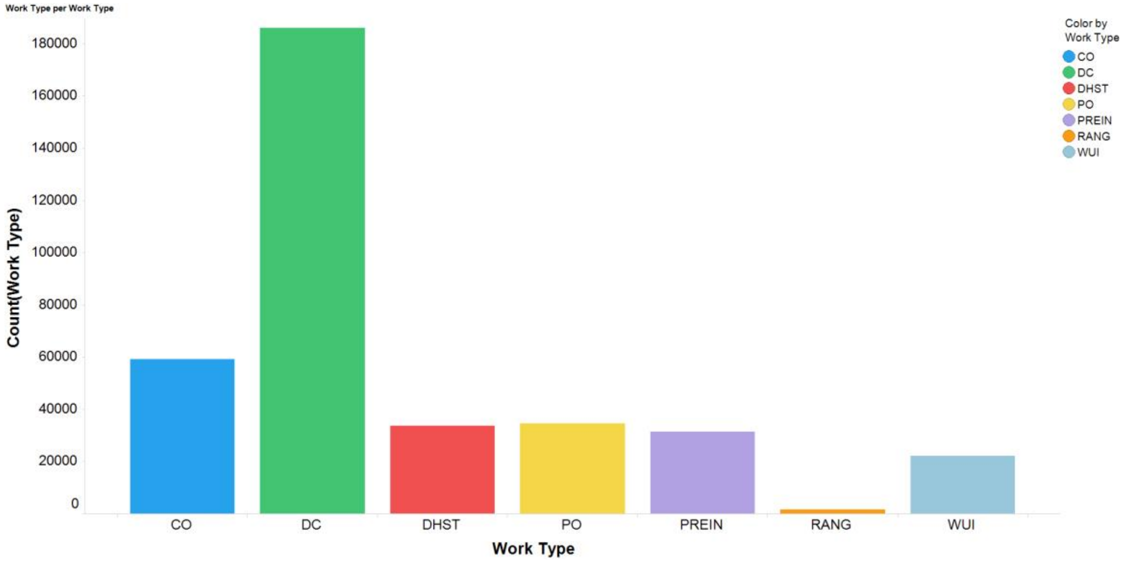


Figure 11: Number of work orders per WO type

Figure 11 shows that almost half of all work orders is a daily check (DC). The second most common work order type is the repair (CO, DHST, PO and WUI, see paragraph 2.7.2), followed by washing (PREIN) and shunting (RANG). As mentioned, cleaning is out of scope. The shunting movements given here are not all shunting movements of all Service Locations, only the movements requested by NedTrain are shown for some locations. The logistic process is not logged directly and therefore, the data is not representative for this task. Nevertheless, it gives insight in the processes with respect to shunting as some routing is included in the shunting work orders.

The dominant role of the daily checks amongst all work orders implies the importance of this task. Besides that, as there is a lot of data available for all locations, and the check task is the same in all cases, these tasks are very interesting for further study. The repair related work orders have a much broader variety and therefore it is harder to show consistent statistics about them. That is only possible if all work order descriptions for that tasks are generalised.

3.1.2 WO distribution per location

In Figure 12, the number of closed work orders for each of the 33 Service Locations is given. Within this distribution, the number of work orders of each type is also shown.

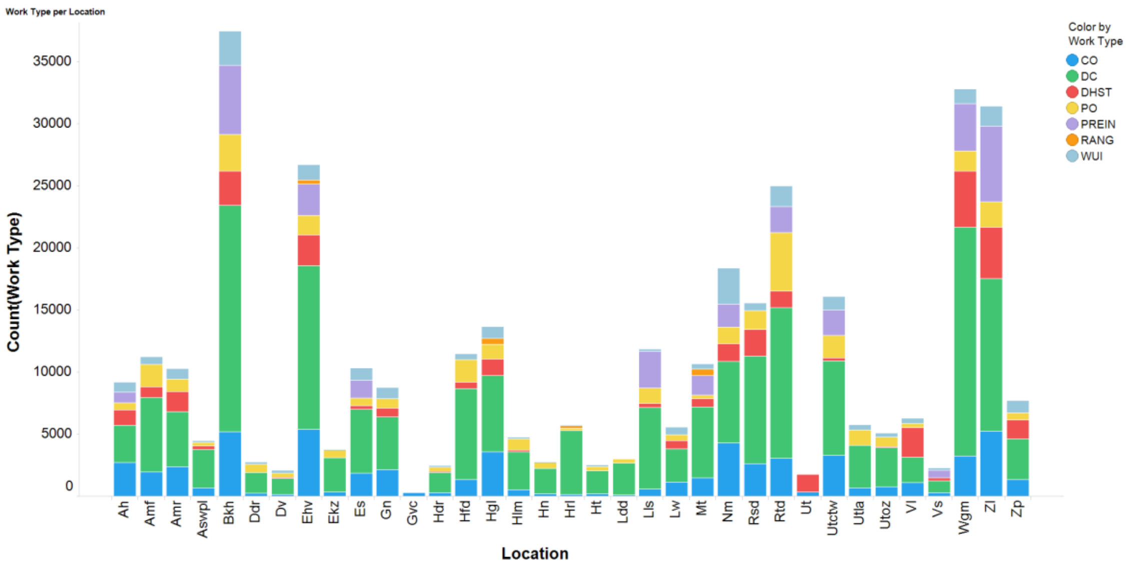


Figure 12: Number of work orders per Service Location with WO types

According to Figure 12, there is some distinction in the scale of the Service Locations. There are five big Service Locations with more than 20.000 work orders annually, eight with a work order production between 10.000 and 20.000 and the majority, 22 small Service Locations, have less than 10.000 work orders per year. The five big Service Locations are Den Haag Binckhorst, Eindhoven, Rotterdam, Amsterdam Watergraafsmeer and Zwolle. The biggest location of all, Binckhorst, is responsible for more than 10% of the total work orders in a year, with over 37.000 closed work orders. The data also include the work orders from the Technical Centres. Only in the case of Hengelo, this results in somewhat distorted results, as the main part of the DM'90 maintenance is done there. However, the Technical Centres are part of the Service Locations and therefore those work orders are processed here.

Regarding the work order type distribution, it can be noticed that for all locations the checks have the highest share in importance. However, some locations perform relatively many repairs, while other locations do almost exclusively checks. The smaller the total throughput of a location, the higher the share of checks. The location Utrecht OZ was temporarily closed since June 2016, resulting in a production difference of about 1000 work orders based on the throughput of the previous 10 months. This influence on the total image is very limited and is therefore ignored in the rest of the study. In Figure 12 also the locations Utrecht Centraal (Ut) and Den Haag Centraal (Gvc) are given, since they appear in the data. These are not proper Service Locations, only in case of disruptions mechanics from respectively Utrecht Cartesiusweg or Binckhorst come to those stations to handle the problems. Those two locations do not appear in the data further on, the related work orders are only included if the total number of work orders is considered.

3.1.3 WO distribution per month

Figure 13 shows the number of closed work orders for all Service Locations together for every month within the year considered. The different work types are given as well.

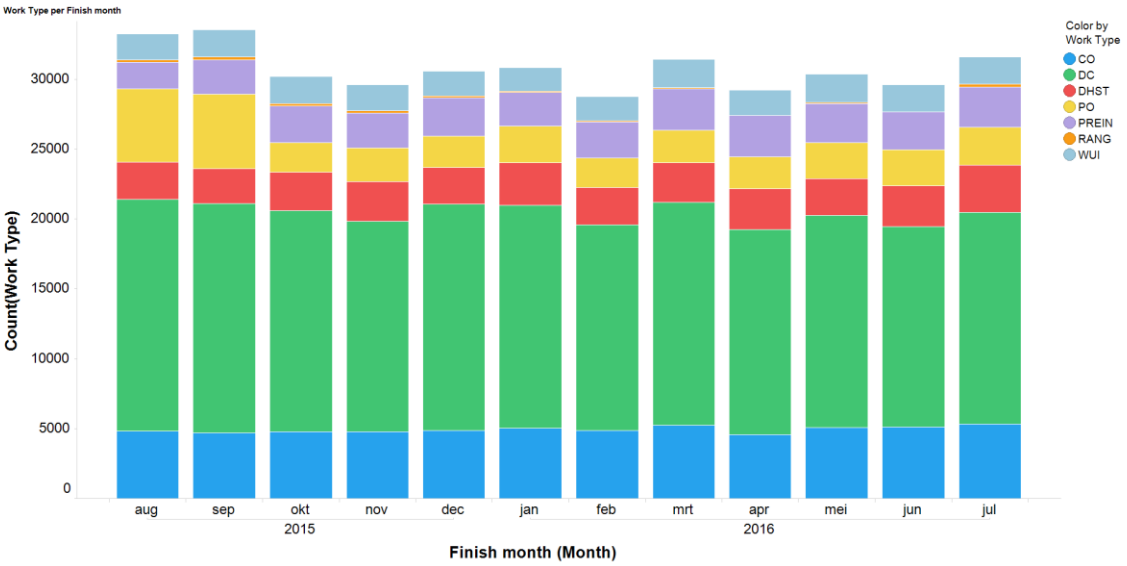


Figure 13: Number of work orders per month with WO types

Figure 13 shows that, within a range of approximately 10%, the number of work orders per month does not vary significantly. All months with 29 or 30 days have slightly less work orders, that explains the variation. The only outliers are August and September, which have the most work orders amongst all months considered in the study. Also the distribution of work types is quite constant, with a dominant role for the checks (DC). Remarkable is the share of preventive orders (PO) in August and September, which can be explained by a specific task at the time to check the condition of the axles of the SLT train sets, resulting in over 3000 work orders per month. By ignoring the influence of this tasks, those two months can be compared perfectly to the other months regarding the number of work orders and the distribution of work order types.

3.1.4 WO distribution per shift

In Figure 14 the total work orders for all Service Locations are given per shift for each month. The early shift is from 7 a.m. to 3 p.m., the late shift from 3 p.m. to 11 p.m. and the night shift from 11 p.m. to 7 a.m..

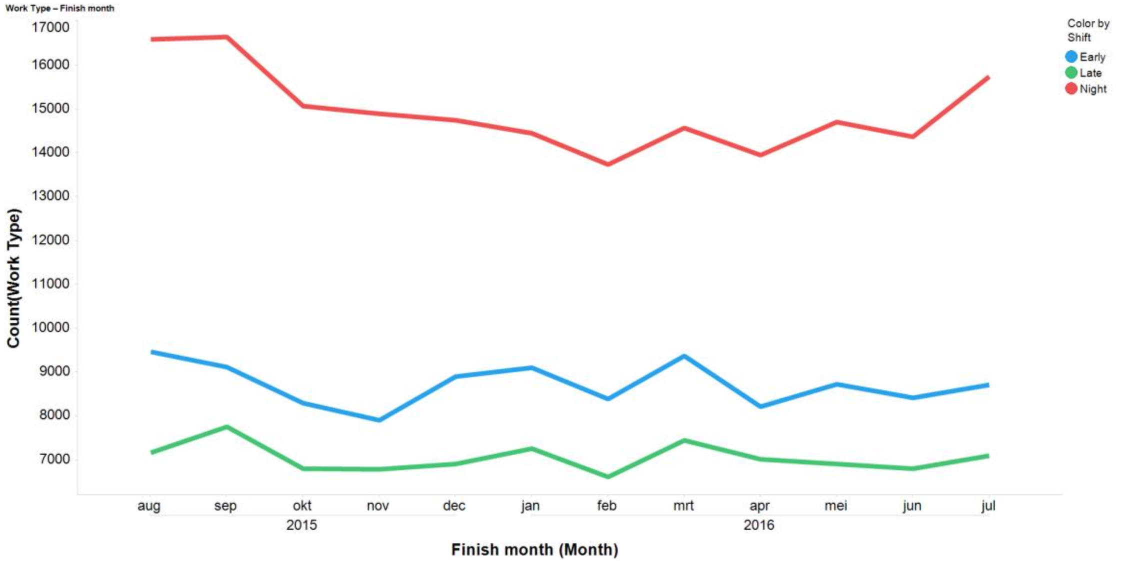


Figure 14: Number of work orders per month for each shift

Figure 14 indicates the importance of each of the three shifts. It is clear that the night shift is the leading shift, as this shift covers about the same number of work orders than the early and late shift together. The early shift has slightly more closed work orders than the late shift. This can be explained by the fact that the afternoon peak hours last longer and thus most of the train units are running during those hours. Therefore, the number of train units that are available for maintenance at the Service Locations is limited. The morning peak lasts shorter, hence the time in revenue service is shorter and the occupation at the Service Locations is higher.

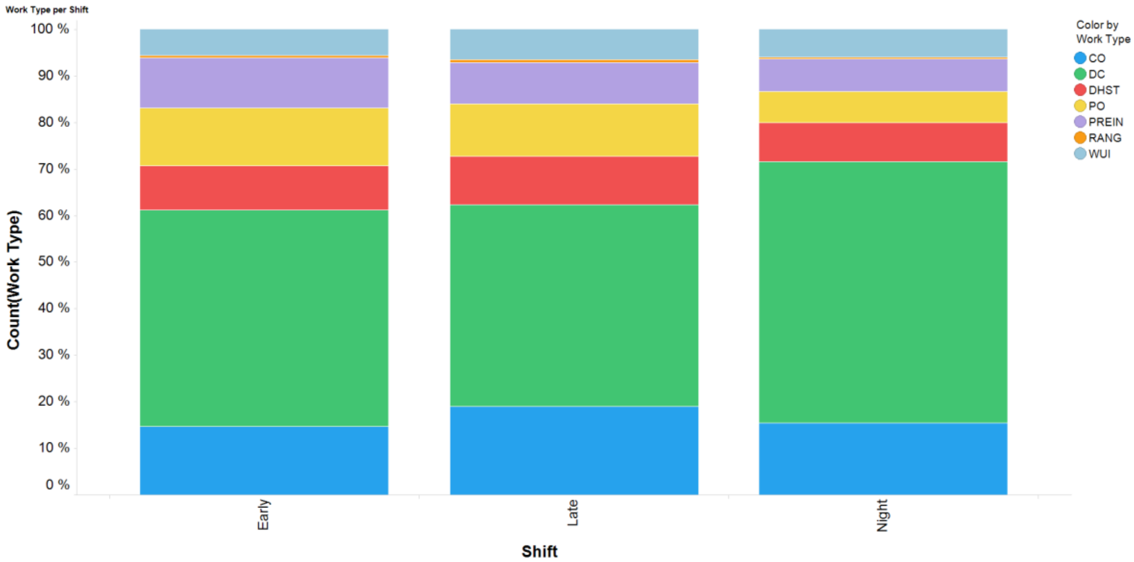


Figure 15: Work order type distribution per shift

Figure 15 gives insight in the distribution of the work order types for each shift at all Service Locations. There is no substantial proportional difference between the shifts regarding the work orders. Proportionally there are more checks during the night shifts and more repairs of various types during the day shifts. In absolute numbers, the night shifts have more work orders in all cases (see Figure 14).

The night shift is by far the busiest shift in terms of numbers of closed work orders. From all work orders, 49% is done during the night shift, 28% during the early shift and 23% during the late shift. Therefore the night shift is chosen as the normative shift amongst all shifts. For the rest of the study, the capacity calculations of the Service Locations are based on the capacity of the night shift.

3.2 Performance per Service Location

In this section, the focus is on the individual Service Locations, to see whether production varies per location. The exact number of work orders per location is already distinguished in chapter 3.1.2, while in this section some production characteristics are determined.

3.2.1 Rolling stock assignment per Service Location

In Figure 16, the ratio of each train type is given for the closed work orders of all locations. For each location, the percentage of all closed work orders for train units of a distinguished type is given. By doing this, the distribution of the train types at the Service Locations can be indicated and thus the variation in the work load of a location.

Bigger train unit series appear on more locations and have on average the highest throughput. The distribution of the various train types is directly influenced by the rolling stock assignment for the various train lines of NSR throughout the country. As a result of this, there are Service Locations where only a limited variety in rolling stock types appears. The majority of the Service Locations have more than four train types in its portfolio. Vlissingen is an extreme example, with almost only VIRM, except for a dozen of work orders.

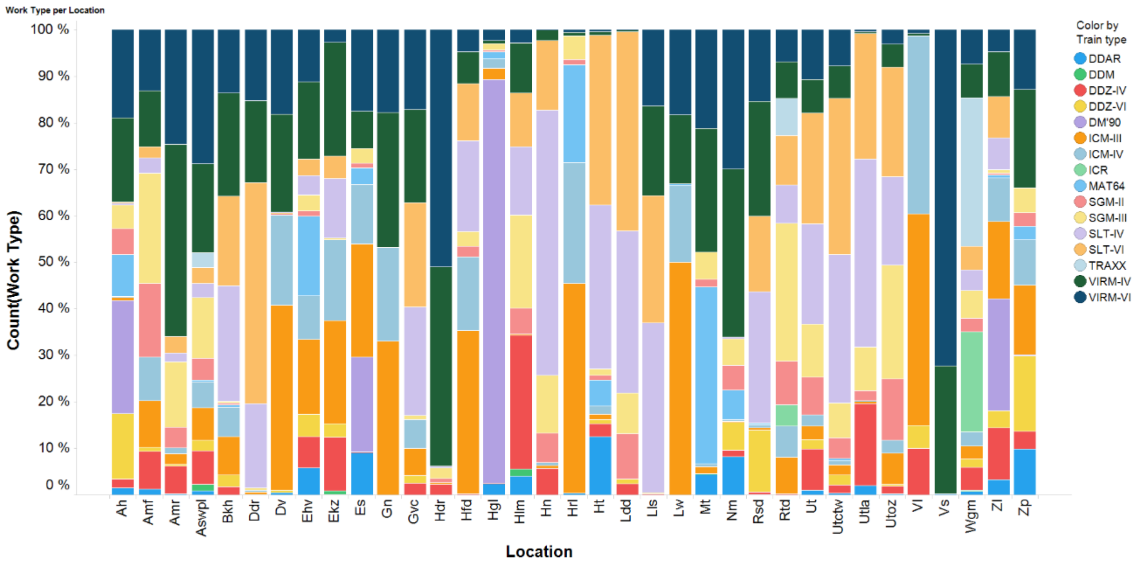


Figure 16: Train type ratio per Service Location

As a general rule, when Figure 16 is linked to Figure 12, it can be stated that the bigger Service Locations have a broader variety in train types. This does not imply that small Service Locations always have a limited variety, like at Zutphen or 's Hertogenbosch for example.

3.2.2 Average check time per Service Location

In Figure 17 and Figure 18 the average reported time and the standard deviation of an A- and B-check are shown. This is done for each Service Location. For this example, the checks of the VIRM-IV train units are chosen, since this type appears at all locations considered. The VIRM-IV is the biggest series within the NSR fleet and is widely spread across the country, which results in much data available.

For this study, it is chosen to give the average value instead of the most common value. By taking the average instead of the mode, the influence of all realised work orders is taken into account. Since a large data set is used, the influence of proper outliers is very limited, but all work orders do have influence. By taking the mode, a less representative insight would have been created, as all outliers would not be taken into account.

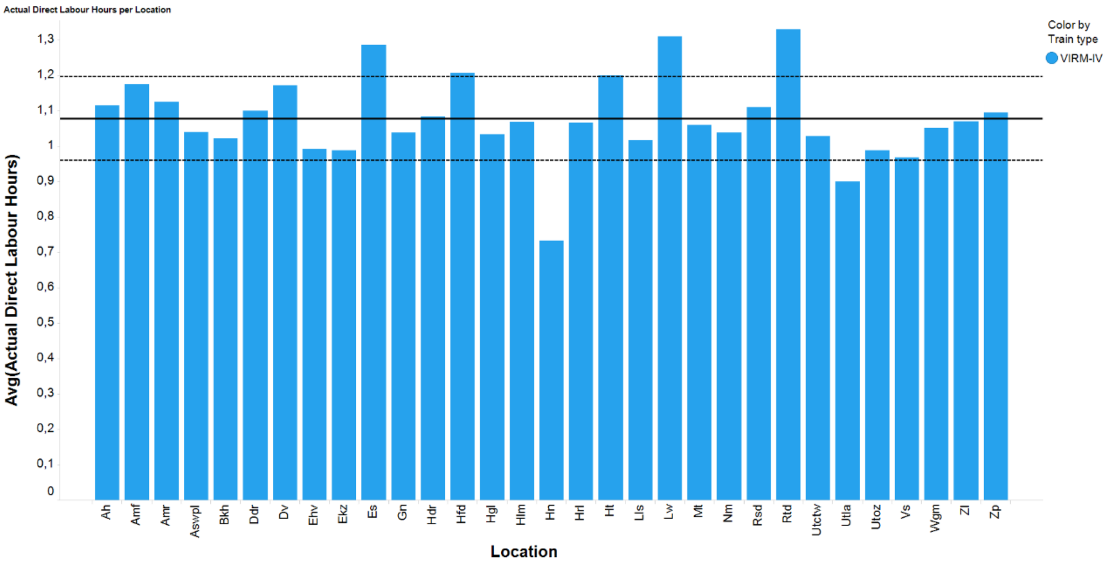


Figure 17: Average time reported per Service Location for A-check VIRM-IV

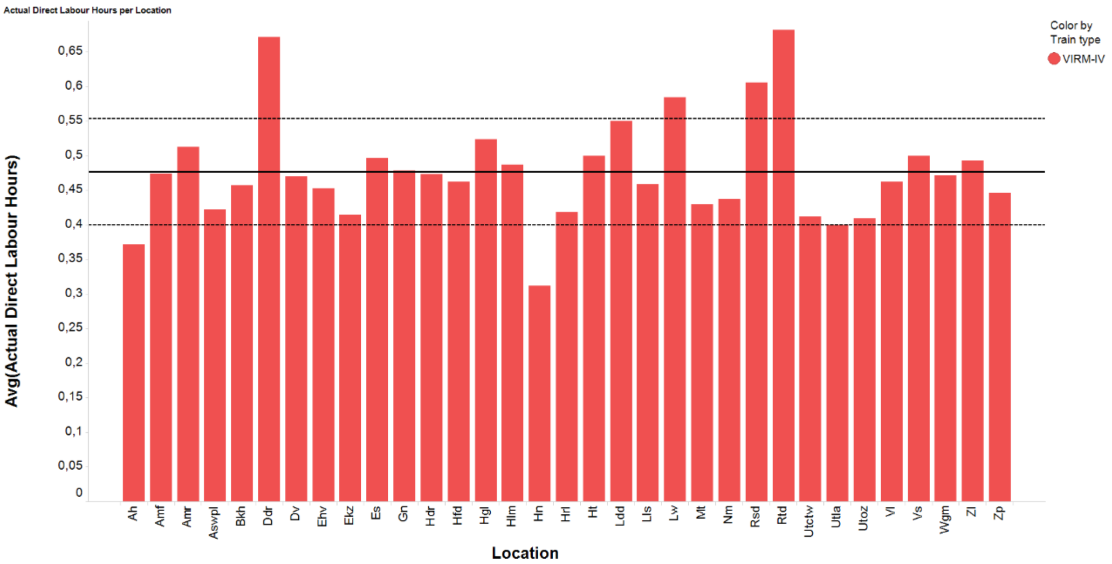


Figure 18: Average time reported per Service Location for B-check VIRM-IV

Figure 17 and Figure 18 show that there is quite some variation in the time consumed for the same task among the various locations. Also the average and the standard deviation for all checks are given in the figures. For the A-check, there is a range between 0,70 hours and 1,30 hours. The average value for the A-check is 1,08 hours with a standard deviation of 0,12 hours. For the B-check, the range is 0,30 to 0,65 hours, with an average value of 0,48 hours and a standard deviation of 0,08. This means that this specific B-check takes on average more than twice as long in Rotterdam as in Hoorn for example. In general, there is spread in the realised work order times, but the standard deviation is not that big, so the values for the average time are representative for the average performance of all locations for those tasks.

3.2.3 Washing per Service Location

In Figure 19 the washing production for all Service Locations with a washing machine is given. The production is the number of train units washed, as every work order represents one train unit washed for a specific washing task. The numbers represent the production for the full year considered.

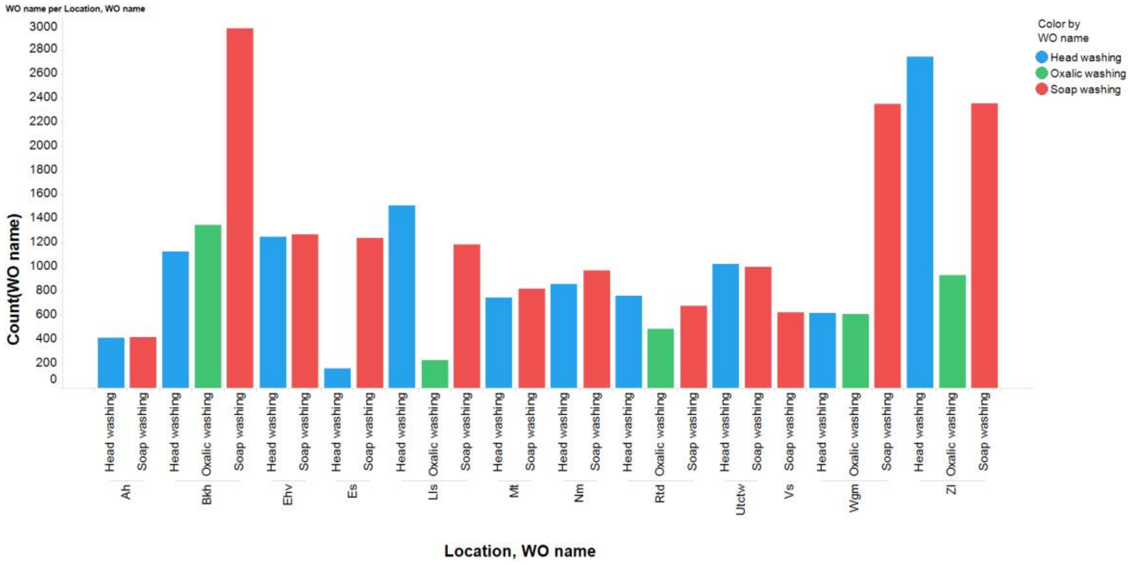


Figure 19: Number of work orders per washing type for all Service Locations

As can be seen in Figure 19, not all washing machines have the same capabilities. There are only five washing machines that can do all three washing tasks, including the oxalic washing. The washing machine in Vlissingen can only do soap washings and nothing else. The washing machines at Binckhorst, Lelystad, Rotterdam, Watergraafsmeer and Zwolle can do all three washing tasks. The total number of oxalic washings is limited, which is in line with the planning, as a train unit has to be washed with oxalic only once per 63 days. When in the case of soap washing the planning is compared to reality, the results are remarkable: with in total 689 train sets in service, with 52 soap washings per year per train unit according to the standards, a total of almost 36.000 soap washings is expected. The soap washings of the twelve locations summed show a total production of only 16.000 soap washings. This is on average equal to slightly less than 44 soap washings per day, with a standard of about 98 washings per day (see Table 4), so only 45% of the planned washings was done.

In some washing machines the program automatically includes a head washing, for example in Utrecht Cartesiusweg or Eindhoven. That shows an almost equal production level of head wash and soap wash. Other washing machines do not do such a head washing automatically, resulting in evident differences in production. In general, the differences in production are remarkable: the washing machines in Arnhem and Vlissingen are responsible for only 10% of the work of the machines at Binckhorst and in Zwolle. Since the washing machines are all located at Service Locations with sufficient rolling stock presence, there are other causes for the difference in production. An average washing machine does 1325 soap washings, 935 head washings and 720 oxalic washings a year.

3.3 Performance per rolling stock type

The performance of tasks may vary for different rolling stock types, since they have different characteristics. Besides length or age, also technical specifications or reliability influences the performance. Insight in this is gathered in this paragraph, related to the time consumption involved with the process, but also the appearance interval of the trains at the various locations.

3.3.1 WO distribution per train type

In Figure 20, the number of work orders for all train types considered in the study is given. Trains of the same type with different length are shown together, with a distinction between the various length types within the columns, since there are a lot of similarities between them.

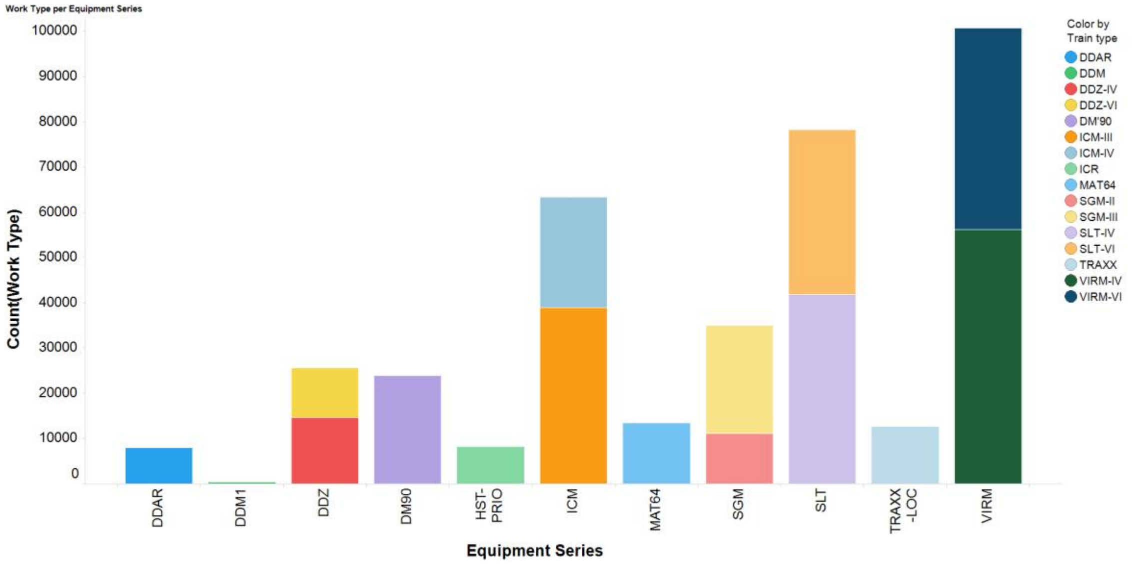


Figure 20: Number of work orders per train type

There is a direct correlation between the number of train units within a type and the number of work orders for a train type. The highest volume of work orders is generated by VIRM-IV, which is 15% of the total number of work orders. The VIRM is the biggest series of NSR's fleet, with also the highest number of work orders, followed by SLT and ICM. DDM is taken back into service since June, with only eleven train units, resulting in a limited number of work orders. Two other small series, MAT64 and DM90, produce relatively many work orders, as they have special regulations for their maintenance, with smaller check intervals. MAT64 is phased out since April 2016 and therefore further study to its characteristics is not relevant anymore.

	Total WO	WO/train unit	WO/coach
DDAR	7.896	439	108
DDM	267	24	5
DDZ-IV	14.268	476	119
DDZ-VI	10.841	542	90
DM'90	23.790	1.081	541
ICM-III	38.986	448	149
ICM-IV	24.289	486	122
ICR	8.151	479	80
MAT64	13.292	359	180
SGM-II	11.091	370	185
SGM-III	23.846	397	132
SLT-IV	41.813	606	152
SLT-VI	36.353	586	98
TRAXX	12.601	371	371
VIRM-IV	56.145	573	143
VIRM-VI	44.438	570	95

Table 6: Number of work orders per train unit per year and per coach per year

The average number of work orders per train unit and per coach for the full year is given in Table 6. The value for DDM is not reliable, since it is only in revenue service since June 2016, so for one month. Also the high value for DM'90 is not reliable, since almost all maintenance for this series is done at Service Location Hengelo. From the larger series, SGM and ICM perform best, with a low number of work orders per train unit. VIRM and SLT perform worse, with a high number of work orders per train unit. For SLT, this can be explained by the fact that the B-check for this type is done every day. Regarding the number of work orders per coach, the conclusions are different. For all series, the train unit with more coaches performs better. However, this is not a reliable indicator for all work order types. The checks are done per train unit, so a check is divided over more coaches in case of longer train units, resulting in less work orders. Repairs are done for individual cases. Therefore, the number of work orders per train unit is not the most reliable indicator of the performance of a train type. A comparison of train types with the same number of coaches is possible. The DDAR performs best for trains with four coaches, while for six coaches, it is DDZ. This is a different result as when work orders per train unit are considered.

Figure 21 shows the distribution of the work order types for each rolling stock type. The work order distribution is given as a percentage of the total number of work orders (from Table 6) done for that train type.

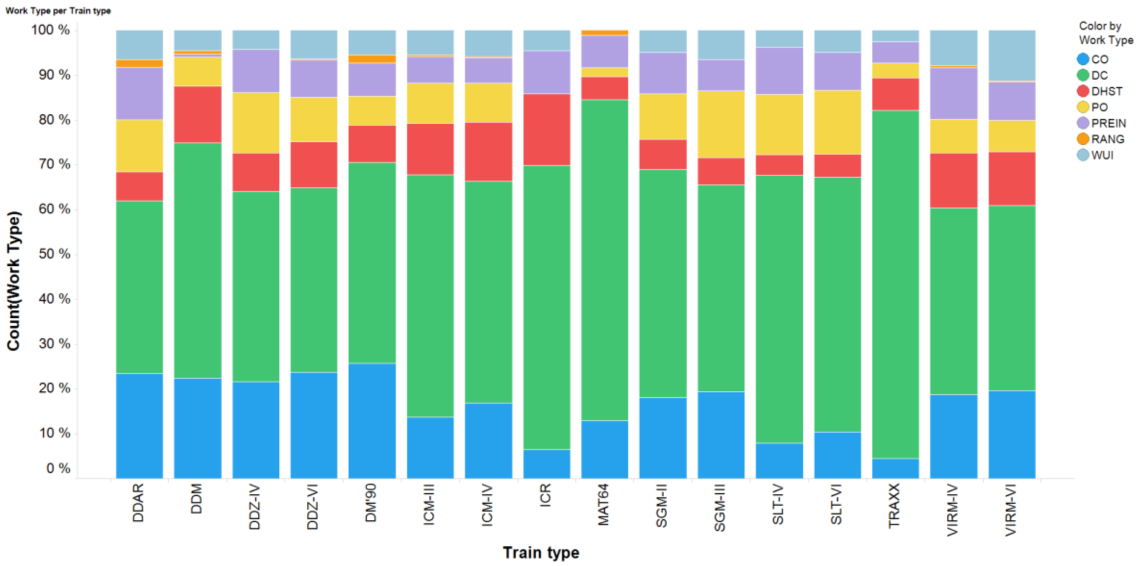


Figure 21: Work order type distribution per rolling stock type

This work order distribution shows several interesting phenomena. As pointed out earlier, the daily checks (DC) are the dominant tasks for each rolling stock type. However, the share of the total work orders for DDAR, DDZ and VIRM is only 40%, which is below average (50%). The share of repairs for those rolling stock types is higher. In contrary to MAT64 and TRAXX, which have a share of more than 70% for the DCs amongst all work orders. This is partly caused by the fact that the more advanced maintenance tasks for the TRAXX-locomotives are not done by NedTrain. That is done by Bombardier, the manufacturer of the locomotives. A low number of corrective orders (CO) is generated by SLT and ICM as well as the IC Direct train compositions, TRAXX+ICR. The relatively highest share of direct repairs (DHST) are generated by DDM and ICR, which is an indication for unreliability. For VIRM, the share of work from inspection (WUI) work orders is notable amongst other types. The share of washing work orders (PREIN) is at the same relative level for all rolling stock types.

3.3.2 Average check time per rolling stock type

In Figure 22 and Figure 23, the average time reported for an A- and B-check per rolling stock type is shown. The time consumption shown is the average over all Service Locations for that train type.

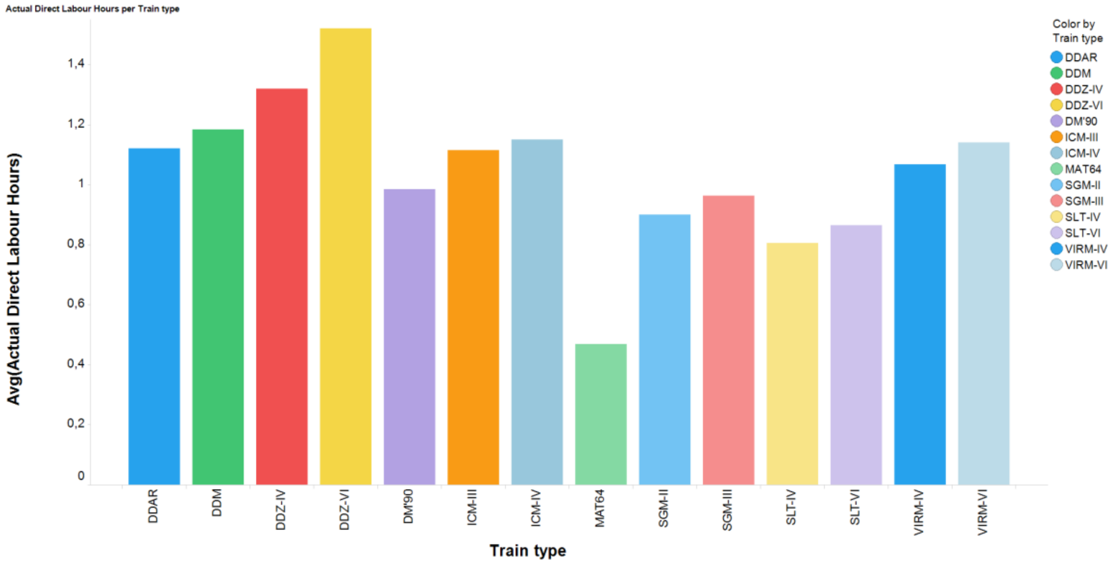


Figure 22: Average time reported per rolling stock type for A-check

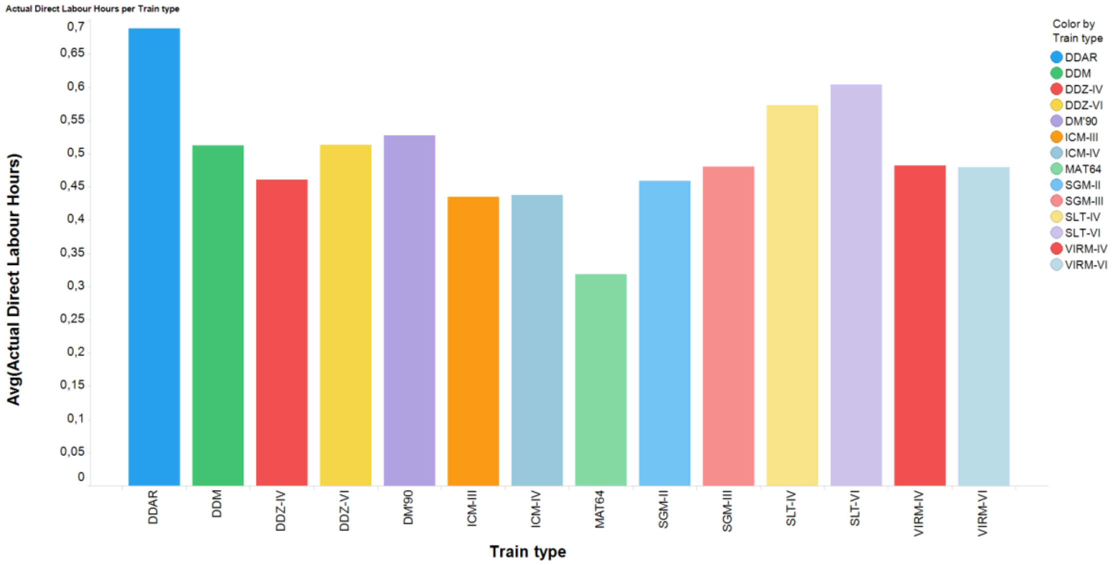


Figure 23: Average time reported per rolling stock type for B-check

Figure 22 and Figure 23 show that there is a variance in the check time for each rolling stock type. Although, most of the checks are within the same time frame: for an A-check between 0,8 and 1,2 hours, for a B-check the range is 0,4 to 0,6 hours. Especially regarding the A-checks, the time consumption for a check for longer trains is different from shorter trains of the same type. Compare for this DDZ-IV to DDZ-VI or SGM-II to SGM-III. Looking at the B-checks, there is less variance, especially not for the ICM- and VIRM-train sets with various length. As the checks partly exist of an inspection of the interior and the exterior of a train, it can be concluded that the walking time to do this is not the most time consuming part of the checks.

Remarkable are the peaks: the high time consumption for the A-checks of DDZ compared to the long duration of the DDAR B-checks. On average, all checks on MAT64 are the shortest. This can be caused by the fact that this train type does not have any electronical features, but only consist of mechanical parts. The more average time consumption is generated by the ICM- and VIRM-train units, that are widely common throughout the whole railway network. Remarkable is the fact that a B-check for a SLT-train unit takes almost the same amount of time as an A-check. This results in a smaller than average A-check time and an above average B-check time.

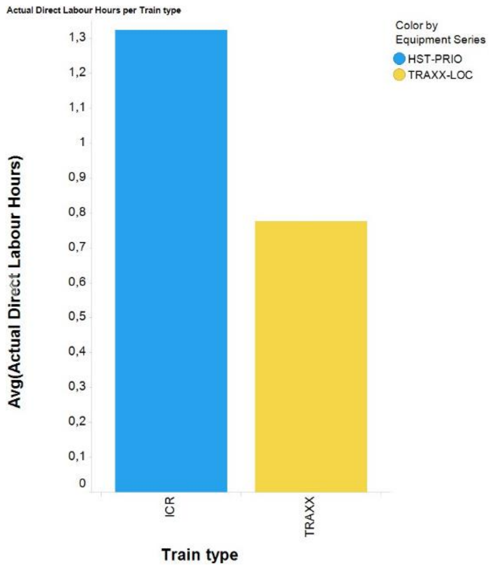


Figure 24: Average time reported per rolling stock type for 24 hour-check

In Figure 24, the average time consumed for a 24-hour check for the TRAXX locomotives and the ICR coach sets is shown. As mentioned in chapter 2.2.1, this check is in line with an A-check and also the time consumed is in the same time frame as the A-checks.

3.4 Performance in planning variation

The workload in types and numbers of rolling stock per shift at a Service Location varies heavily. As a result of that, the process is not stable, as it is almost never according to plan and never the same as any other day. The variance in what is planned and actually arrives at a Service Location, is high. This aspect of the process is described in this paragraph.

3.4.1 Overall variance

Figure 25 shows the number of train units worked on at any Service Location during all Mondays of the year considered, regarding closed work orders. The horizontal axis shows the finish date, the vertical axis the number of train units present at that Monday. The line connecting the data points is only given to gather insight in the pattern between the data points, it does not indicate anything itself.

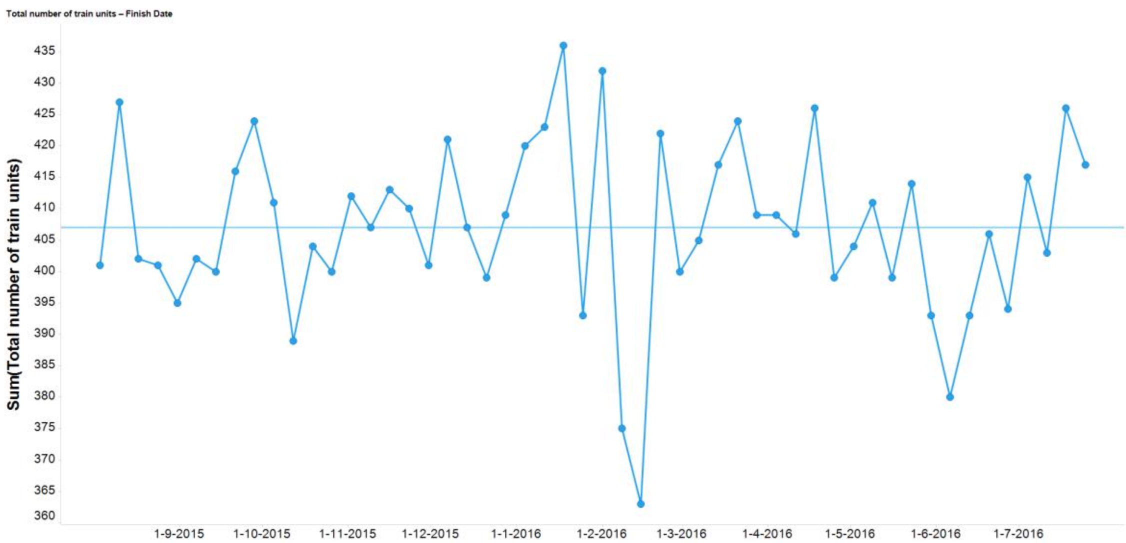


Figure 25: Total number of train units worked on at all Service Locations on all Mondays

The average number of train units at the Service Locations resulting from Figure 25 is 407. This indicates that, for all Service Locations on 407 train units any work is done, resulting in closed work orders that appear in the data. The total number of train units in the NSR fleet is 689. Given the assumption that 6% of the fleet is on average in a Maintenance Location, 648 train units are available for service. These 648 train units could appear on a Service Location during the day. There is no data available to directly get insight in the number of train units present at a Service Location that did not have work orders. In order to get insight in this topic, Figure 26 is introduced. In that figure, the percentage of the total fleet that is not in a Maintenance Location is given for each Monday. The Monday is chosen since it is a representative week day.

Figure 26: Percentage of train units of the full fleet worked on at all Service Locations on all Mondays

Figure 26 shows that the percentage of train units at a Service Location is pretty stable. The average percentage is 63% with a standard deviation of 2%. So, 63% of the running fleet is having at least one work order during a day. This implies that the resulting 37% of the running fleet is either running, stabled outside a Service Location or stabled at a Service Location, without having work orders. The last option is the most common option.

3.4.2 Variance per location

Figure 27 shows the variance in the demand and the production of a Service Location by giving the number of train units during the day and the number of closed work orders for all Mondays of the full year at Service Location Rotterdam, including averages. All 52 data points are indicated, representing one Monday. The data points are connected by a line that is only introduced in order to emphasize the pattern between the number of trains units and the associated number of work orders. The average results for all Service Locations are given in Appendix 5.

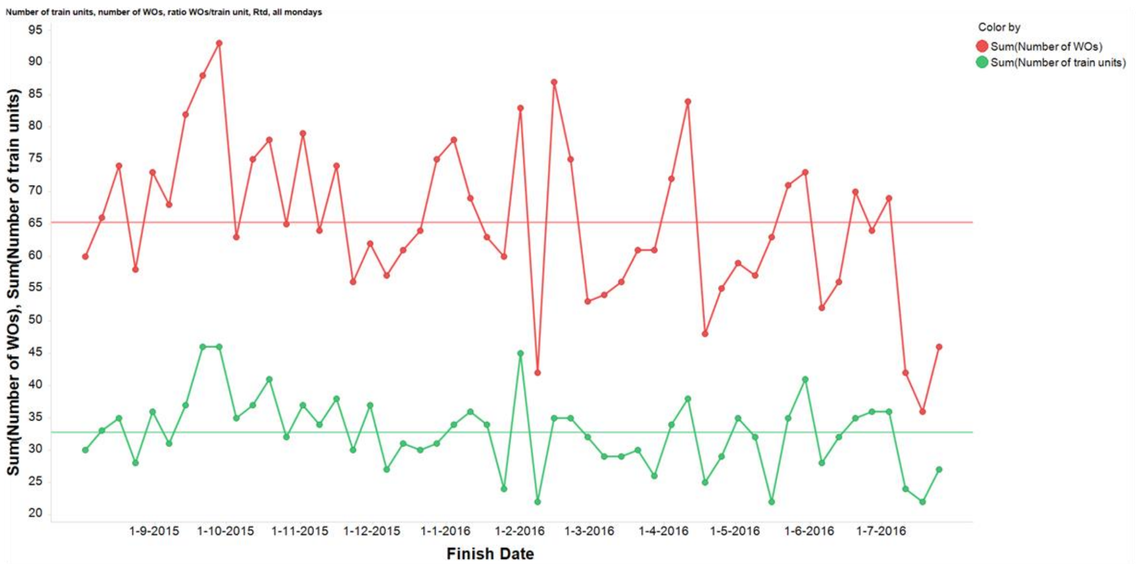


Figure 27: Number of train units worked on, number of closed WOs for all Mondays at Rotterdam

Theoretically, Figure 27 should show a more or less consistent course, as the timetable is almost the same over the year and the same number of mechanics is working on the day. This would also result in a more stable work order production, as on average the trains need the same processes. However, both the number of train units present at Rotterdam and the number of work orders finished give a more unstable course. There is a gap between the minimum train units present (22) and the maximum (46). The same goes for the closed work orders: 36 minimum, with a maximum of 93. The maximum is always more than double the minimal number. For this Rotterdam case, the average number of trains present is 33 and the average number of closed work orders is 65. The same is done for Eindhoven, resulting in matching results regarding variance and throughput.

By taking the ratio between the two, the results become even more interesting, what can be seen in Figure 28. The ratio indicates the number of work orders per train unit with at least one work order. The figure shows the ratio between train units present and closed work orders for all Mondays of the year for the Service Locations Rotterdam and Eindhoven. Figure 28 also includes the average value and the standard deviation for both locations. The ratio is more or less stable, with an average value of 2,05 work orders per train unit in Rotterdam and 2,19 in Eindhoven. The range of this ratio is 1,70 to 2,90 for Rotterdam and 1,50 and 3,00 for Eindhoven. On average, Eindhoven generates more work orders per available train unit, but the performance is more capricious. An influence factor for this is the Technical Centre in Eindhoven, that generates a lot of work orders if a train unit is serviced there. However, there are many days without any closed work order from the Technical Centre, resulting in significantly less closed work orders. The ratio for all Service Locations is given in Appendix 5.

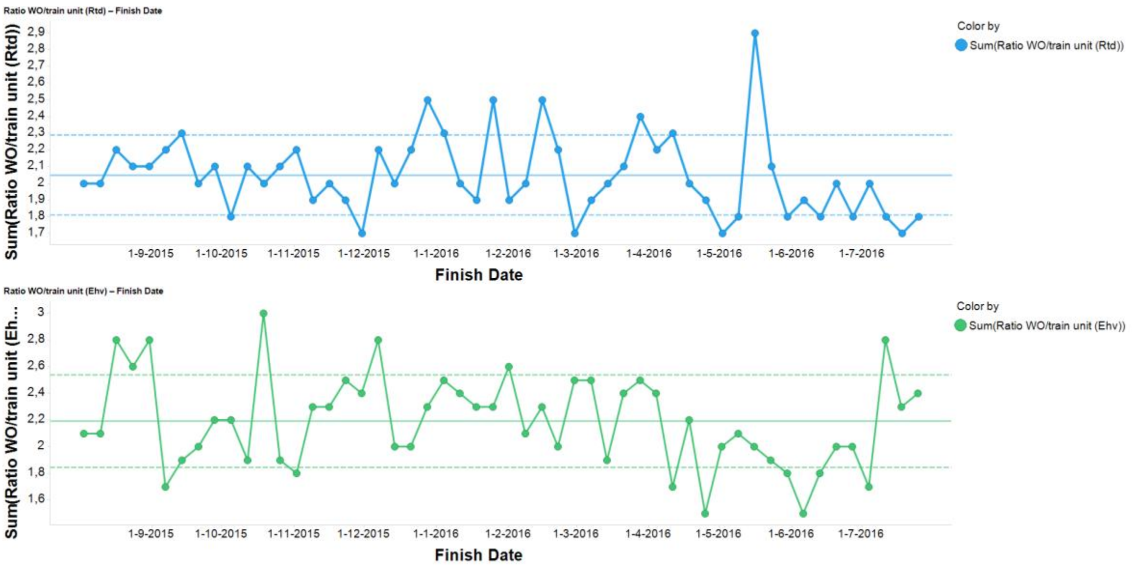


Figure 28: Ratio between number of train units worked on and number of closed WOs for Rotterdam (above) and Eindhoven (below) for all Mondays

Considering the standard deviation however, for both locations the conclusion is that there is no stable process. In Rotterdam, the standard deviation is 0,24 WO/train unit, for Eindhoven this is 0,35 WO/train unit. In general it can be stated that, especially in Eindhoven, the ratio is higher if the number of train units present is above average. This is also valid for Rotterdam, but less strong. A conclusion is that more supply results in a higher throughput. In other words: mechanics work harder if there is a need to do so. Conversely, they do less if there is less demand. The service process is able to manage variance in supply and therefore it is not directly sensitive for disruptions. There is a flexibility to handle these disruptions regarding the generated throughput. This result is only valid for the technical aspects of the work at the Service Locations. Concerning the shunting and stabling process, other conclusions will be drawn later on.

3.5 Conclusions quantitative analysis

In this chapter, the performance of the Service Locations is described based on findings from data analysis. This data set consists of all realised work orders for all locations from August 2015 until July 2016. In total, 368.000 work orders are considered. In the data, all information about the checks, repairs and washing is logged. For shunting and cleaning, there is no reliable data source.

Roughly half of all work orders is a check: an A-check, B-check or 24 hour-check. The work order distribution is comparable for all Service Locations and for all months. An exception are the washings, which can only be done at locations with a washing machine.

From the three shifts at the Service Locations, the night shift is the busiest one. This shift has as much work orders as the work orders from the early and late shift together. Besides that, also the majority of the fleet is stabled at a Service Location during the night shift. Therefore, proportionally the most checks are done during that shift.

Some locations have a large variety in train types that appear, others only handle a handful of types. The average reported times for each location vary, but no direct link to the work composition can be found. For the A- and B-check, the realised times for most locations are within a band width of 10% from the average value. Although this are several minutes, the consequences are limited. The washing production of each of the washing machines is very different: some do ten times as much as others. The washing machines at bigger locations perform better in general.

There is a distinction within the work order distribution of the various train types. Some train types have a higher share of repairs, also resulting in more work orders per train unit for those train types. This is mainly for DDAR, DDZ and VIRM. The highest number of work orders is for SLT, since they have a B-check every day. Series with the lowest number of work orders are SGM and ICM. There also exist a distinction between the average check times for the same check of the train types.

The variance within the process is shown by the number of train units present at a location, compared to the total number of realised work orders. Some dates are more than twice as busy as others, although the same day of the week is studied in all cases. The ratio between the number of train units present and the work orders done is more or less stable, whether days are busy or not.

Summarised

- A data set with 368.000 work orders from one year is considered
- Half of the work orders is a A-check, B-check or 24 hour-check
- Work order distribution is stable over the locations and over the months
- The night shift has the same number of work orders as the two daytime shifts together
- There is a variety between the train types at each location
- The performance of the washing machines is very different
- DDZ, SLT and VIRM have more work orders than other train types
- SGM and ICM have less work orders than the other train types
- There is a large variance within the number of train units present at a location
- If more train units are present, more work is done
- The number of work orders per train unit per day at a location is stable

4 Data analysis

In this chapter, the information gathered in chapter 2 and 3 is used to draw conclusions regarding the performance of the Service Locations. The chapters mentioned are more descriptive, while this chapter is analytical. This is done by combining the results from the qualitative and quantitative analysis in a reasonable way, by drawing hypotheses and checking correlations between factors. Testing the hypotheses and drawing relationships results in a comprehensive analysis of the performance of the Service Locations. This analysis is used to base the model on in chapter 5. Criteria are defined to create conditions for the modelling. Also the uncertainties are indicated.

4.1 Factors of influence

The qualitative analysis in chapter 2 and the quantitative analysis in chapter 3 are in this chapter combined to gather a complete overview of the current state of the Service Locations. The focus in those chapters is at three main points of interest: the infrastructure at a location, the rolling stock visiting the location and the variance in the process. In order to finally describe the influence of those factors to the whole process and the throughput of a Service Location, it has to be explained if there is any influence and what the nature of that influence is.

4.1.1 Infrastructure

Related to the infrastructure are the layout (chapter 2.3) and size (appendix 3). These characteristics are linked to the throughput of a Service Location. It is relevant to study the influence of the layout of a location, to find out whether a particular layout type generates a larger throughput and if so, why.

First, it is expected that the size of a location will influence the throughput of a location. Besides that, it must be clarified if the track layout influences the production, also related to the number of tracks available and the corresponding length. At multiple Service Locations, it became clear that the layout strongly influences many processes. The position of a washing machine at a Service Location can be a bottleneck on itself, as the washing process takes much time and therefore, the tracks are occupied for a considerably long period. At some locations, for example Utrecht Cartesiusweg, the available space is highly limited, resulting in trains in the washing machine occupying the central switch complex. Also the influence of the position of the arrival and departure tracks is considered to see whether notable divergences exist. The findings are linked to the performance of a location.

4.1.2 Rolling stock

The rolling stock at the Service Locations have different characteristics, regarding the size (chapter 2.4) and related standards (appendix 4). In this study, it is made clear what the influence of the various rolling stock types is on the throughput of the Service Locations. Locations with a specific mix of rolling stock types may perform better than others. This depends on the rolling stock characteristics and the mixture of rolling stock types within a Service Location.

The main topics of interest are the differences between various train types for the same work type. The realised check times, for example, are linked to one another, to find out what the differences are. Besides that, the realised work from all Service Locations is compared to one another, to find out whether there is any distinction between the locations. Also the scale of a Service Location is involved in the topic: a high number of train units of the same type may influence the performance of a location.

4.1.3 Process variance

The process variance is described qualitatively in chapter 2.5. There is a limited amount of data given about the process variance in chapter 3.4 regarding the technical performance. However, the influence of the variance is stated earlier and therefore, it is input for the hypotheses further on. The limited amount of available data is mainly caused by the fact that this phenomenon can only be studied for individual cases. This study focuses on the more broader viewpoint and therefore, these individual cases are too detailed. However, the influence of varying supply and demand is described. This description mainly focusses on the average throughput of a Service Location and the expected additional capacity needed to guarantee sufficient availability of trains for revenue service. The aspects for stabling trains during the night are described as well.

4.2 Correlations

In line with what is described in chapter 4.1, the expected correlations are studied by analysing the data gathered in the previous chapters. This is done by combining factors which are described in chapter 4.1 and explain what they show. By doing this, the hypotheses are accepted or refused. Based on that, the influence of each of the factors on the throughput of the Service Locations is clarified. The location visits and interviews resulted in findings, that are also integrated in the outcomes in this chapter. That findings explain the outcomes from the hypothesis testing, combined with literature reviews.

In this chapter, hypotheses are drawn: the null hypothesis (H_0) and the alternative hypothesis (H_1). The hypothesis are tested to correlation. The data generated in this chapter is not normally distributed, as it is not based on a sample of probabilistic data. The data points indicate a combination of properties, resulting in a position relatively to other data points, sorted out in a graph. This representation is made for all data in this chapter. Any exceptions to this are indicated if needed.

4.2.1 Statistical methods

This paragraph described concisely the statistical methods to check the hypotheses and check correlations between factors. The approach for the complete paragraph 4.2 is equal. The information is based on (Dekking, Kraaikamp, Lopuhaä, & Meester, 2005) and (Berkum & Bucchianico, 2007).

The p-value is used to test a hypothesis. It states the evidence that a random value from the data sample has to provide against the null hypothesis. To calculate it, the test statistic *T* is used, expressed as the maximum value of the random variables *X*. The p-value is calculated by taking the tail probabilities for the sample by taking the value *t* for that sample. Both the right and left tail probabilities are calculated.

$$p = 2\min(P(T \geq t|H_0), P(T \leq t|H_0))$$

The critical value for *p* is for this case 0,05. For p-values higher than 0,05 the null hypothesis is accepted, otherwise it is rejected.

The ρ-value is the Pearson correlation coefficient, also called R. It expresses the correlation between two variables *X* and *Y*. The correlation coefficient is independent from a change of units. It uses the variance and the covariance between the variables. The covariance describes to what extent both variables depend on each other. The variance described the spread of the values within the sample.

$$\rho(X,Y) = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}}$$

The value of ρ is between -1 and 1. A value of 0 indicates no correlation, 1 is perfect positive correlation and -1 is perfect negative correlation.

The R^2-value indicates the fit of the linear regression line and indicates which part of the correlation can be explained by the dependency on the other value. It is calculated by taking the square of the Pearson correlation coefficient. A R^2-value higher than 0,5 indicates a good fit, the data points within the graph are therefore in general well described by the regression line. Also the standard error for the regression line is given. This indicates the average error of a data point and indicates how good the regression line represents the data points.

The linear regression line describes the relation between the two variables *X* and *Y*. It can be used only if the R^2-value is indicating good fit. C_1 is the regression coefficient, C_0 is the intersection point with the y-axis, ε is de error term.

$$Y_i = C_0 + C_1 X_i + \varepsilon_i$$

4.2.2 Size versus production

For the first hypothesis, the correlation between the size of a location and the throughput of a location is considered. The null hypothesis and alternative hypothesis are stated as follows.

- H_0: There is no correlation between the size and the throughput of a location
- H_1: There is a correlation between the size and the throughput of a location

By combining the information gathered earlier, Figure 29 is created. In Figure 29, the total available stabling capacity of a Service Location, expressed as the number of coaches to be stabled, is compared to the total number of realised work orders for the full year of that Service Location. Every data point (dot) represents a Service Location, having the specific combination of stabling capacity and total realised work orders.

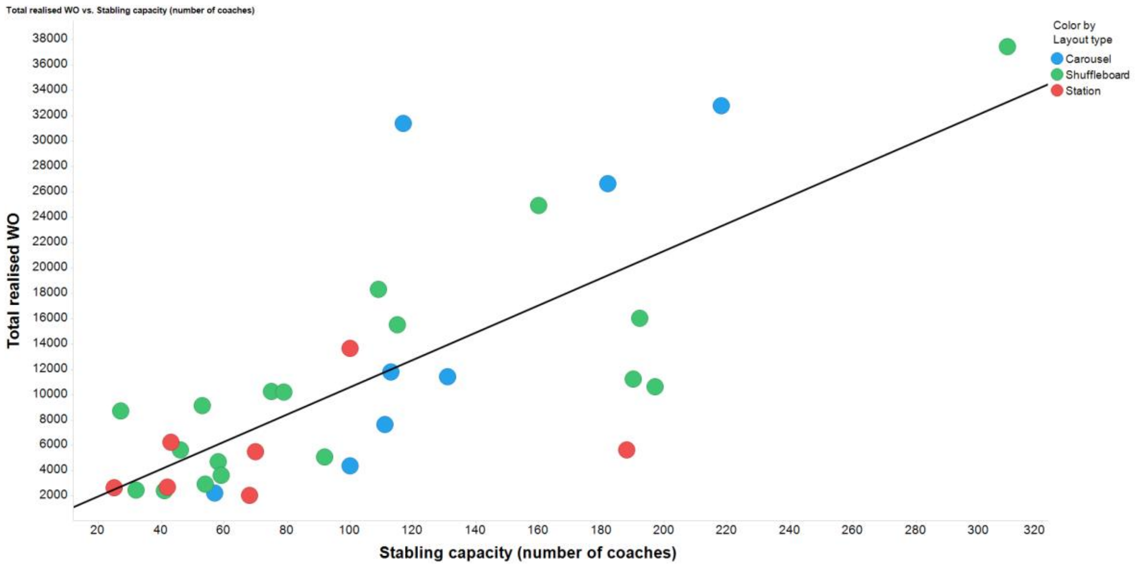


Figure 29: Stabling capacity in number of coaches versus total work orders

Stabling capacity vs. realised WO	General
p-value	4,42*10^-7
ρ-value	0,75
R^2-value	0,57
Standard error	6.427,27
regression coefficient C_1	107,57
y-intercept C_0	-160,83

Table 7: Statistical values from Figure 29

Figure 29 shows the correlation between the available stabling capacity and the total realised work orders of all types of the Service Locations for the full year. The stabling capacity is calculated by dividing the total available track length by the length of a coach. This results in a stabling capacity expressed as the theoretical number of coaches to be stabled. The fit of the line through the data points is explained later on. Table 7 shows the statistical values associated with Figure 29. A linear regression line is chosen, since the fit of this line is the best. The data points are not normally distributed, since they are not outcomes of a probability test. The values in the graph are representations of locations and therefore they are a combination of the properties on both axes. The conclusion from Figure 29 and Table 7 is that there is a correlation between the physical size of a location and the number of realised work orders, that are handled at the locations. The p-value is low and therefore, the null hypothesis can be rejected. The correlation is strong, regarding the Pearson correlation coefficient, with a value of 0,75.

Besides that, the R^2-value of 0,57 is indicating that the fit of the line for the data points given is good. By taking the line indicated in Figure 29, connecting the data points, 57% of the spread in the realised work orders can be explained by the available stabling capacity. Regarding Table 7, for every extra available stabling coach length at a Service Location, 107,57 work orders per year are expected for that location. The interception point of the regression line with the y-axis intercepts is at a value of -160,83. This indicates that with an available stabling coach length of 0 coaches -160,83 work orders could be expected over the year. This value is very low compared to the total number of realised work orders. The fit of this regression line is the best possible for this case. The standard error is 6.427,27 work orders. This indicates the average variance of the data points relative to the regression line.

Moreover, the number of realised work orders is taken into account. That does not implicitly say everything about the productivity of a Service Location, since there are work orders that take much time to complete, while others do not. To use Figure 29 as an indication of the productivity, it is assumed that each Service Location has a comparable mixture in work orders with corresponding lead times, which is also described in chapter 3.1.2.

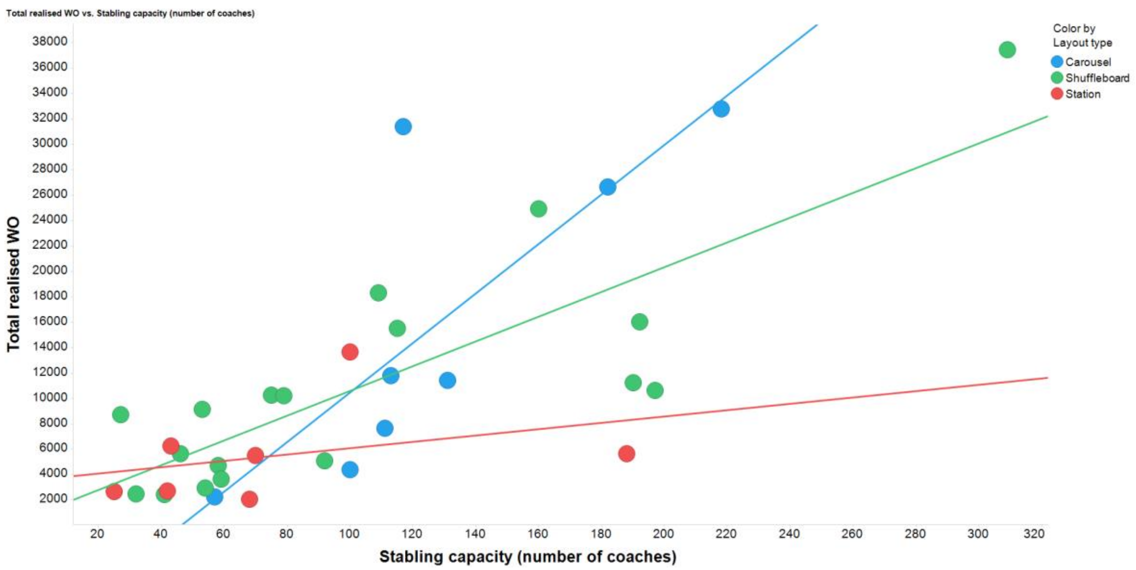


Figure 30: Stabling capacity in number of coaches versus total work orders, per layout type

Stabling capacity vs. realised WO	Carousel	Shuffleboard	Station
p-value	1,95*10^{-2}	2,48*10^{-5}	0,45
ρ-value	0,79	0,83	0,35
R^2-value	0,63	0,68	0,12
Standard error	8.150,05	5.227,16	4.075,44
regression coefficient C_1	194,96	97,44	24,95
y-intercept C_0	-9.029,41	849,02	3.589,22

Table 8: Statistical values from Figure 30

Although the null hypothesis is already explained, the influence of the layout on the production is a main study goal. In order to show the influence of each of the three layout types described, each data point is categorised in Figure 30. A linear correlation line is drawn through the data points of each of the layout types, as Figure 29 shows.

In Figure 30, the basic input is the same as in Figure 29, while in this figure the correlation for each layout type is given as well. Table 8 gives the statistical values associated with Figure 30. For each layout type, the correlation between the two factors is given by the line drawn between the data points. A conclusion is that for both the carousel and the shuffleboard layout, there is still a correlation between the size and the production of a Service Location, since the p-values are low. For the station layout, the combined layout type, there is no correlation, with only seven data points as a reference. Like in Figure 29, the regression lines in Figure 30 for values below a stabling capacity of 20 coaches are not defined, given the intercept points with the vertical axis. For the R^2-values of both the carousel and shuffleboard layout types, the fit is better than in Figure 29, so the regression lines represent the data points better.

From Figure 30 is concluded that the layout type influences the production. The shuffleboard layout generates more realised work orders below a stabling capacity of 100 coaches and the carousel layout does that for locations with a capacity of more than 100 coaches, according to the regression lines. Individual locations around this stabling capacity show different behaviour. Those locations have more work orders than locations with a comparable stabling capacity but different layout type. The influence of the station layout is neglected in this case, since there is too much dispersion in the data points and there is not enough data available. Besides that, the shuffleboard layout has the most data points and the result for that layout type matches best with Figure 29. The results from Figure 30 do not meet the expectations beforehand, based on experience and literature (Lentink, 2006). Since a carousel layout generates overall more shunting movements, the process is more sensitive for disruptions. Therefore, the shuffleboard layout in general is more accepted for a higher number of train units present.

Also when the total repair work orders and check work orders are compared to the stabling capacity, the correlation and results for the layout types are comparable to what is found in Figure 30. For washing, it is found that the carousel layout type always produces more realised work orders, but the total of twelve data points is too limited to draw conclusions from. The statement regarding the influence of the layout can be explained from the logistic point of view. The bigger the location is in terms of available track length, the more train units can visit the location, making the shunt planning more complex. The carousel layout offers more flexibility to the shunt process than the shuffleboard layout and therefore, the turning point is at about 100 coaches. An important remark is that this conclusion is based on only 35 data points in total, which does not seem to be a significant number on itself. However, these data points exist of thousands of data points each, so the data is reliable and the conclusion is feasible.

Figure 31 shows the same as Figure 30, but without the isolated data point for the carousel layout type, Service Location Zwolle. The influence of this intervention is remarkable regarding the fit of the line: this fit changes from 0,63 in the original situation to 0,95 in the situation of Figure 31. Besides that, the line somewhat shifts to the right side of the graph, resulting in a changed intersection point with the shuffleboard line. The turning point between the layout types now is at 130 instead of 100 coaches stabling capacity. Zwolle is an outlier amongst the other data points, since it has a relatively low stabling capacity, but produces a lot. This Service Location is working more at the limits of its capacity, in contrary to some other locations.

Also for the shuffleboard layout, the most influencing data point is removed in order to prove the validity of the graph. This is done by removing Service Location Binckhorst, resulting in Figure 32. Binckhorst is chosen, since it has the biggest influence on the fit of the line. Besides that, Binckhorst is a combination of two layout types and therefore does not fully describe the shuffleboard layout. Therefore, it is interesting to study the influence of this location on the performance of the layout type.

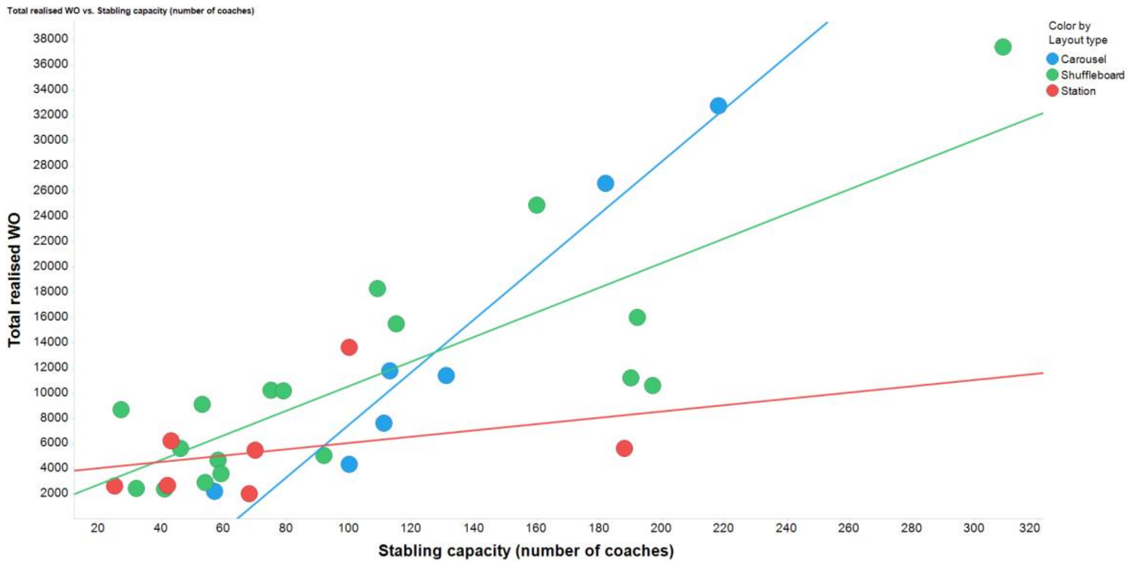


Figure 31: Stabling capacity in number of coaches versus total work orders, per layout type, without Service Location Zwolle

Stabling capacity vs. realised WO	Carousel	Shuffleboard	Station
p-value	$2{,}17*10^{-4}$	$2{,}48*10^{-5}$	0,45
ρ-value	0,97	0,83	0,35
R^2-value	0,95	0,68	0,12
regression coefficient C_1	208,45	97,44	24,95
y-intercept C_0	-13.300,48	849,02	3.589,22

Table 9: Statistical values from Figure 31

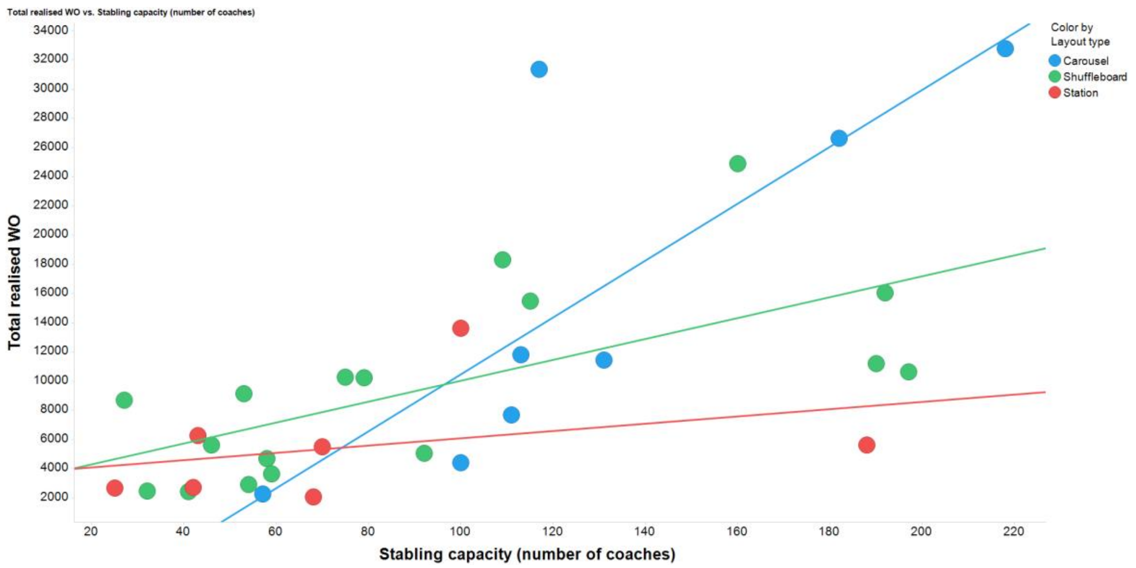


Figure 32: Stabling capacity in number of coaches versus total work orders, per layout type, without Service Location Binckhorst

Stabling capacity vs. realised WO	Carousel	Shuffleboard	Station
p-value	$2{,}17*10^{-4}$	$3{,}78*10^{-3}$	0,45
ρ-value	0,79	0,66	0,35
R^2-value	0,95	0,44	0,12
regression coefficient C_1	208,45	71,61	24,95
y-intercept C_0	-13.300,48	2.866,08	3.589,22

Table 10: Statistical values from Figure 32

Figure 32 shows the same as Figure 30, but without Service Location Binckhorst. The fit of this graph, that was 0,68 in the original case, changed to 0,44. This implies that the fit is worse, while the line is less steep. This intervention does not result in any improvement for the fit of the graph for the shuffleboard layout type. For both Figure 31 and Figure 32, removing Service Locations Zwolle and Binckhorst from the data set is a radical measure. Those Service Locations are amongst the five biggest location with the highest number of realised work orders. Their influence can therefore not be ignored by removing them. For the station layout, the same is done as described above by removing Service Location Heerlen, the most isolated value. This results in a fit that is better: R^2 is 0,56. The station layout itself is also not a main study goal and thus the results are not taken for further purpose.

To conclude this paragraph: there is a correlation between the size of a Service Location and the amount of work that is done at that location. The correlation is described with a regression formula. By making distinction between the three layout types, the shuffleboard and carousel layout type are described better, but the start values for the regression formulas are not realistic.

4.2.3 Tracks versus production

Apart from the stabling capacity, a Service Location is also defined by other track characteristics. Those characteristics are described in chapter 2.3 and appendix 3. In order to identify their influence, the hypothesis below is drawn.

- H_0: There is no correlation between the number of tracks and the realised work orders
- H_1: There is a correlation between the number of tracks and the realised work orders

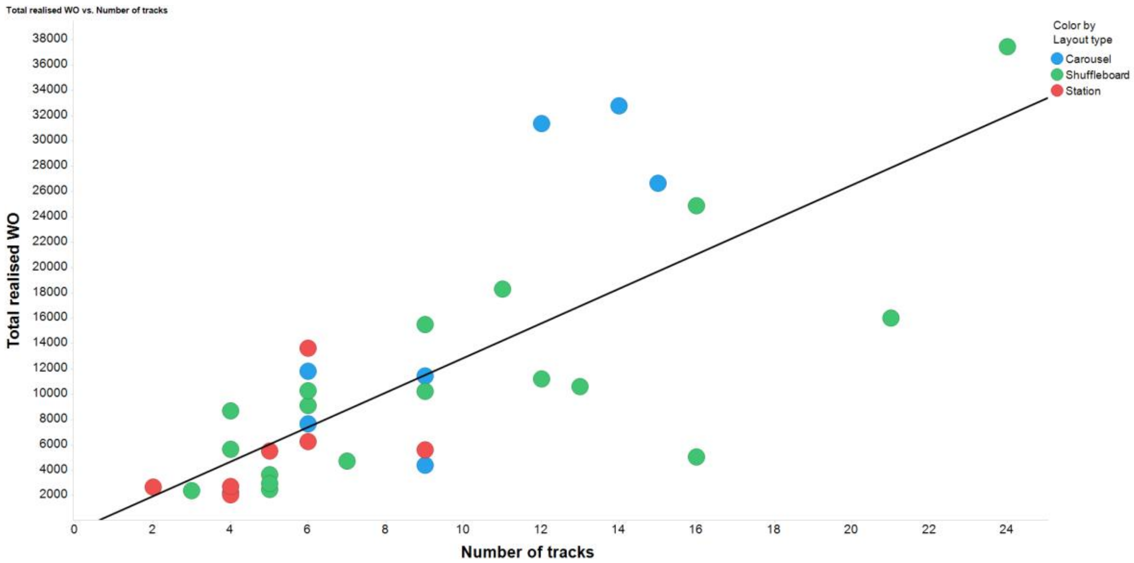


Figure 33: Number of tracks versus total work orders

Number of tracks vs. realised WO	General
p-value	$4{,}39*10^{-7}$
ρ-value	0,75
R^2-value	0,57
Standard error	6.417,08
regression coefficient C_1	1.364,89
y-intercept C_0	-775,99

Table 11: Statistical values from Figure 33

Figure 33 and Table 11 show the linear relationship between the number of tracks at each Service Location and the realised number of work orders for that location. The figure shows a correlation, also supported by the statistical data. The p-value is low and the correlation coefficient is good. The fit of the graph (R^2) is also sufficient. So, H_0 is rejected and H_1 is accepted.

Figure 34 shows the correlation between the number of tracks and the total realised work orders for the distinct layout types.

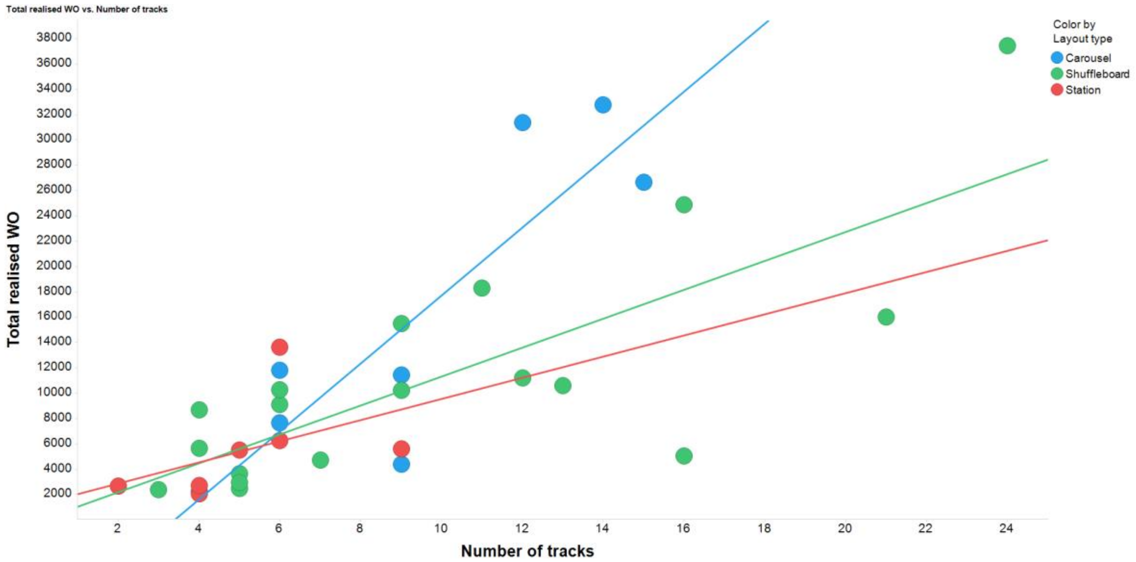


Figure 34: Number of tracks versus total work orders per layout type

Number of tracks vs. realised WO	Carousel	Shuffleboard	Station
p-value	$4{,}94 \times 10^{-3}$	$1{,}33 \times 10^{-4}$	0,30
ρ-value	0,87	0,78	0,46
R^2-value	0,76	0,61	0,21
Standard error	6.556,71	5.787,49	3.851,42
regression coefficient C_1	2.682,55	1.141,41	835,09
y-intercept C_0	-9.101,85	-91,46	1.204,97

Table 12: Statistical values from Figure 34

The correlation between the factors for the carousel and shuffleboard layout type are good, but not for the station layout. Therefore, the regression formula for the carousel and shuffleboard layout will be used for further study, since the p-value is low and the R^2-value indicates a good fit. The graph indicates that the carousel layout is preferable for a Service Location with more than six tracks. However, that conclusion is based on a very limited number of data points.

The number of tracks thus has influence on the production: the more tracks, the more work orders can be expected. An overview of this production is shown in Table 13. Table 13 shows a large variation between the realised work orders of each Service Location. The values for the total work orders per track and per available meter of track length are a normalisation of the data set. This approach implies that locations can be better compared to each other. The average produced number of work orders for a year at a track is 1226. Groningen, Hengelo, Zwolle and Watergraafsmeer are far above average. Groningen simply has a very limited number of official stabling tracks; many work is done at the station tracks. Hengelo has many work orders for DM'90 at the Technical Centre. Zwolle and Watergraafsmeer have a relatively limited stabling capacity compared to the work that is done. Regarding the realised work orders for a meter of track length, the average is 4,05 work order per year. Especially Groningen and Zwolle are notable locations. Groningen is a small Service Location with a low number of work orders and also limited infrastructure. Zwolle is a big location, with limited stabling capacity, but a high number of work orders: it is the third biggest location of all, looking at number of realised work orders. So, Zwolle does more with less infrastructure.

Service Location	Total realised WO	Total realised WO per track	Total realised WO per available meter of track length
Ah	9.134	1.522	6,47
Amf	11.213	934	2,23
Amr	10.219	1.135	4,88
Aswpl	4.393	488	1,66
Bkh	37.441	1.560	4,58
Ddr	2.718	680	2,40
Dv	2.066	517	1,15
Ehv	26.652	1.777	5,55
Ekz	3.634	727	2,33
Es	10.267	1.711	5,15
Gn	8.699	2.175	12,12
Hdr	2.416	805	2,20
Hfd	11.436	1.271	3,30
Hgl	13.637	2.273	5,15
Hlm	4.707	672	3,06
Hn	2.680	1.340	3,94
Hrl	5.631	626	1,13
Ht	2.468	494	2,89
Ldd	2.933	587	2,05
Lls	11.807	1.968	3,93
Lw	5.505	1.101	2,95
Mt	10.628	818	2,04
Nm	18.324	1.666	6,36
Rsd	15.505	1.723	5,09
Rtd	24.907	1.557	5,86
Utctw	16.037	764	3,16
Utla	5.641	1.410	4,63
Utoz	5.069	317	2,07
Vl	6.261	1.044	5,40
Vs	2.255	564	1,49
Wgm	32.786	2.342	5,68
Zl	31.375	2.615	10,13
Zp	7.673	1.279	2,60

Table 13: Production parameters for each Service Location for the full year considered

There are Service Locations that are performing weak, given a number of tracks. Especially Heerlen, Vlissingen, Amsterdam Zaanstraat and Deventer produce a low number of work orders per available meter stabling length. This implies that there is in those cases an inequality between what is asked and what can be delivered: the locations are too big for the need. Especially for Vlissingen, this is remarkable, since this location even has a washing machine, which should result in more work orders.

Another characteristic of the tracks is the average track length. The most probable link is between the average track length and the total realised work orders per track. The result is shown in Figure 35.

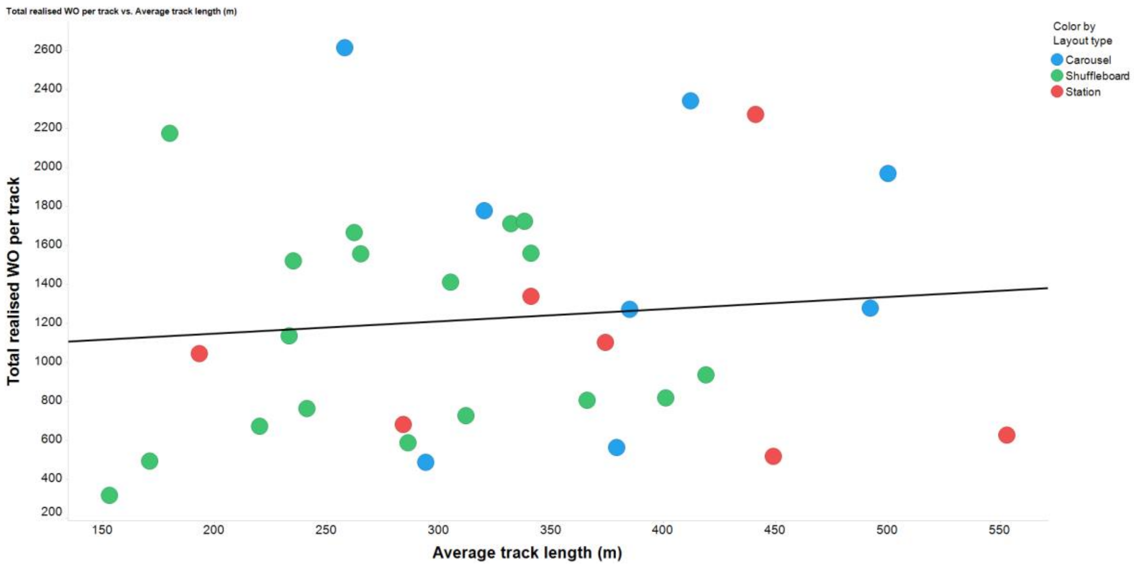


Figure 35: Average track length versus total realised work orders per track

Average track length vs. realised WO	General
p-value	0,58
ρ-value	0,10
R^2-value	0,01

Table 14: Statistical values from Figure 35

Figure 35 and Table 14 indicate that there is no correlation between the factors, since the p-value is high and the graph does not represent the data points. This is explained by the fact that the average length of a track does not directly implicate anything for the whole location. The average track length will therefore not be included for further study. The same result is obtained when the average track length is compared to the total realised work orders: there is no correlation.

4.2.4 Scale versus lead times

The second hypothesis is about the scale of a Service Location, which is described as both the number of work orders finished and the stabling capacity in number of coaches for each location. The assumption is that a larger location has a positive influence on the average B-check time of a VIRM-IV train unit. This B-check is also used (in chapter 3.3.2) as a representative work order, as it is the most common check of the most common train type. The assumption is translated into an hypothesis as stated below.

- H_0: There is no correlation between the scale and the realised times of a Service Location
- H_1: There is a correlation between the scale and the realised times of a Service Location

In Figure 36 the average realised B-check time of a Service Location is compared with the total coach capacity of that Service Location. Each data point represents a Service Location, having the specific combination of stabling capacity and average labour time.

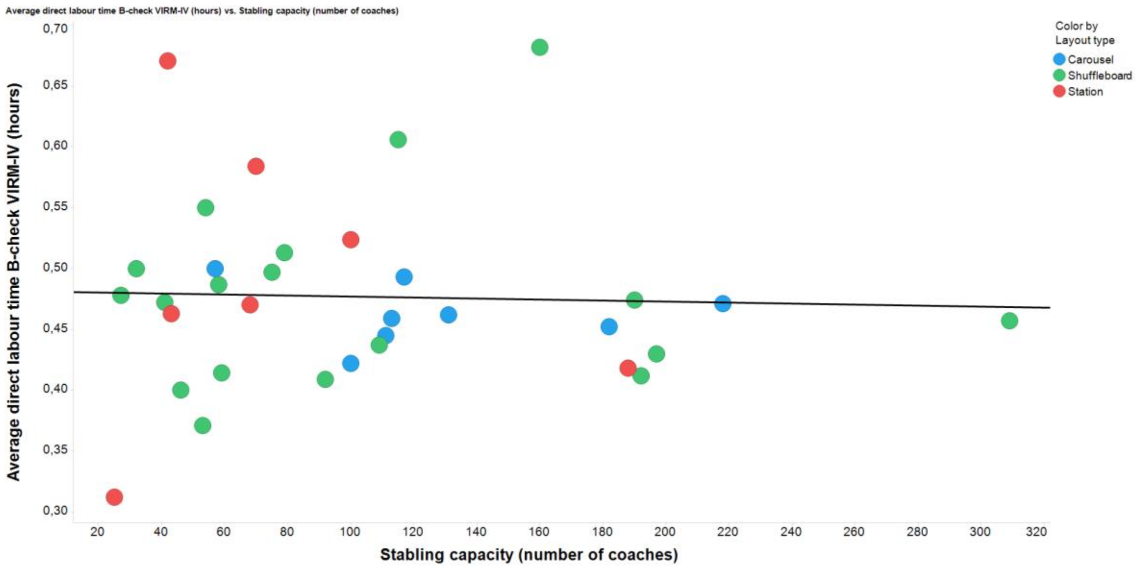


Figure 36: Stabling capacity in number of coaches versus average B-check time VIRM-IV, per layout type

Stabling capacity vs. average time	General
p-value	0,84
ρ-value	-0,04
R^2-value	0,0

Table 15: Statistical values from Figure 36

Figure 36 does not give a coherent impression: the data points are spread and not aligned. This is a first indication that there is no correlation. Table 15 gives the values related to Figure 36. The line indicated in the figure does not represent the data points correctly, although it should indicate the correlation, according to the fit of the line. The R^2-value is very low. The p-value is very high. The average of all three layout types shows a nearby horizontal line, indicating that there is no direct correlation between the factors. These results confirm the visual results from Figure 36. The data points are too much spread to give coherent results.

Based on the location visits, it was expected that a larger location would have longer average B-check times. Opposite, small Service Locations would have smaller B-check times, following this reasoning. This would have resulted in correlation between those factors. The assumption is based on the fact that walking times are on average shorter on small Service Locations (see chapter 2.7.2). Since walking times are of high influence, this result was expected. Figure 36 shows that there is no link between the factors and therewith, size has no identifiable influence on the performance of a Service Location. Since the B-check is a representative task, it is assumed that this conclusion also holds for the others tasks at a Service Location. The divergence between what was expected and what the data shows is partly caused by the data itself: the labour times are based on manual acts. Besides that, the realised times are based on human behaviour. Since most locations have fixed staff and methods, results can be influenced.

Another point of view regarding the scale is shown in Figure 37. In this figure, the number of realised work orders for the B-check of a VIRM-IV for the full year is compared to the average B-check time for each individual Service Location. The data points again indicate a Service Location with the respective combination of characteristics.

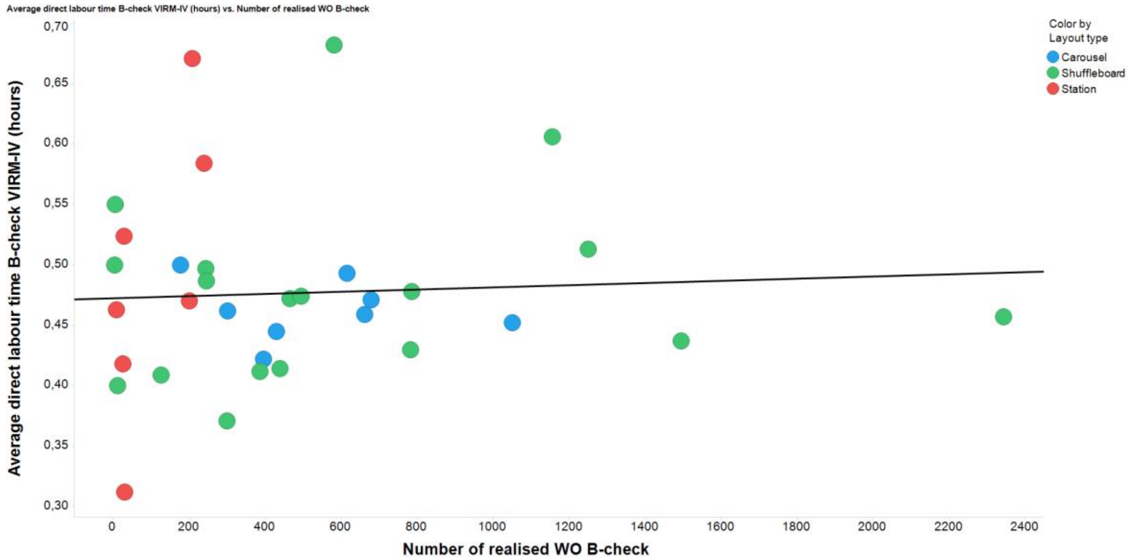


Figure 37: Realised B-check work orders versus average B-check time VIRM-IV, per layout type

Realised WO vs. average time	General
p-value	0,74
ρ-value	0,06
R^2-value	0,0

Table 16: Statistical values from Figure 37

What was, based on scale benefits, expected to see in Figure 37, was a negative correlation between the two factors. In that case, more of the same work type per location would have resulted in smaller average check time. However, the contrary is the case: there is no correlation between those two factors. The p-value is high (0,74) and the R2-value is 0. Figure 37 shows, equal to Figure 36, a large spread in data points. The regression line therefore is not very representative, however the fit is good. This does not indicate that more work of the same type at the same location results in shorter average check times. In contrary, the average check times at large locations is slightly higher, so the individual performance on average deteriorates a bit with a larger throughput. This is not in line with expectations on beforehand.

Given the results from Figure 36 and Figure 37, the null hypothesis is accepted. There is no proof for correlation between the scale and the realised time of work orders at the Service Locations. The results from Figure 36 and Figure 37 show that no distinction has to be made regarding the size of a location if the performance is considered. On average, the time consumption is almost the same for all cases looking at different aspects. The variance within the results however is very large.

4.2.5 Infrastructure capacity versus infrastructure occupation

Chapter 3.4 describes the variance in the performance of the Service Location, given the number of realised work orders for that location. In order to define the infrastructure occupation, the actual present number of coaches have to be defined. This has to be done based on the available data of realised work orders, since other reliable data sources are not available. Therefore, the hypothesis stated is that the present number of coaches can be estimated by taking the closed work orders of a shift.

- H_0: There is no correlation between the number of coaches present and the closed work orders
- H_1: There is a correlation between the number of coaches present and the closed work orders

The hypothesis is tested by checking the available data for June and July 2016 for the work orders during the night shift. These months are chosen, since this are the most recent months of the year considered. Two months secure a sufficient amount of data. Also a data set for the planned number of coaches stabled at each Service Location during the night is available for these months (NedTrain, Nachtoverstanden, 2016).

The available stabling capacity and the planned stabled number of coaches for a night are combined into a graph in Figure 38. Theoretically, the stabled number of coaches according to planning is always smaller than the capacity. At some of the locations however, the number of coaches planned is higher than the capacity according to the data. This is for example in Groningen or Zwolle. Stabling can still be feasible in this case, since during the night also some station tracks can be used for stabling. This is a common approach throughout the country. Figure 38 also shows that various locations have considerably less coaches planned than the actual capacity is. Examples are Heerlen and Maastricht. Geographical characteristics are of influence on this: although many tracks are available, not many train units are used for service in those areas.

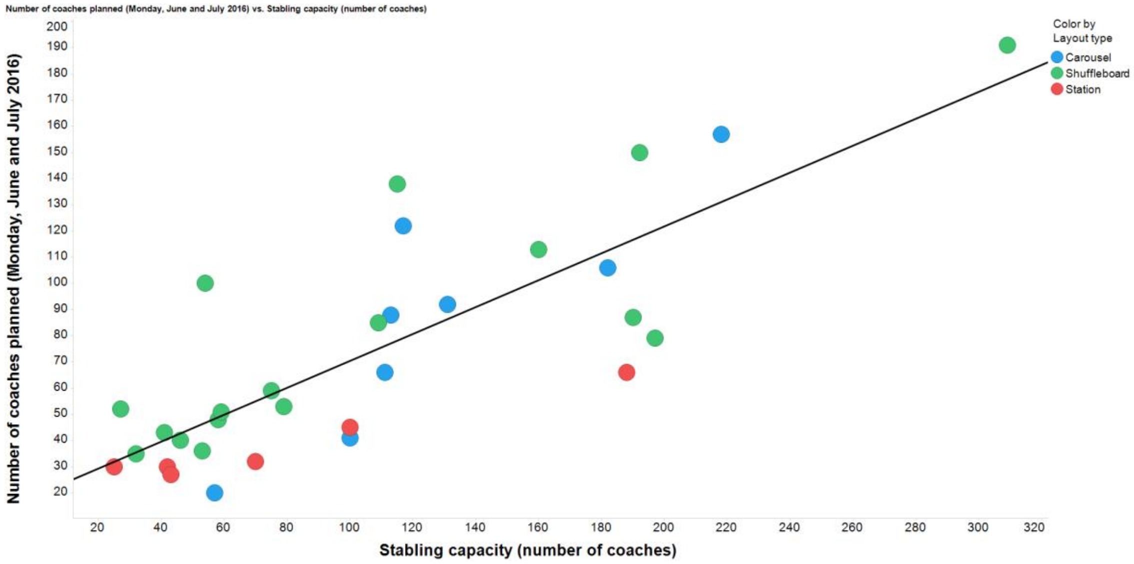


Figure 38: Stabling capacity in number of coaches versus number of coaches planned, for Mondays in June and July 2016

Stabling capacity vs. planned coaches	General
p-value	$2{,}94 \cdot 10^{-8}$
ρ-value	0,81
R^2-value	0,66
Standard error	25,88
regression coefficient C_1	0,51
y-intercept C_0	19,01

Table 17: Statistical values from Figure 38

Figure 38 and Table 17 show that there is a correlation between the stabling capacity and the planned coaches. The p-value is very low and the fit of the regression line is good. The y-axis intercept value is not 0, as would be expected, but the fit of this linear line is best, compared to the best fitting exponential line. Based on this, the null hypothesis can be rejected. The alternative hypothesis is accepted and therefore, there is a correlation between planned capacity and theoretical stabling capacity.

In order to further define the occupation of the Service Locations in practice, some method is developed to state what this occupation is. Combining the data for closed work orders and planned stabling results in Figure 39. Based on the closed work orders, insight is gathered in the number of train units (and thus coaches) present at a Service Location, for a sample of the data.

Figure 39: Planned stabled number of coaches during the night shift versus present stabled number of coaches based on work orders, for Rotterdam (above) and Eindhoven (below), June and July 2016

Figure 39 shows the planned number of coaches stabled at the Service Locations Rotterdam and Eindhoven versus the number of coaches present, based on realised work orders. Also the average present number of coaches is given, including the standard deviations. This is done for all Mondays during June and July 2016. For Rotterdam, the planned number of coaches to be stabled is 113, for Eindhoven this is 106. Figure 39 shows a variation in the present number of coaches based on work orders: on average 64 in Rotterdam and 72 in Eindhoven. Remarkable is that Eindhoven has less stabled coaches planned, but has more coaches present at the location based on the number of work orders. The standard deviation for Rotterdam is 13, for Eindhoven this is 11. This says that the process input (the number of coaches) in Rotterdam is less stable than in Eindhoven.

The most important data for all other Service Locations is given in Table 18, the complete dataset can be found in appendix 7. Remark that for Deventer, there is no data available, since there are no train units planned to be stabled. Utrecht OZ was closed for maintenance during these months, also resulting in no data. The average realised values are for the number of coaches present, based on work orders during the night shift. The standard deviations indicate that there is a variance in the realised number of coaches, since many of the values are not within the range of the standard deviations. The planned number of coaches originate from the data set. The data for all other days are given in appendix 7.

The ratio given for the standard deviation as a result of the average number of coaches present indicates the variance within the data. A small value for this ratio indicates that the number of coaches serviced is stable, while a large value demonstrates that the variance is high. This indicates that a simple approximation only based on the information given is not possible. The highest values for the ratio (shown red in Table 18) are given for Service Locations at the end of the network. That may explain that those locations are sensitive for disruptions and therefore have less stable occupation.

Finding a factor that describes the overall difference between planned number of coaches during a night and coaches that are present based on realised work orders, will help to predict the occupation of a Service Location. However, since the standard deviation/average-ratio varies, simply taking a single factor for all Service Locations does not seem to be feasible for individual days. Therefore, the same approach is followed for all week days (Monday-Friday) for June and July 2016, see Table 19. Following from this is an average factor of 2,03. So, for each train unit in the data set, on average another train unit is also present during the night at each Service Location.

Service Location	Average realised	Standard deviation	Ratio stdev/avg	Planned
Ah	26	5,4	0,20	36
Amf	35	10,8	0,31	87
Amr	20	8,9	0,45	53
Aswpl	23	7,8	0,34	41
Bkh	128	22,5	0,18	191
Ddr	22	3,3	0,15	30
Ehv	72	14,3	0,20	106
Ekz	11	5,4	0,49	51
Es	27	7,8	0,29	59
Gn	22	9,2	0,41	52
Hdr	20	4,7	0,24	43
Hfd	52	13,9	0,27	92
Hgl	19	5,5	0,29	45
Hlm	24	5,9	0,25	48
Hn	15	5,3	0,37	30
Hrl	11	4,2	0,38	66
Ht	20	4,8	0,24	35
Ldd	32	8,5	0,26	100
Lls	60	12,3	0,21	88
Lw	13	4,7	0,36	32
Mt	19	8,4	0,44	79
Nm	51	9,1	0,18	85
Rsd	58	10,2	0,18	138
Rtd	64	8,4	0,14	113
Utctw	58	22,7	0,39	100
Utla	28	6,1	0,22	40
Vl	19	3,5	0,18	27
Vs	16	5,0	0,32	20
Wgm	133	28,5	0,21	157
Zl	87	14,4	0,17	122
Zp	21	6,4	0,31	66

Table 18: Average value, standard deviation, ratio standard deviation, all based on realised work orders and planned value for number of coaches at each Service Location, for Mondays in June and July 2016

The ratio standard deviation/average value indicates the variance within the presence of coaches at a Service Location. For Service Locations with a low ratio, using the factor would give a reliable indication of the occupation. For Service Locations with a high ratio, the indication would be much less variable for an individual day. Regarding the variance in the ratio for all five week days: the larger the Service Location, the smaller the variance is. Large Service Locations (Eindhoven, Zwolle, Binckhorst) are less sensitive for deviations from the planning than small Service Locations. However, since the study goal is to gather overall insights, taking a general solution is a feasible option.

With respect to the position within the network, two types of Service Locations exist: those that are at an end point of the network and those that are in the middle of the network. Being at the end of the network implies that trains have to arrive and depart in the same direction. Being in the middle of the network means that trains can arrive and leave in multiple directions. The Service Locations at the end of the network (Ekz, Es, Gn, Hdr, Hrl, Ldd, Lw, Mt, Vl and Vs) show all high factors in Table 19. Thus, this are the Service Locations where the number of coaches stabled cannot be explained properly by the factor 2,03. This can be explained if those Service Locations simply produce less work orders. Another reason may be that the number of coaches indeed is less than planned, since locations at the end of the network are more sensitive for disruptions.

Service Location	Monday	Tuesday	Wednesday	Thursday	Friday	Avg	Stdev	stdev/ avg
Ah	1,36	2,78	1,40	1,73	1,17	1,69	0,57	0,34
Amf	2,47	2,40	1,94	2,03	1,64	2,09	0,30	0,15
Amr	2,65	2,09	2,13	1,43	1,68	2,00	0,42	0,21
Aswpl	1,77	1,19	2,36	1,92	2,01	1,85	0,38	0,21
Bkh	1,50	1,50	1,62	1,54	1,64	1,56	0,06	0,04
Ddr	1,38	2,59	1,46	1,46	1,46	1,67	0,46	0,28
Ehv	1,47	1,70	1,62	1,41	1,58	1,56	0,10	0,07
Ekz	4,58	2,87	2,63	4,00	3,31	3,48	0,72	0,21
Es	2,20	1,64	1,68	1,95	1,95	1,88	0,20	0,11
Gn	2,32	1,64	2,05	1,86	2,02	1,98	0,23	0,11
Hdr	2,18	2,41	2,58	2,41	2,51	2,42	0,14	0,06
Hfd	1,77	1,71	1,74	1,86	1,66	1,75	0,07	0,04
Hgl	2,34	1,87	1,94	1,62	1,81	1,91	0,24	0,12
Hlm	2,03	1,99	2,24	2,69	3,10	2,41	0,42	0,18
Hn	2,07	1,96	1,86	2,12	2,67	2,14	0,28	0,13
Hrl	6,00	4,10	4,07	5,76	4,16	4,82	0,87	0,18
Ht	1,76	2,01	2,57	1,89	1,79	2,01	0,29	0,15
Ldd	3,10	2,72	3,82	3,40	4,44	3,50	0,59	0,17
Lls	1,47	1,81	2,06	1,79	1,78	1,78	0,19	0,11
Lw	2,49	2,15	2,25	2,25	2,29	2,29	0,11	0,05
Mt	4,16	2,55	2,78	2,50	2,51	2,90	0,64	0,22
Nm	1,68	1,45	1,55	1,47	1,45	1,52	0,09	0,06
Rsd	2,39	1,93	2,04	2,24	2,05	2,13	0,16	0,08
Rtd	1,77	1,59	1,63	1,24	1,37	1,52	0,19	0,12
Utctw	1,71	1,45	1,39	1,17	1,58	1,46	0,18	0,13
Utla	1,44	1,32	1,77	1,40	1,84	1,55	0,21	0,14
Vl	1,42	1,21	1,54	1,69	1,49	1,47	0,16	0,11
Vs	1,27	1,05	1,36	1,11	0,90	1,14	0,16	0,14
Wgm	1,18	1,07	0,99	1,02	0,89	1,03	0,10	0,09
Zl	1,40	1,30	1,34	1,32	1,35	1,34	0,03	0,03
Zp	3,18	1,76	1,74	1,56	2,04	2,06	0,58	0,28
Factor	2,21	1,93	2,01	1,99	2,00	2,03		

Table 19: Forecast factors for all Service Locations for all weekdays of June and July 2016, including average values (avg), standard deviations (stdev) and ratio standard deviation/average (stdev/avg)

This paragraph shows the deviations from planning regarding the number of coaches present at a Service Location. By calculating average values for all Service Locations, the influence of the size of a Service Location is neglected. However, this influence is already described in chapter 4.2.2 and therefore this simplification is acceptable. Still, for the larger Service Location, the chosen approach is more reliable than for smaller Service Locations.

4.3 Other relationships

Besides the correlations, described in chapter 4.2, also other relationships exist and can be proved with the data, but not with statistical hypothesis testing. This is done in this chapter, still based on the factors described in 4.1 and based on the information from chapter 2 and 3.

4.3.1 Fixed process ordering versus random process ordering

Regarding the process analysis, it is important to find out whether there is a fixed ordering in the various process steps described in chapter 2.2. Based on that, a dependency could be described. By doing that, the performance of different steps can be coupled to each other. The hypothesis is that work at a Service Location is done in a fixed order.

- H_0: There is no fixed ordering in the process steps at a Service Location
- H_1: There is a fixed ordering in the process steps at a Service Location

An example of all visits of a train unit type ICM-III to the Service Locations is given in Appendix 6. In that example, also an impression of the ordering can be gathered. In order to explore the hypothesis, individual cases like this example have to be studied. This is done by taking realised work orders of the Service Locations Rotterdam, Binckhorst and Eindhoven in the last week of July 2016. The data of the individual days is sorted by train unit number and by work order finish time. From this data sample, 42% of the train units has more than one work order reported per day. The ordering of the work orders for this part is studied in more detail. This is done based on the finish times of the work orders, which is the most reliable source for this.

Based on the location visits and the data available, it turns out that there is not a fixed ordering in the process. Based on the outcomes of the study of a sample of the data, it cannot be stated that certain process steps follow up other process steps consequently. The only valuable finding is that a train unit that needs a check and repair(s) (65% of the cases), in about 73% of that cases is first subject to the check, followed by the repair. In total, these are 20% of all train units with at least one realised work order at that Service Location during the week. This work order is not defined in work descriptions, so it is not a rule and therefore it is not done consistently. The combination of tasks is also not the most common, a majority is only a separate check. Besides that, this combination of tasks is only possible if the work is done by a mechanic who is certified to do both checks and repairs. There are also many mechanics who can do checks only. Therefore, most of the checks are done without further repairs or with the tasks not being coupled to each other. However, the combination of a check followed by a repair is a logical one. If a defect is found during a check, it may be repaired directly afterwards, if possible.

From the data, no other tasks can be linked to each other. From practice however, it is known that washing is always coupled to shunting, before as well as after the washing. This is necessary to reach and leave the washing machine, independent from the location layout or the washing machine type. So for washing, always two shunting tasks are necessary. On top of that, also the arrival and departure of a train unit at a Service Location always is a shunting movement. The service process consequently always starts and ends with a shunting movement.

By concluding this, it is clear that the hypothesis can be rejected. Some tasks at the Service Locations are always done after another, but most of them are done separately. There is no statistical evidence that tasks are always followed by the same other task and therefore, H_0 is rejected. The findings about shunting are still valid.

4.3.2 Planned work versus unplanned work

Washing is a task which is not planned, as mentioned in chapter 2.7.2 and 3.2.3. The same applies for shunting: although this task is roughly planned on beforehand, in practice it is done ad hoc. This is in contrast to the repairs and checks. Those tasks are planned by the planners and have to be finished in order to release the train unit for revenue service. This results in a division of tasks: the planned and the unplanned work. Apart from the relevance of each of the tasks, the performance of each of them is an interesting study topic, to find out which approach works best. The hypothesis therefore is that the performance of the planned work is better than the unplanned work.

- H_0: There is no correlation between the performance of a task and the way it is planned
- H_1: There is a correlation between the performance of a task and the way it is planned

The hypothesis is explored by looking at the difference in realised work orders for washing and for checks, compared to what was expected from standards. In Figure 40, the expected number of B-checks and soap washings for the full fleet is compared to the realised work orders from the data for the full fleet. The expected number of work orders are based on the findings in chapter 2.6.1. The B-check and soap washing are chosen since these are tasks which are highly homogeneous: a B-check is always a B-check, the same for a soap washing. They are also very common at many Service Locations and therefore representative.

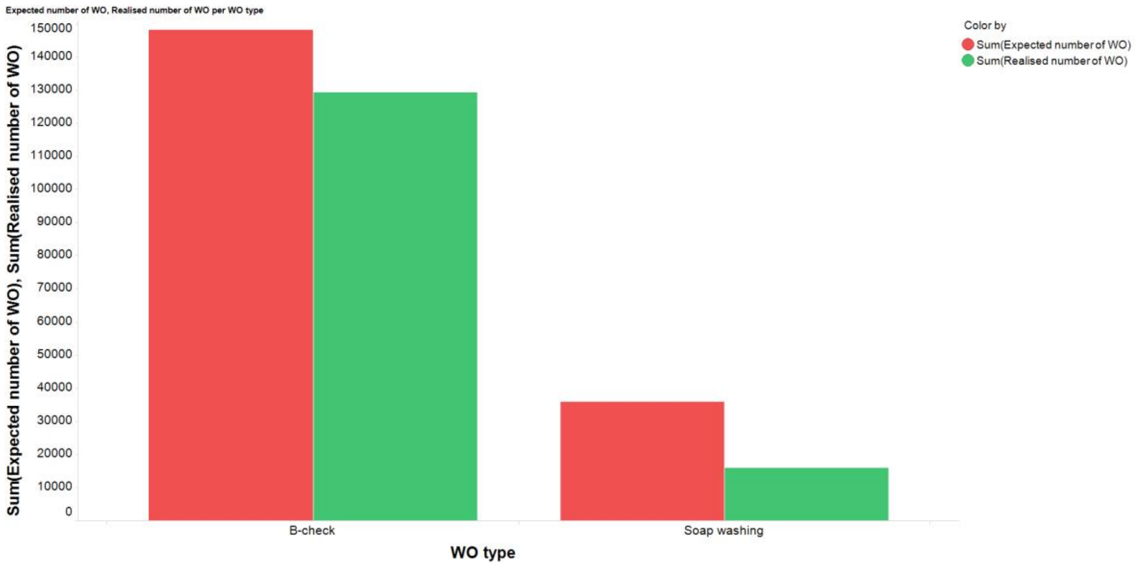


Figure 40: Expected versus realised number of work orders for B-check and soap washing

There are scale differences in Figure 40: there are much more B-checks than soap washings expected as well as realised. Regarding the purpose of the figure, it is interesting to mention the ratio between the expected and the realised work orders. For the B-check, this is 87%, for the soap washing it is 44%. Although 87% seems to be a good score for the B-check, it is still remarkable that this score is not further close to 100%. That phenomena can partly be explained by the fact that the expected values are based on rough estimations about rolling stock availability. Over the year, on average 94% of the fleet is available for service, but still a lot of reasons exist why a B-check is not done and the train is not running in revenue service. Besides that, for example the TRAXX+ICR train compositions do not have a B-check, while they are part of the expectations. Also the fact that a B-check can be postponed twelve hours without consequences will influence the total realised number of work orders.

As explained before, the B-checks are fully planned for every shift, in contrary to the washings. The results from Figure 40 cannot completely be due to the fact that B-checks are planned and washings are not planned. Also the effect is of influence on these results: a B-check that is not done, results in a train unit that is not available for revenue service. A train which is not washed, can run infinitely, up to certain limitations. This is mainly a result of policy: it is a management decision to accept that trains can run in revenue service being dirty. Moreover, things are changing: during this study project, constraints for dirtiness are sharpened. Train units that are not washed for weeks, can be forced to go to a washing machine before they will released for revenue service. The effects of this decision cannot be measured from the study, as the data used is generated before that measure is implemented.

However, also based on the theory, it can be explained that planned tasks perform better. Also for the service process overall, it is better to plan everything. By doing that, the process is more predictable and will therefore become more reliable. With all tasks known on beforehand, the planning can be made much more precise than is done nowadays. Another important remark is that washing nowadays highly influences the overall performance of a location if a train unit is washed. At most of the Service Locations, going to the washing machine has direct influence on other tasks, since a shunting movement is necessary. The physical effect of that decision is that a train unit has to be moved, resulting in infrastructure occupation. Regarding the layout of many locations, infrastructure occupation can cause delays in other processes, especially when the occupation lasts as a long as a washing can take (up to 20 minutes). The ad hoc approach of the washing planning, done during the shift at a location, results in a lack of overview in the other processes. Due to this lack of overview other processes, also later on, can be disturbed, for example as a result of a shortage in infrastructure availability. This depends very much on the location and the layout characteristics.

Based on the example from the data only, the null hypothesis H_0 cannot be accepted. Also the explanation does not help to fully except the hypothesis. However, theory tells that the basic idea from the hypothesis is correct. Planning the full process will result in a more stable process. This can only be achieved by integrating the processes from NSR and NedTrain to come to an optimal result.

4.3.3 Washing machine versus no washing machine

One third of all Service Locations has a washing machine. For all analyses done before, the work orders for washings are included in the data set. In order to find out to what extend the presence of a washing machine has any influence on the remaining work orders, those work orders have to be excluded from the data set. This is done in Figure 41, to study the hypothesis below.

- H_0: There is no link between the presence of a washing machine and the total work orders
- H_1: There is a link between the presence of a washing machine and the total work orders

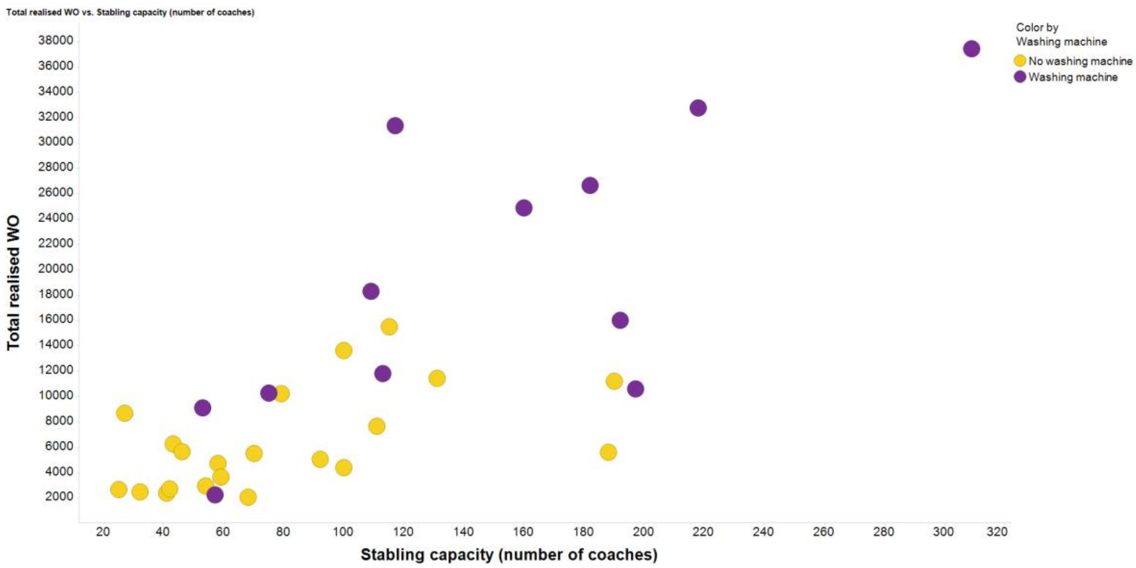


Figure 41: Stabling capacity in number of coaches versus total realised work orders, washing work orders excluded, for all Service Locations

Figure 41 shows that most washing machines are located at bigger Service Locations, as stated earlier. Since the washing work orders are excluded from the data, the remaining work orders (checks, repairs) can be directly compared to each other. All Service Locations can do the remaining work orders. Only Service Locations with a more or less equal stabling capacity can be compared to each other, since it is proved that there is a correlation between the stabling capacity and production (paragraph 4.2.2). This results in three groups of Service Locations that are compared to each other: with a stabling capacity of about 60 coaches (Vs, Ldd, Ekz, Hlm, Ah), 110 coaches (Zp, Lls, Rsd, Nm, Zl) and 190 coaches (Mt, Hrl, Amf, Utctw, Ehv).

Stabling capacity	Average washing machine	Average no washing machine
60	4.958	3.758
110	16.883	11.589
190	15.719	8.422

Table 20: Average realised work orders per stabling capacity cluster

The average values for the realised work orders for each of the three capacity clusters are given in Table 20. For each of the capacity clusters, the Service Locations perform better on average. Still, the washing machines in Vlissingen and Maastricht perform worse than the average of their capacity clusters. However, the conclusion is that on average the Service Location with a washing machine perform better than Service Locations without a washing machine. So, the null hypothesis can be rejected. On average, Service Locations with a washing machine have 1,5 as much work orders realised than locations without washing machine. However, there are other factors that influence the number of realised work orders. Those factors are not taken into account here, but are covered in other chapters. In other words: the average better performance of locations with a washing machine cannot be explained only by the fact that they have a washing machine.

4.4 Conclusions and criteria data analysis

In this chapter, the information gathered in the previous chapters is combined to describe the Service Locations. This is done by finding correlations between factors and describe other relationships. The input for all correlations come from previous chapters.

The stabling capacity of each Service Location is compared to the total realised work orders of that location. There is a correlation between the two factors. There is also a correlation between the two if the a distinction is made towards the layout types. The shuffleboard layout performs better with a stabling capacity below 100 coaches, the carousel otherwise. Also the number of tracks have a correlation with the realised work orders: the more tracks, the more work orders.

The size of a Service Location does not have any direct influence on the production times. Nor the size of the infrastructure, nor the number of work orders of the same type have any provable correlation. This implies that using average values for realised times of the tasks is a feasible approach.

The number of coaches at a location defines the required stabling capacity. To give an estimation of the number of coaches present, the planning is compared to the number of coaches present based on work orders. By doing this, the number of coaches that do not have a work order can be identified. On average, there are roughly twice as much coaches present at a location than coaches that have a work order.

The service process itself is studied as well. It turns out that there is no fixed ordering between the different tasks. The only feasible conclusion is that each visit to a Service Location starts and ends with a shunting movement. Another conclusion is that planned tasks perform much better than unplanned tasks. The percentage of work orders done based on standards for the B-check is 87%, compared to only 44% for the washings. Planning of tasks therefore helps to finish them.

Service Locations at the end of the network have less work orders than centrally located ones with comparable characteristics. It can be explained by the higher sensitivity for disruptions, resulting in less train units.

Having a washing machine results in boosting other work for a Service Location. Next to the washing work orders, also more work orders of the other types can be expected for locations with a washing machine. This is explained by the higher number of train units that are send to those locations. Besides that, most Service Locations with a washing machine are the bigger locations.

Summarised

- A higher stabling capacity results in more realised work orders
- The shuffleboard layout performs best with a stabling capacity below 100 coaches
- The carousel layout performs best with a stabling capacity above 100 coaches
- A higher number of tracks results in more realised work orders
- The size of a location does not influence average lead time of a work order
- A higher stabling capacity results in more train units planned for stabling
- For each train unit with a work order during the night, there is one without a work order
- There is no fixed ordering within the service process
- Planning work orders results in more realised work orders than unplanned work
- A washing machine results in more work orders for checks and repairs
- A location at an end of the network has less work orders than other comparable locations

5 Model design

In this chapter, all information previously gathered is translated into a model to describe the expected number of work orders and stabling capacity of a Service Location. First, a basic formula for the capacity is given. The process within the Service Locations is described in a qualitative and quantitative way in the previous chapters. Based on this description, it is known what the key elements of the performance of these locations are. Those key attributes are used to make a model, which is described in this chapter. It is tested and the results from the model are given. A sensitivity analysis indicates how reliable the model is and what changes if attributes change.

5.1 General capacity description

The final goal of the study is to describe the estimators for the capacity of the Service Locations. By determining the actual production of the Service Locations, which is the throughput of a location, the capacity can be defined. This is explained later on. For this study, it is assumed that the production of the Service Locations of today is equal to the capacity of those locations. The current production therefore is assumed to be the maximum production of each Service Location. Since the established time consumption based on the data from realised work orders is not reliable, the influence of time on the capacity is ignored in this study. Moreover, it is assumed that time is not the limiting factor to finish the work orders and tasks required, since time strongly depends on available staff.

The theoretical capacity is depending on time. The production is the number of tasks finished in a certain time span, which can be translated into the number of coaches serviced and the associated service times. The capacity is described as the number of coaches handled per unit of time, which can be a shift (8 hours) or a day (24 hours). Therefore, the process has to be defined by the time consumption of the process steps. This capacity definition is for a Service Location, so not for individual trains, with $cap = capacity,\ prod = production,\ n = number,\ t = time,\ l = length$:

$$Cap_{location} = \max(Cap_{stay};\ Cap_{shunt};\ Cap_{clean};\ Cap_{wash};\ Cap_{check};\ Cap_{repair}) = Prod_{location}$$

All parts of the capacity formula are defined separately, with *t* the time and *n* the number:

$$\begin{aligned}
Cap_{stay} &= \Sigma\, n_{coaches} * t_{stay} / \Sigma\, l_{tracks} \\
Cap_{shunt} &= \Sigma\, n_{trains} / t_{shunt} \\
Cap_{clean} &= \Sigma\, n_{coaches} / t_{clean} \\
Cap_{wash} &= \Sigma\, n_{coaches} / t_{soap} + \Sigma\, n_{coaches} / t_{oxalic} + \Sigma\, n_{trains} * 2 / t_{head} \\
Cap_{check} &= \Sigma\, n_{coaches} / t_{check,A} + \Sigma\, n_{coaches} / t_{check,B} \\
Cap_{repair} &= \Sigma\, n_{trains} / t_{repair}
\end{aligned}$$

For this basic model, the capacity is assumed to be constant over time. All individual capacity indicators have to be compared to the maximum defined capacity for that indicator for the specific Service Location. In this study, the data analysis is used to specify the characteristics of the different influence factors per process step, for each location. This is used to predict the production of a location, given the characteristics.

5.2 Model description

The capacity of the Service Locations in this study is described based on the realised work orders of the Service Locations, as used throughout the study. Consequently, the capacity prediction is in fact a production prediction, since it is based on empirical analysis. As a result of this approach, the capacity formulation is different from the formulation of the theoretical capacity from paragraph 5.1.

From the previous study, it turned out that the stabling capacity during the night is the normative indicator amongst all indicators of the performance of a Service Location. Therefore, the capacity is described as a result of the number of realised work orders. This approach has the advantage that it defines both the capacity and the production (number of realised work orders) of a Service Location.

The following approach is used to describe the stabling capacity, based on an estimation from the model:

- The theoretical stabling capacity is the maximum number of coaches to be stabled at the same time at a Service Location (paragraph 5.1).
- An estimation of the stabling capacity of a Service Location is described by a model, based on the expected number of work orders for that location and the mix of train types for that location.
- The expected number of work orders is the result of the product of all influence factors that describe the production of a Service Location, based on the findings from chapter 2, 3 and 4.

First, the influence factors are described, afterwards the model is formulated.

5.2.1 Influence factors

The analysis done in chapter 2 and 3 results in chapter 4 in the description of correlations and relevant links. Based on this, the influence factors can be described. All factors that appeared to have no influence are excluded from this chapter.

Layout type
The two distinguished layout types have different correlations with respect to the other influence factors. Therefore, two separate models are made to describe the capacity, based on the layout type selected. Furthermore, a general model is made, that defines the expected production for the case the layout type is not specified.

Stabling capacity
The total theoretical stabling capacity has an influence on the production, as is proven in paragraph 4.2.2. Each of the two layout types have a separate regression model. This model is used to define the stabling capacity in the full model.

Number of tracks
The available number of tracks describes the production of a Service Location, as shown in paragraph 4.2.3. Each of the layout types have a separate regression model. The regression model is used to define the influence of the number of tracks on the production in the full model.

Washing machine
A washing machine itself generates extra work orders, as shown in paragraph 3.2.3. Besides that, the performance of Service Locations with a washing machine is different from the performance of locations without a washing machine, as described in paragraph 4.3.3. Therefore, a factor is introduced to describe the influence of having a washing machine. Two types are identified: an oxalic washing machine and a regular washing machine.

Position within the network
Service Locations positioned at the ends of the network are in general smaller but have a less stable input of rolling stock, see paragraph 4.2.5. Estimating the production of a Service Location at the end of the network gives a less reliable indication of the actual capacity than the estimation for a more

central Service Location. In order to account for this, a factor is added to the model, indicating the position within the network.

Rolling stock
The distribution of rolling stock types at a location influences the total number of work orders, since some rolling stock types generate on average more work orders than other types. This is described in paragraph 3.3. The average number of work orders for each rolling stock type is indicated.

5.2.2 Basic model formulation

The statements from the previous paragraph are used to base a model on. It is based on the linear regression formula for the stabling capacity (paragraph 4.2.2) and number of tracks of each of the layout types (paragraph 4.2.3). SPSS statistics is used to combine these regression models into a multiple linear regression model. The regression model is defined based on the assumption that the expected number of work orders from that model is the maximum number of work orders for that location. This can be stated, since the initial assumption from the analysis in chapter 2, 3 and 4 is that all Service Locations are performing at their maximum capacity during the period considered.

Based on the findings from paragraph 5.2.1, three multiple linear regression models are made. The output for the multiple linear regression model is given in Table 21.

Number of tracks and stabling capacity vs. realised WO	General	Carousel	Shuffleboard
p-value number of tracks	0,08	0,14	0,53
p-value stabling capacity	0,08	0,68	0,07
ρ-value	0,78	0,88	0,83
R^2-value	0,61	0,77	0,69
Standard error	6.197,85	7.046,92	5.327,04
regression coefficient C_1	734,27	2.206,70	302,41
regression coefficient C_2	57,73	44,73	75,55
y-intercept C_0	-1.332,14	-10.394,28	187,60

Table 21: Statistical values multiple linear regression models

Table 21 shows that for all three models, the correlation factor (ρ) and the fit (R^2) of the models is good. Regression coefficient C_1 is for the number of tracks, C_2 for the stabling capacity. However, the p-values are insignificant in all cases. The variables number of tracks and stabling capacity are dependent variables and therefore they cannot be used together. Those variables are correlated themselves. Given this, one of the two variables implicitly also defines the other one and thus, one of the two initial linear regression models (paragraph 4.2.2 for stabling capacity and 4.2.3 for number of tracks) is sufficient to predict the expected number of work orders for a Service Location. Thus, the model outcomes from Table 21 will not be used for further study. The table is only given to support this decision.

The relevant tables with statistical values for both variables from previous paragraphs are again given as Table 22 and Table 23.

Stabling capacity vs. realised WO	General	Carousel	Shuffleboard	Station
p-value	$4{,}42 \cdot 10^{-7}$	$1{,}95 \cdot 10^{-2}$	$2{,}48 \cdot 10^{-5}$	0,45
ρ-value	0,75	0,79	0,83	0,35
R^2-value	0,57	0,63	0,68	0,12
Standard error	6.427,27	8.150,05	5.227,16	4.075,44
regression coefficient C_1	107,57	194,96	97,44	24,95
y-intercept C_0	-160,83	-9.029,41	849,02	3.589,22

Table 22: Statistical values for correlation stabling capacity versus number of realised work orders

Number of tracks vs. realised WO	General	Carousel	Shuffleboard	Station
p-value	$4{,}39*10^{-7}$	$4{,}94*10^{-3}$	$1{,}33*10^{-4}$	0,30
ρ-value	0,75	0,87	0,78	0,46
R^2-value	0,57	0,76	0,61	0,21
Standard error	6.417,08	6.556,71	5.787,49	3.851,42
regression coefficient C_1	1.364,89	2.682,55	1.141,41	835,09
y-intercept C_0	-775,99	-9.101,85	-91,46	1.204,97

Table 23: Statistical values for correlation number of tracks versus number of realised work orders

A distinction between the two linear regression models (for stabling capacity or number of tracks) cannot easily be made based on the statistical values of the two. Both are very much comparable (see Table 22 and Table 23) regarding fit and correlation factors. However, the p-values are slightly more significant in the case of the number of tracks. Besides that, the start values for the graphs are better in this case: all are negative, indicating that no tracks will not result in work orders. In both cases, the start value for the carousel layout is very low, not defining the model for low values of number of tracks and stabling capacity. But still, the values are comparable. Therefore, the regression model for the number of tracks is chosen to be the basic model for this formulation. For the station layout type, the general model is chosen instead of the regression model for the station layout. This is done since the model for the station model has insignificant p-values and a bad fit. All basic parameters for the model are given in Table 24.

Number of tracks vs. realised WO	General	Carousel	Shuffleboard
regression coefficient C_1	1.365	2.683	1.141
y-intercept C_0	-776	-9.101	-91

Table 24: Statistical values basic model

The basic model formulation therefore will be:

$$E[N_{wo}^0(i)] = C_0(i) + C_1(i) * N_{track}$$

$$C_0(i) = \begin{cases} -776 \\ -9.101, \\ -91 \end{cases} \quad C_1(i) = \begin{cases} 1.365 \\ 2.683 \\ 1.141 \end{cases} \quad for\ i = \begin{cases} general \\ carousel \\ shuffleboard \end{cases}$$

$$N_{wo}^0(i) = number\ of\ work\ orders\ for\ Service\ Location\ of\ type\ i$$
$$N_{track} = number\ of\ tracks$$

The model gives the expected number of work orders for a Service Location for one year. It will be extended in the next paragraphs.

5.2.3 Final model formulation

The influence factors, described in paragraph 5.2.1, are added to the basic model (from paragraph 5.2.2) to formulate the final model, from now on called model. The model formulation is described in this paragraph.

The washing machine has a clear influence on the production of a Service Location. In order to identify the influence of a washing machine on the total number of realised work orders, Figure 42 is used. The clusters that are used in Table 25 are indicated in the figure as well. These clusters are generated by choosing the Service Locations that have a comparable stabling capacity. By doing that, three clusters are considered, including 9 out of 12 washing machines. The other locations with a washing machine are out of the range of any other Service Location and therefore cannot be compared in this approach.

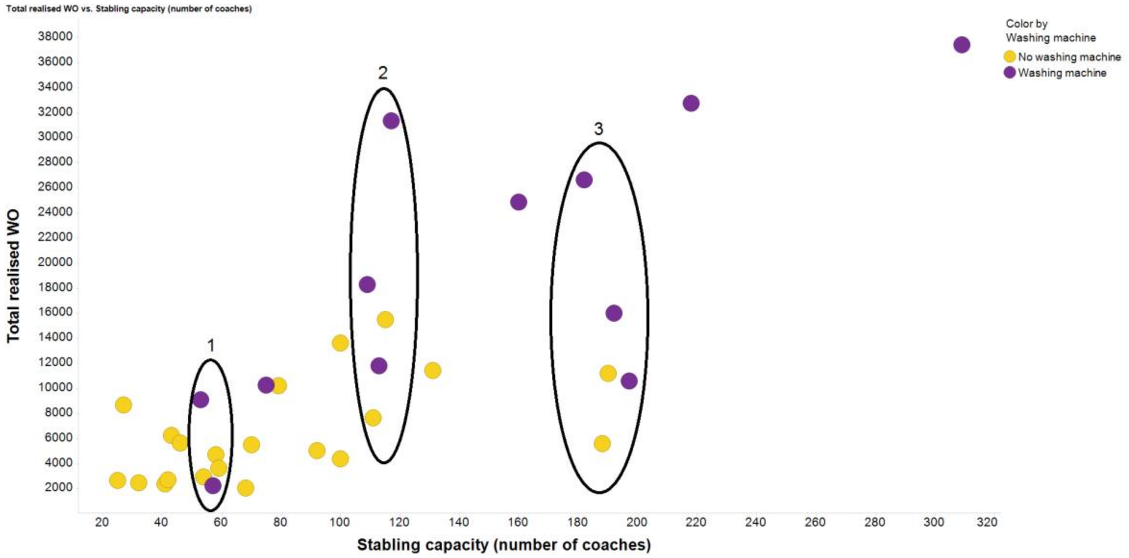


Figure 42: Clusters of compared Service Locations

Stabling capacity	Washing machine	Average number of WO	Total average number of WO	Factor
Cluster 1	No washing machine	3.758	4.238	0,89
(60 coaches)	Washing machine	4.958		1,17
Cluster 2	No washing machine	11.589	14.765	0,78
(110 coaches)	Washing machine	16.883		1,14
Cluster 3	No washing machine	8.422	12.800	0,66
(190 coaches)	Washing machine	15.719		1,23
Average	No washing machine	23.769	30.665	0,8
	Washing machine	37.560		1,2

Table 25: Influence of washing machine on realised work orders

Table 25 indicates the non-washing work orders only, so the washing work orders are excluded. The factor ($C_{wash}(j)$) indicates the divergence from the average value of realised work orders. For locations with a washing machine, a factor of 1,2 is used, for locations without one, the factor 0,8 is used. The basic model is multiplied by this factor, since it is proven that the presence of a washing machine has direct influence on the production of a Service Location.

The work orders for washing ($N_{wash}(j)$) are added to the model: 2.260 extra work orders per regular washing machine, 2.980 extra work orders per oxalic washing machine (see paragraph 3.2.3). This is the average number of washings for each washing machine type from the data: 1.325+935=2.260 for a washing machine and 2.260+720=2.980 for an oxalic machine. Including this in the model results in:

$$E[N_{wo}^{1}(i)] = (C_0(i) + C_1(i) * N_{track}) * C_{wash}(j) + N_{wash}(j)$$

$$C_0(i) = \begin{cases} -776 \\ -9.101, \\ -91 \end{cases} \quad C_1(i) = \begin{cases} 1.365 \\ 2.683 \\ 1.141 \end{cases} \quad for\ i = \begin{cases} general \\ carousel \\ shuffleboard \end{cases}$$

$$C_{wash}(j) = \begin{cases} 0,8 \\ 1,2, \\ 1,2 \end{cases} \quad N_{wash}(j) = \begin{cases} 0 \\ 2.260 \\ 2.980 \end{cases} \quad for\ j = \begin{cases} no\ washing\ machine \\ soap\ washing\ machine \\ oxalic\ washing\ machine \end{cases}$$

$$N_{wo}^{1}(i) = number\ of\ work\ orders\ for\ Service\ Location\ of\ type\ i$$
$$N_{track} = number\ of\ tracks$$

The position within the network indicates how constant the number of realised work orders is and thus how reliable the estimation for the expected number of work orders is. Locations at the end of the network have less work orders. Those Service Locations are shown in Figure 43.

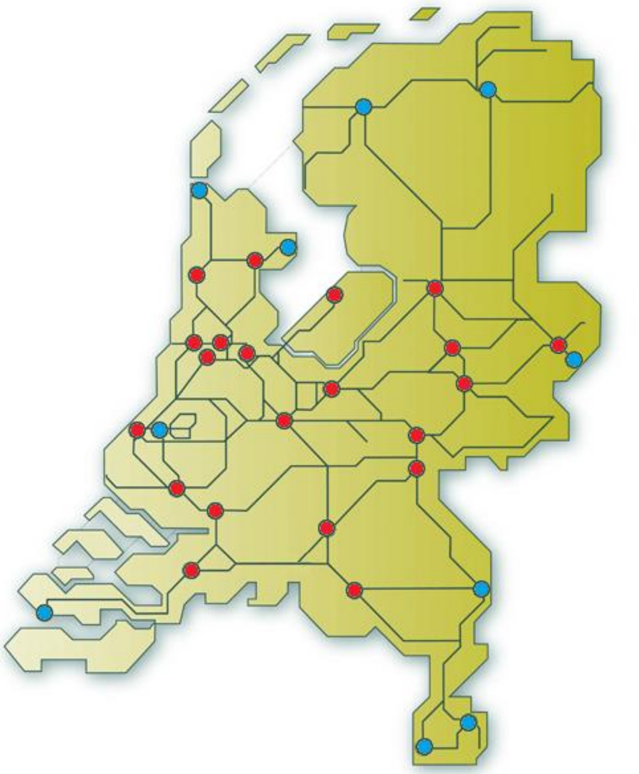


Figure 43: Service Locations (red) and Service Locations at the end of the network (blue) (based on (NedTrain, Overzicht van de NedTrain locaties, 2015))

The divergence for the locations at the end of the network from the basic model is given in Table 26.

Service Location	Number of tracks	Total number of realised WO	Expected number of WO	Difference	Deviation factor
Ekz	5	3.634	5.614	-1.980	-0,35
Es	6	10.267	6.755	3.512	0.52
Hdr	3	2.416	3.332	-916	-0,27
Hrl	9	5.631	11.506	-5.875	-0,51
Ldd	5	2.933	5.614	-2.681	-0,48
Lw	5	5.505	6.046	-541	-0,09
Mt	13	10.628	14.742	-4.114	-0,28
Vl	6	6.261	7.411	-1.150	-0,16
Vs	4	2.255	1.631	624	0,38
				Average	-0,14

Table 26: Work order variation for Service Locations at the end of the network

Table 26 shows all locations at the end of the network that are performing weak and the corresponding factor, indicating the difference between the model and realised work orders. The average value for a Service Location at the end of the network is minus 14% work orders. The factor for Service Locations at the end of the network ($C_{end}(k)$) therefore is 0,86, otherwise it is 1. The model therefore becomes:

$$E[N_{wo}^2(i)] = (C_0(i) + C_1(i) * N_{track}) * C_{wash}(j) * C_{end}(k) + N_{wash}(j)$$

$$C_0(i) = \begin{cases} -776 \\ -9.101, \\ -91 \end{cases} \quad C_1(i) = \begin{cases} 1.365 \\ 2.683 \\ 1.141 \end{cases} \quad for\ i = \begin{cases} general \\ carousel \\ shuffleboard \end{cases}$$

$$C_{wash}(j) = \begin{cases} 0,8 \\ 1,2, \\ 1,2 \end{cases} \quad N_{wash}(j) = \begin{cases} 0 \\ 2.260 \\ 2.980 \end{cases} \quad for\ j = \begin{cases} no\ washing\ machine \\ soap\ washing\ machine \\ oxalic\ washing\ machine \end{cases}$$

$$C_{end}(k) = \begin{cases} 0,86 \\ 1 \end{cases} \quad for\ k = \begin{cases} end\ of\ network \\ middle\ of\ network \end{cases}$$

$$N_{wo}^2(i) = number\ of\ work\ orders\ for\ Service\ Location\ of\ type\ i$$
$$N_{track} = number\ of\ tracks$$

Based on the model with the influence factors described above, the expected number of work orders for a Service Location can be estimated using this final model. The input to execute the model is the following:

- the layout type of a location
- number of tracks of the location
- presence of a washing machine, including washing machine type (oxalic or soap)
- location at the end of the network or not

5.2.4 Stabling capacity prediction

Based on the expected number of work orders, calculated with the model from paragraph 5.2.3, an estimation of the average number of coaches stabled at a location can be made. This estimated stabling capacity is made for the night shift, since in this shift the majority of the train units is serviced and stabled at a Service Location (see paragraph 3.1.4). Based on the model outcomes for expected work orders, 49% of them are expected to be realised during the night.

	WO/train unit/night visit
DDAR	1,64
DDM	1,37
DDZ-IV	1,88
DDZ-VI	1,96
DM'90	2,31
ICM-III	1,68
ICM-IV	1,77
ICR	1,32
MAT64	1,80
SGM-II	1,39
SGM-III	1,49
SLT-IV	1,54
SLT-VI	1,64
TRAXX	1,14
VIRM-IV	1,87
VIRM-VI	2,02

Table 27: Average number of work orders per train unit per visit during the night shift

For the estimation of the stabling capacity based on work orders, information about the train types that are stabled is needed. This is generated based on the information given in Table 27, with the average number of work orders for a train unit per night shift that the train unit has at least one work order at a Service Location.

The mix of train types that has a work order during a night shift has to be defined. The number of work orders predicted from this mix of train types has to be compared to the expected number of work orders from the model. Based on Table 27, the predicted number of work orders for a mix of trains types can be calculated.

$$E[N_{mix}] = \sum N_{unit}(l) * N_{avg}(l)$$

$$l = train\ unit\ type, \qquad N_{unit}(l) = maximum\ available\ number\ of\ train\ unit\ of\ type\ l$$

$$N_{mix} = number\ of\ work\ orders\ from\ mix\ of\ train\ types$$
$$N_{unit}(l) = number\ of\ train\ units\ of\ type\ l$$
$$N_{avg}(l) = average\ number\ of\ work\ orders\ for\ train\ unit\ of\ type\ l$$

This predicted number of work orders has to be calculated based on the mix of train types, so the exact number of train types of each type has to be given in. This calculation finally results in a predicted number of work orders, that are compared to the expected number of work orders from the model.

$$E[N_{mix}] \leq E[N_{wo}(i)]$$

The expected number of work orders from the model is defined as the maximum number of work orders that can be realised at that location. If the predicted number of work orders, based on the mix of train types, is smaller than the expected number of work orders for a location, the mix of train types can be serviced there. To estimate the stabling capacity of a location, the predicted number of work orders always has to be smaller than the expected number of work orders to have a feasible solution.

If the mix of train types results in a feasible solution, the number of coaches based on that mix is determined. Consequently, the total number of coaches stabled based on the number of work orders is calculated.

$$E[N_{coach}] = \sum N_{unit}(l) * N_{coach}(l)$$

$$N_{coach} = number\ of\ coaches\ with\ a\ work\ order$$

Based on the findings from paragraph 4.2.5, there are on average 2,03 times more coaches stabled during each night shift at a location than there are present based on work orders. The total number of coaches stabled can be calculated.

$$E[N_{stabled}] = E[N_{coach}] * 2{,}03$$

$$N_{stabled} = number\ of\ coaches\ stabled$$

Using this formulation, the total number of coaches that are stabled at a location can be predicted, based on the work orders that are expected. The complete formulation to predict the number of coaches stabled, based on expected work orders, becomes:

$$E[N_{mix}] = \sum N_{unit}(l) * N_{avg}(l)$$

$$E[N_{mix}] \leq E[N_{wo}(i)]$$

$$E[N_{coach}] = \sum N_{unit}(l) * N_{coach}(l)$$

$$E[N_{stabled}] = E[N_{coach}] * 2{,}03$$

$$l = train\ unit\ type, \qquad N_{unit}(l) = maximum\ available\ number\ of\ train\ unit\ of\ type\ l$$

$$N_{mix} = number\ of\ work\ orders\ from\ mix\ of\ train\ types$$
$$N_{unit}(l) = number\ of\ train\ units\ of\ type\ l$$
$$N_{avg}(l) = average\ number\ of\ work\ orders\ for\ train\ unit\ of\ type\ l$$
$$N_{coach} = number\ of\ coaches\ with\ a\ work\ order$$
$$N_{stabled} = number\ of\ coaches\ stabled$$

The input parameters to calculate the predicted number of coaches stabled are:

- the layout type of a location
- number of tracks of the location
- presence of a washing machine, including washing machine type (oxalic or soap)
- location at the end of the network or not
- mix of train types: number of train units of each type

The combination of models given in paragraph 5.2.3 and 5.2.4 results in an estimation of the total number of work orders expected at a Service Location. Also an estimation of the total number of coaches stabled during the night shift can be made. The results and sensitivity of the model are discussed in the next paragraphs.

5.2.5 Model results

The model described in the previous paragraphs (5.2.3 and 5.2.4) can be checked by using the real generalised work orders, known from the data. By giving in the characteristics of that Service Locations, the results from the model can be determined. The results are given in Figure 44.

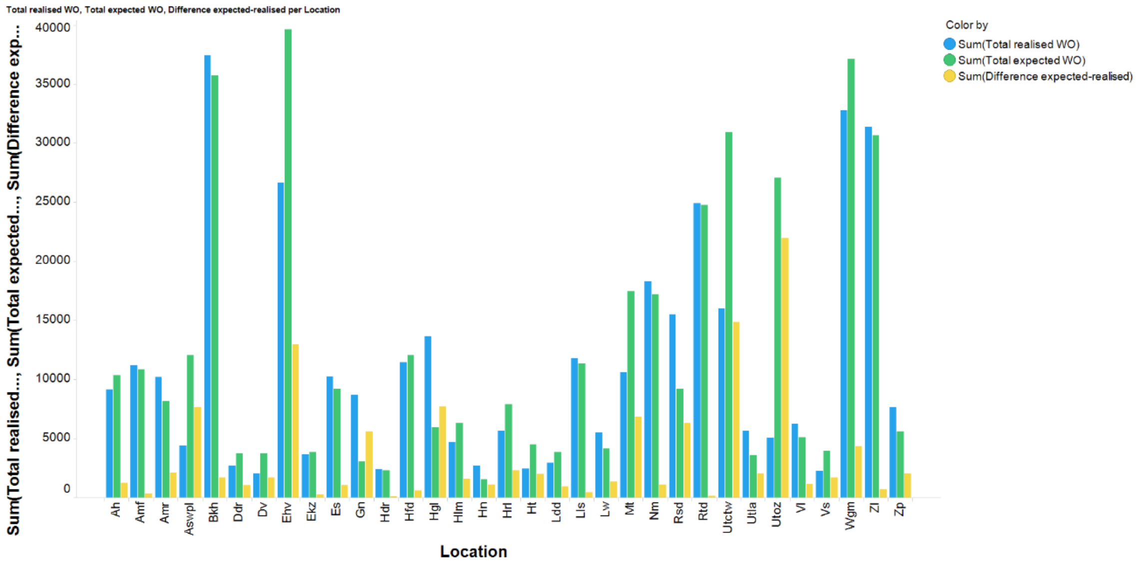


Figure 44: Total number of realised work orders, total expected number of work orders, difference between realised and expected work orders, per Service Location

The results of the expected number of work orders using the model $E[N_{wo}^2(i)]$ for the 33 Service Locations are given in appendix 8.1. The table shows the number of tracks of a Service Location and the total number of realised work orders from the data. The expected number of work orders is the model outcome. The difference in the last column shows the divergence between the total realised and the expected number of work orders as a percentage. A positive percentage indicates that the model overestimates reality, a negative percentage indicates an underestimation of reality. The results of the model $E[N_{wo}^2(i)]$ are visualised in Figure 44 and the relevant statistical values for the model are given in Table 28.

Number of tracks vs. realised WO	General	Carousel	Shuffleboard
p-value	0,06	$0{,}24 \cdot 10^{-2}$	$2{,}09 \cdot 10^{-6}$
ρ-value	0,73	0,84	0,90
R^2-value	0,54	0,70	0,81
Standard error	3.890,98	6.942,47	4.230,28

Table 28: Statistical values model $E[N_{wo}^2(i)]$

In the table in appendix 8.1, the last column is indicating the accuracy of the model $E[N_{wo}^2(i)]$. A negative value indicates that the model underestimates the real realised work orders, a positive value means that the model expects more work orders than are realised. Using the model, from the total of 33 Service Locations, 9 are within a range of 10% deviation, 12 within a 20% range and 18 within a 30% range. So, by using the model, the expected number of work orders for a location can be determined. The stabling capacity can only be calculated given a mix of train types. An example of the model for work order prediction is given in Table 29. This is an example based on a real day in Rotterdam, in order to check the results from the model. Each mix of train types that results in a feasible solution for the number of predicted work orders versus expected work orders is possible.

Expected number of WO per day Rtd		24.778/365= 68	**Expected number of WO per night Rtd**	68*0,49= 33	
Train unit	**WO/train unit**	**Coaches per train unit**	**Number of train units**	**Number of WO**	**Number of coaches**
ICM-III	1,68	3	2	3,36	6
ICM-IV	1,77	4	0	0	0
ICR	1,32	6	3	3,96	18
SGM-II	1,39	2	3	4,17	6
SGM-III	1,49	3	4	5,96	12
SLT-IV	1,54	4	2	3,08	8
SLT-VI	1,64	6	0	0	0
TRAXX	1,14	1	4	4,56	4
VIRM-IV	1,87	4	3	5,61	12
VIRM-VI	2,02	6	0	0	0
			Total	30,7	66

Table 29: Example of work order prediction for Rotterdam

The total number of coaches stabled, given in Table 29, has to be multiplied by 2,03. This finally gives an estimation for the total number of coaches that are stabled at a Service Location during the night shift. In the case of Rotterdam, this can be compared to the actual stabling capacity (Appendix 3), to check whether the solution is feasible. For this example, the expected number of coaches to be stabled is 134. This is less than the actual stabling capacity, which is 160, so the solution of the model is feasible.

5.3 Model validity

The model described in the previous paragraphs is based on real data. This is a reliable base for this model, but there are some irregularities in the data. This paragraph compares the model $E[N_{wo}^{2}(i)]$ outcomes to the original outcomes from the basic regression model $E[N_{wo}^{0}(i)]$ and introduces a test dataset. Also a sensitivity analysis is done.

5.3.1 Model comparison

The model $E[N_{wo}^{2}(i)]$, as described in paragraph 5.2, is based on the basic model $E[N_{wo}^{0}(i)]$. The model is an improvement of the basic model. The results from the model are compared to the results from the basic model, in order to determine the value of the model. A table with the results is given in appendix 8.2. The table shows the expected number of work orders for each Service Location, using the basic model and the model (same results as paragraph 5.2.5). The results are also visualised in Figure 45. This figure shows the total realised number of work orders from the data, the expected number of work orders from the model $E[N_{wo}^{2}(i)]$ and the expected number of work orders from the basic model $E[N_{wo}^{0}(i)]$. A comparison of the two expectations shows that the model is more accurate than the basic model in almost all cases.

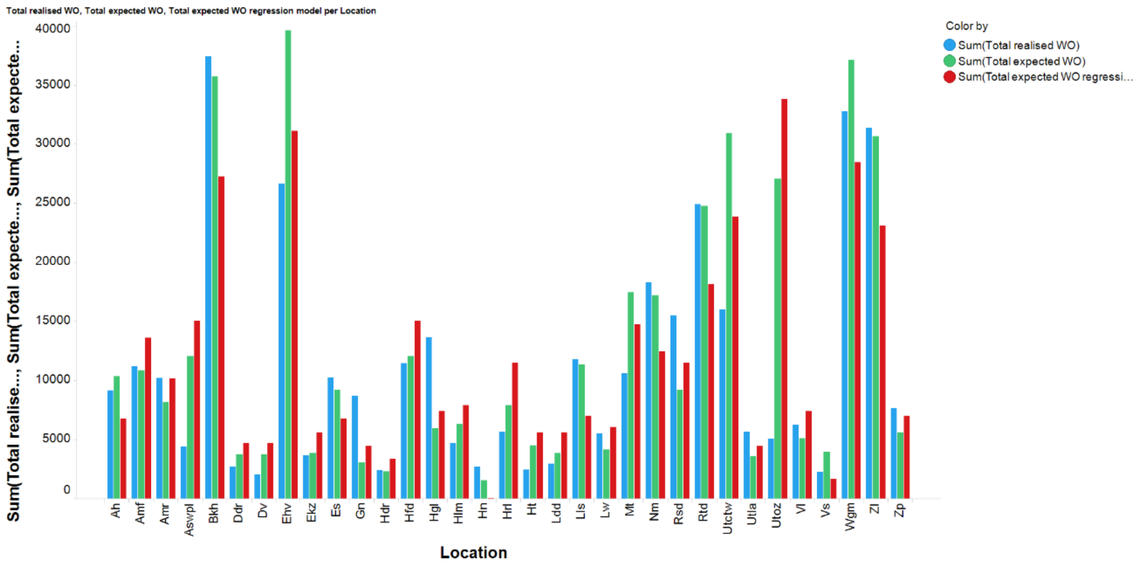


Figure 45: Total number of realised work orders, total expected number of work orders from model $E[N_{wo}^{2}(i)]$, total expected number of work orders from basic model $E[N_{wo}^{0}(i)]$, per Service Location

The difference between the expected and realised number of work orders is also indicated as a percentage. This is called the accuracy of the model and it is for both models indicated in Table 30. The table shows the number of Service Locations that are estimated within an accuracy of 10%, 20% and 30%. There is an accumulating number of Service Locations with a decreasing accuracy, so within the 20% accuracy, also the 10% is included.

Model accuracy	**Basic model $E[N_{wo}^{0}(i)]$**	**Final model $E[N_{wo}^{2}(i)]$**	**Final model without station layout $E[N_{wo}^{2}(i)]$**
Within 10%	3	9	9
Within 20%	7	12	12
Within 30%	11	18	15

Table 30: Model accuracy

Table 30 shows that the model $E[N_{wo}^{2}(i)]$ estimates the number of work orders more accurate than the basic model $E[N_{wo}^{0}(i)]$. With a maximum divergence between expected and realised number of work orders of 30%, the model makes a reliable estimation for 55% (18 out of 33) of the Service Locations.

The basic model is reliable with this accuracy interval for only 33% of all Service Locations. Therefore, the model can make a reliable estimation of the expected number of work orders. Causes for inaccurate estimations are given in paragraph 5.3.3. If only the model for carousel and shuffleboard layout are considered, the model becomes more accurate. The estimation for the station layout is not accurate, since the model parameters are not statistically acceptable and this hybrid layout type is not in the scope of the study. For the estimation of work orders within 30% of the realised work orders, now 15 out of 26 estimations are good, so 58%.

Number of tracks vs. realised WO	**General $E[N_{wo}^0(i)]$**	**General $E[N_{wo}^2(i)]$**	**Carousel $E[N_{wo}^0(i)]$**	**Carousel $E[N_{wo}^2(i)]$**	**Shuffleb. $E[N_{wo}^0(i)]$**	**Shuffleb. $E[N_{wo}^2(i)]$**
p-value	4,42*10^{-7}	0,06	1,95*10^{-2}	0,24*10^{-2}	2,48*10^{-5}	2,09*10^{-6}
ρ-value	0,75	0,73	0,79	0,84	0,83	0,90
R^2-value	0,57	0,54	0,63	0,70	0,68	0,81
Standard error	6.427,27	3.890,98	8.150,05	6.942,47	5.227,16	4.230,28

Table 31: Statistical values model compared

Table 31 shows the statistical results of the model $E[N_{wo}^2(i)]$. This is given for the expected number of work orders versus the realised number of work orders. This table is compared to the original values of the regression model $E[N_{wo}^0(i)]$ in Table 23. The most important is the standard error, that is especially better for the shuffleboard layout type, but also for the carousel layout,. All other values indicate a good fit of the model $E[N_{wo}^2(i)]$.

5.3.2 Model test

Up till now, all results given for the model $E[N_{wo}^2(i)]$ compare the model outcomes to the realised number of work orders for all Service Locations. This realised number of work orders is the same as the input of the model. Therefore, a test data set is introduced, to check whether the model gives reliable outcomes for the expected number of work orders. The test data set is a different data set than the data set used earlier. This test data set consists of all realised work orders for all Service Locations for the months February until July 2015, so half a year, in total 198.517 work orders. Since it is already proven that all months are mostly comparable to each other (paragraph 3.1.3), this is a good sample. To indicate the accuracy, the model $E[N_{wo}^2(i)]$ is used for both the test and the regular data set. The results for the model $E[N_{wo}^2(i)]$ using the test data set are given in appendix 8.3. The results are visualised in Figure 46 and the statistical values are given in Table 32.

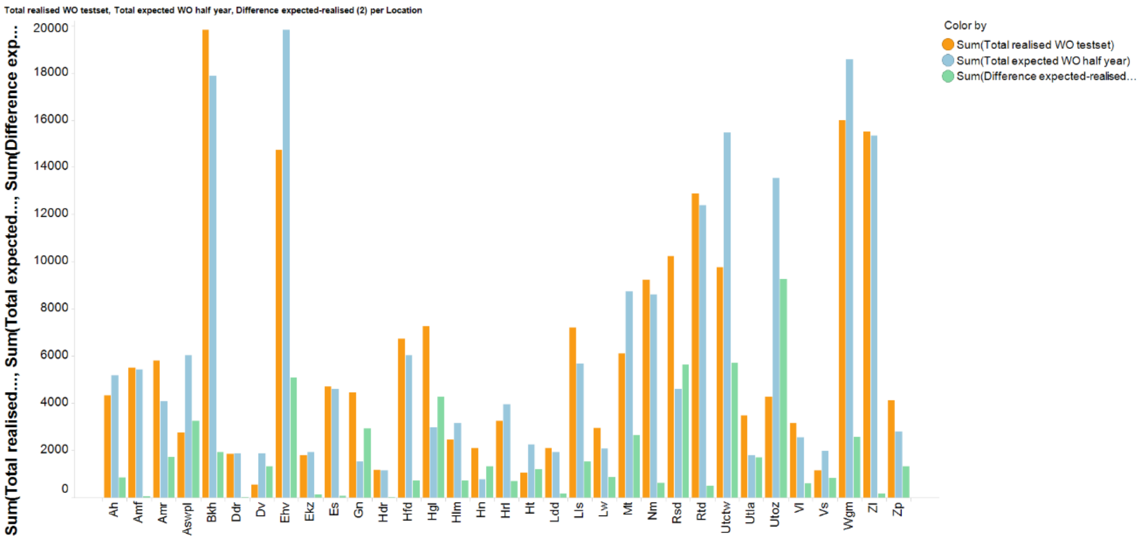


Figure 46: Total number of realised work orders from test set, total expected number of work orders for half a year, difference between realised and expected work orders, per Service Location

Figure 46 and Table 32 show, compared to Figure 44 and Table 28, an equal course and comparable values. The standard error in Table 32 should be doubled to give an indication of the performance for a year. Although there are changes and inequalities between the results for both data sets, the overall performance is the same. This is also proved by looking at the results for the expected number of work orders. These results are given in Table 33.

Number of tracks vs. realised WO	General	Carousel	Shuffleboard
p-value	0,04	$0,11*10^{-2}$	$3,54*10^{-7}$
ρ-value	0,78	0,87	0,92
R^2-value	0,62	0,76	0,85
Standard error	2.286,99	3.017,71	1.976,83

Table 32: Statistical values model $E[N_{wo}^2(i)]$ for test set

Model accuracy	Regular data	Test data
Within 10%	9	9
Within 20%	12	14
Within 30%	18	17

Table 33: Model accuracy results for regular and test data set

Table 33 shows the model accuracy for the test data set, compared to the results for the model accuracy for the regular data set. There are small differences in the accuracy between the model outcomes for both data sets. However, in general the results of both models are the same. The Service Locations that could not be estimated correctly in paragraph 5.2.5 using the regular data set can also not be estimated correctly using the test data set. So, the regular and the test data set are comparable and can therefore be interchanged to give the results from the model.

5.3.3 Model sensitivity

The results of the model $E[N_{wo}^2(i)]$ (Figure 44) show that for only 55% of all existing Service Locations a reliable estimation of the number of work orders can be made. There are many explanations and especially in individual cases an explanation is feasible. However, the model is based on the mix of properties from all locations and describes the existing Service Locations the best. There are two more general explanations for the deviation between real number of realised work orders and expected work orders based on the model.

The first explanation has to do with the number of tracks. This is a dominant variable in the model. If the number of tracks does not match the number of realised work orders, the model does not estimate the expected work orders correctly. Groningen for example has only four official NedTrain-tracks, but has many work orders on those tracks. Many work is done on regular tracks. Groningen also has relatively many work orders since the nearest other Service Location (Zwolle) is an hour away. Besides that, Groningen also receives work orders since the Maintenance Location in Onnen is located close by. Trains have to visit a Service Location before they can enter revenue service after they have been in a Maintenance Location. Vlissingen also has four tracks, but has significantly less realised work orders, although it has a washing machine. The bigger Service Location in Roosendaal is apparently more attractive for train units in that part of the country. The farther away a location is from the regression line in the basic model, the less reliable the estimate for the expected number of work orders based on the number of tracks can be made.

The second explanation is that the real Service Locations are examples from practice. This implies that there are many characteristics of a location that cannot be described by the modelling approach chosen. The model is a simplification of the reality. Utrecht Cartesiusweg is a remarkable example: it is the biggest Service Location in Utrecht. It has less work orders than expected, since the location cannot be reached from the majority of station tracks of Utrecht Centraal. Therefore, it can never get the theoretical capacity. Hengelo has much more work orders than expected, due to the Technical Centre at that Service Location. There will be many more reasons for the other Service Locations with a high difference between expected and realised number of work orders.

Another influencing factor, regarding the predicted stabled number of coaches at a location, is the variation within the process. The model is based on averages: the outcomes are given for an average day. As shown in paragraph 3.4.2, the number of work orders between the days varies. At some days, the number of realised work orders can be double the number of another day. Also shown is that the Service Location is in that case still capable to handle those work orders. The estimated outcome of the model is accurate for the average day, but a busy day can imply up to 30% more realised work orders than average. However, the stabling capacity is not 30% higher than the average, so that has to be estimated based on the outcomes of the model.

Since the model has only six input variables, each parameter has relevant influence on the model outcomes. The number of tracks at a location has a big influence, since one extra track will result in up to 2.683 more work orders. This number is based on the regression model and therefore, it cannot be varied. Thus, it is not sensitive to changes.

The other variables in the model are more sensitive to changes. Since the variables are estimated based on averages, they do not apply correctly in some cases, as the previous paragraph already indicated. Especially washing results in extra work orders that sometimes overestimate the real production of a Service Locations with a washing machine. The presence of a washing machine itself can result in up to 2.980 work orders, which is a considerable number. A more reliable distinction between washing machines cannot be made. In the case of the distinction between location at a place in the network, the factor chosen sometimes results in an overestimation of work orders, but sometimes the model underestimates the real production. The influence of individual, existing locations on the outcomes of the model is considerable. However, to give a general estimation of the performance of Service Locations, the outcomes of the model are useful.

5.4 Conclusions and results model

This chapter gives a model to predict the number of work orders that can be realised at a Service Location and the number of coaches stabled at a location. The theoretical capacity of a Service Location is defined as the maximum capacity of each of the different tasks: stabling, shunting, cleaning, washing, checking or repairing. In practice, the stabling capacity is always leading, since this is the most fixed of all. The theoretical stabling capacity is described as the maximum number of coaches that can be stabled at a Service Location at the same time.

A model is described to the number of work orders and the related stabling capacity, based on the analysis from the study. All variables having a positive correlation with the number of realised work orders from chapter 4 are used in the model. The model can be summarised into the following steps:

1. Select input variables: layout type, number of tracks, presence of a washing machine, washing machine type (soap or oxalic), at the end of the network or not
2. Calculate expected number of work orders using the model
3. Select mix of train types: number of train units of each train type
4. Calculate predicted number of work orders for mix of train types
5. Check: predicted number of work orders has to be smaller than expected number of work orders
6. Calculate predicted number of coaches stabled based on work orders
7. Calculate total predicted number of coaches stabled

The outcomes from the model give a reliable indication of the average expected number of work orders for a location. Based on that, a prediction can be made about the stabling capacity. Since the model is based on average values and is a simplification of reality, there is some difference between the outcomes of the model and the actual production of real Service Locations. The model can calculate the expected number of work orders within a range of 30% accuracy for 55% of the existing Service Locations. This is explained by the influence factors that are not taken into account in the model. Overall, the standard errors using the model are smaller and the model fit is good, so the model is an improvement of the basic model.

The model is verified by using a test set, consisting of the realised work orders of the previous half year, compared to the regular data set. The outcomes of this test set are mostly comparable to the outcomes for the regular data.

Since the model is based on only six variables, the influence of each of them is considerable. By changing the layout type, adding an extra track or adding a washing machine, the number of work orders will increase quickly. However, adding more variables is not feasible, since from the study only the variables involved turned out to be of provable influence.

Summarised

- The theoretical capacity is the maximum number of coaches stabled at the same time
- A model is made to calculate the expected number of work orders for a location
- Based on the expected number of work orders, the number of coaches stabled is predicted
- The model input is the layout type, number of tracks, presence of a washing machine, washing machine type (soap or oxalic), end of the network or not, mix of train types
- The model estimation is for most cases accurate with a range of 30%
- Standard errors are smaller by using the model for expected number of work orders
- The model outcomes give an estimation for the average performance of a location
- The model gives an estimation of the number of coaches to be stabled at a location
- The model is verified by a test set of data

6 Conclusions and recommendations

In this chapter, the results from the study are presented and discussed. The main research question is answered, which is:

How can the capacity of the Service Locations and stabling yards, given the characteristics of that location, be estimated, as a result of empirical analysis?

The results from all previous chapters are combined and based on that, recommendations for NedTrain and for further study are made.

6.1 Conclusions

The process at the Service Locations is complex. This study focusses on most parts of the processes and describes them. Cleaning is predominantly excluded from the study. Since all Service Locations are different, a general description as made in this study never completely covers the characteristics of all locations. Therefore, all locations are included in the study, but the final description is mainly based on averages and includes assumptions and simplifications. This study is a comprehensive description of the characteristics of the Service Locations, based on the three main topics of interest: the infrastructural and rolling stock characteristics and the variance within the process. The conclusions of the study are described in this paragraph. First, the sub questions are answered and afterwards, the general conclusion of the study is described.

6.1.1 Sub conclusions

In order to answer the main research question, various sub questions are posed. Those questions are all answered here.

The data needed to get insight in the production of the Service Location are the realised work orders, train units stabled and a log of all shunting movements. Only full data of the realised work orders was available. This data includes the location, rolling stock type, work order type, work order description, finish date and time, train unit number and in the majority of cases also the associated direct labour hours spent.

The variance within a Service Location depends on the planning of that location. The divergence between planned visits and realised visits can have various causes. The result can be that an extra train arrives, no train arrives, the train arrives earlier or later than planned, the ordering of the arriving trains is different, the train type is different or the train has other work than planned. Due to this, some days have more than twice as much realised work orders than other days.

The layout characteristics are based on physical appearance, combined with the related process arrangement. Two layout types are defined: the carousel and the shuffleboard. A more hybrid layout variant exist, the station layout. Each of the layout types perform different regarding the number of realised work orders.

The variation within train types at a location result in a variety of work orders. Each train type produce a different number of work orders per visit and also a different distribution in work order types. The train type mix is a factor of influence on the estimated production.

The exact function of the process at the Service Locations is to have safe, reliable and clean trains at the right time. The related processes are checking, repairing, shunting, cleaning and washing. A combination of processes together is the service process of a day.

The difference between planned work and realised work is marginal. Many work is planned in some way and that work is done regularly. This are the checks and shunting. Another part of the work, including washing and a part of the repairs is not planned, and the performance of those tasks is much worse. A majority of those unplanned tasks is not executed according to the expected finish time.

The main existing correlations within the Service Locations are between stabling capacity and number of realised work orders and number of tracks and number of realised work orders. There also exist correlations between the presence of a washing machine and the number of realised work orders and the position within the network and the number of realised work orders.

The bottlenecks within the process are shunting and planning. This combination results in missed work orders. This is caused by the layout of the locations, including the switches. Mainly switches at central points are highly occupied and that are the physical bottlenecks at the Service Locations.

The difference between theory and practice is mainly related to the process times of the different tasks. For each task, a standard time is set, but those times are different for all tasks. However, time consumption is roughly the same for all Service Locations, so the influence of this is limited.

6.1.2 Main conclusions

Many people are involved in the Service Locations, with a variety of responsibilities. Communication and planning is a main challenge. The planning problem is related to the variance within the process: a day at a Service Location is never the same as a previous or next day. Integration of tasks and responsibilities, for example between NedTrain and NSR, will result in process improvements. Stabling and the related shunting are the most critical tasks at the Service Locations. Especially for these tasks, the least information is available and the tasks and responsibilities are particularly shared between NedTrain and NSR. This lack of data influences the results of this study.

There exists a lot of variation between the Service Locations, especially regarding the size and number of realised work orders of the locations. Therefore, a direct comparison cannot be made. This study is meant to indicate how the Service Locations are comparable to each other. The most common and also most standardised processes for all Service Locations are the daily checks. Consequently, the performance of the locations can be compared by checking these tasks. However, since the registration of lead times of all processes within the management system is not reliable, comparing the time consumption of tasks is not a reliable approach. For other tasks than the daily checks, especially repairs, comparing of lead times is not possible at all, since there is a large variety within the tasks. The approach chosen in the study results in a comparison of number of realised work orders. The data study proves that the distribution of work orders over locations and rolling stock types is in general comparable.

The average performance of the Service Locations is considered to be the maximum performance. However, it is shown that the locations are capable to handle up to 30% more work orders on top of the average performance. The normative shift is the night shift. During some night shifts, twice as much work orders can be realised than during other night shifts. It can be a result of weather circumstances, but also staff occupation or the availability of open work orders can explain this.

Some relevant correlations between attributes of the Service Location performance are found. All correlations are based on the limited number of data points, each representing one out of 33 Service Locations. For some attributes, if the spread in data points is high, then no correlation is found. The correlations found all apply to one of the three layout types defined. Each of them performs different for all attributes and therefore, the process layout is the first attribute to describe a Service Location. The main correlations are found between theoretical stabling capacity and number of realised work orders and number of tracks and number of realised work orders.

Washing is an important task. The washing performance is the weakest of all tasks. Many trains are washed too late, according to the standards. Washings are not really planned, resulting in missing washings. However, data proves that the availability of a washing machine at a location is an attractor for work orders of other tasks. Also the rolling stock type influence the performance of a Service Location. The position within the train service network also is of influence of the number of realised work orders of a Service Locations. Although it is assumed that there are more attributes to the production, no further influence is proved from the data.

Based on the correlations found, a model is made. This model estimates the stabling capacity, based on average values for all Service Locations. The relevant variables to describe a Service Location are the layout type, the number of tracks, the presence of a washing machine, the position within the railway network and the mix in train types at that location. The model estimates the expected number of work orders for a location and given that, an estimation is done about the number of coaches that are stabled at that location during the night shift. The model can estimate the production of most of the Service Locations with an accuracy of 30%. This implies that the model does not declare all influence factors for the production of a Service Location. However, the model is useful, to get insight in the general expected production of a location. The possible influence factors that explain the difference between the model outcomes and real production are described based on experience from the empirical analysis. The model is the best estimation for the capacity possible, given the assumptions and simplifications in the study.

6.2 Recommendations

Based on the study and the conclusions, some recommendations are done, for NedTrain as well as for further study.

The accessibility of data is poor. All realised work orders from the Service Locations are logged in the system, but not all information is useful. On top of that, information about the most critical tasks is not accessible. The information about stabled trains is available in the system, but is not stored, making it inaccessible. Also information about shunting movements is available, but is constantly renewed and therefore not logged and stored for analysis in the system. A more intelligent use of existing data would improve and simplify study possibilities to the performance of the Service Locations.

Within the Service Locations, there are many more or less shared responsibilities, between NedTrain and NSR, but also between people. Despite that, there is nobody having the full overview over the operations and the final interest of the operations. A person who is really equipped to manage the process, at least at the large Service Locations, will definitely result in a smaller number of missed tasks.

The planning of tasks is now done for a part of the tasks: checks and some repairs. Also the shunting movements are planned. However, some other repairs and especially washing is not planned. This results in many missed washings, compared to the standards. If a train is washed, it has to be shunted towards and from the washing machine. This will in many cases result in disruptions of other processes. The integrated planning of all tasks within the Service Location will result in a more stable process. If the basic process is stable, it is well equipped to handle disruptions if necessary.

All Service Locations are different from each other and therefore, they cannot be compared directly. This study is a part of an approach to improve comparability of the Service Locations. In order to improve this and more important, to optimise planning and operations, the locations should be designed more following a standard layout. This study shows some parameters to compare the locations to each other.

6.3 Further research

To improve the results from this study, some suggestions for further study are made. First, the layout types and process activities regarding shunting have to be studied more in depth. The characteristics, definition and influence of the layout types on the logistic process have to be made clear. From the data in this study, the carousel layout performs better for bigger locations, while experts state the contrary, with the shuffleboard being better equipped as a bigger location. This has to be explained and elaborated further.

This study is a simplification of the service process. Many other influence parameters should be studied, including the influence of people, planning methods and lead times on the process. To enable the last, secure logging of lead times in the actual process is necessary. The actual data is too simple and not reliable enough, since it is recorded by people. The data registration has to be automated and standardised. Results from that will be a reliable basis for further study to differences between Service Locations.

Shunting is a crucial task and therefore it has to be studied more in depth. The actual movements of train units over a location, including occupation of tracks or switches in detail will explain the real capacity of this process. Also influence factors for duration of shunting movements have to be studied. That will be, combined with the data, give insight in process improvements for this task.

Bibliography

Bergvall-Kareborn, B., Mirijamdotter, A., & Basden, A. (2004). Basic Principles of SSM Modeling: An Examination of CATWOE From a Soft Perspective. *Systemic Practice and Action Research*, 55-73.

Berkum, E. v., & Bucchianico, A. D. (2007). *Statistical Compendium.* Eindhoven: Technische Universiteit Eindhoven.

Bokhoven, M. v. (2011). *Lean manufacturing.* MARKonTarget Productiviteitsbureau BV.

Broek, J. v., & Kroon, L. (2007). A Capacity Test for Shunting Movements. *International Dagstuhl Workshop* (pp. 108-125). Berlin: Springer.

Busstra, M., & Dongen, L. v. (2015). Creating value by integrating logistic train services and maintenance activities. *International Conference on Through-life Engineering Services* (pp. 250-254). Utrecht: Elsevier.

Chris. (2016, May 30). Team leader. (S. Janssens, Interviewer)

Daamen, W., Loot, M., & Hoogendoorn, S. P. (2010). Empirical Analysis of Merging Behavior at Freeway On-Ramp. *Journal of the Transportation Research Board*, 108-118.

Dekking, F., Kraaikamp, C., Lopuhaä, H., & Meester, L. (2005). *A modern introduction to probability and statistics.* London,: Springer Verlag.

Dewilde, T., Sels, P., Cattrysse, D., & Vansteenweg, P. (2013). Robust railway station planning: An interaction between routing, timetabling and platforming. *Journal of Rail Transport Planning & Management*, 68-77.

Freling, R., Lentink, R. M., Kroon, L. G., & Huisman, D. (2005). Shunting of Passenger Train Units in a Railway Station. *Transportation Science*, 261-272.

George, M. (2003). *Lean Six Sigma for service.* Mc Gill.

Goverde, R. (2006). A delay propagation algorithm for large scale scheduled rail traffic. *11th IFAC Symposium on Control in Transportation Systems*, (pp. 245-250). Delft.

Goverde, R. M., Corman, F., & D'Ariano, A. (2013). Railway line capacity consumption of different railway signalling systems under scheduled and disturbed conditions. *Journal of Rail Transport Planning & Management*, 78-94.

Hansen, I., & Pachl, J. (2014). *Railway Timetabling & Operations.* Hamburg: Eurailpress.

Jeroen. (2016, June 21). Planner. (S. Janssens, Interviewer)

Kees. (2016, May 23). Long term planner. (S. Janssens, Interviewer)

Lean Enterprise Institute. (2003). *Learning to See: value-stream mapping to create value and eliminate muda.* Brookline, MA.

Leclercq, L., Knoop, V. L., Marczak, F., & Hoogendoorn, S. P. (2014). Capacity Drops at Merges: New Analytical Investigations. *2014 IEEE 17th International Conference on Intelligent Transportation Systems (ITSC)*, (pp. 1129-1134). Qingdao.

Lentink, R. (2006). *Algorithmic Decision Support for Shunt Planning.* Rotterdam: Erasmus Research Institute of Management.

NedTrain. (2015, December). Overzicht van de NedTrain locaties.

NedTrain. (2016, June 12). Nachtoverstanden. Utrecht.

NedTrain. (n.d.). *Locaties.* http://www.nedtrain.nl/over-nedtrain/locaties/, Utrecht.

NS. (2016). Beeldbank. *http://nieuws.ns.nl/media-archief*. Utrecht.

NS. (2016, September). OBIEE. Utrecht.

NS.nl. (2016). *Aanschaf nieuwe intercity's.* Retrieved from http://nieuws.ns.nl/aanschaf-nieuwe-intercitys/

Olba, X. B., Daamen, W., Vellinga, T., & Hoogendoorn, S. (2014). An approach to port network capacity. *International Workshop on Next Generation Nautical Traffic Models.* Wuhan.

Peter. (2016, June 6). Manager Washing. (S. Janssens, Interviewer)

ProRail. (2016). Capaciteitsverdeling emplacementen . Utrecht.

ProRail. (2016, March). Perron- en spoorlengten. Utrecht.

Quaglietta, E., Corman, F., & Goverde, R. M. (2013). Impact of a stochastic and dynamic setting on the stability of railway dispatching solutions. *16th International IEEE Annual Conference on Intelligent Transportation Systems*, (pp. 1035-1050). The Hague.

Ramon. (2016, October 3). Manager Service Locations. (S. Janssens, Interviewer)

Robin. (2016, June 30). Mechanic. (S. Janssens, Interviewer)

Theo. (2016, July 4). Team leader. (S. Janssens, Interviewer)

Veeke, H., Lodewijks, G., & Ottjes, J. (2006). *Conceptual design of industrial systems: an approach to support collaboration.* London: Springer Verlag.

Veeke, H., Ottjes, J., & Lodewijks, G. (2008). *The Delft Systems Approach.* London: Springer Verlag.

Veld, J. i. (1971). Bedrijfsinformatie. *TED.*

Veysel. (2016, September 13). Team leader. (S. Janssens, Interviewer)

Wilson, M., Roos, N., Huisman, B., & Witteveen, C. (2011). International Conference on Through-life Engineering Services. *International Conference on Through-life Engineering Services.* Ghent.

Wu, C.-L., & Caves, R. E. (2002). Modelling of aircraft rotation in a multiple airport environment. *Transportation Research* (pp. 265-277). Elsevier.

Zeegers, S. (2016, September). Retrieved from Sporenplan: www.sporenplan.nl

Appendix 1 Definitions

Various terms are used throughout the study. In order to make sure those terms are interpreted correctly, these terms are defined here. Only the terms which are not common sense are involved here.

Aerial platform	platform at a height to work on the roofs of the train
Capacity	number of train units that can be serviced during a shift or a day at a Service Location
Carriage	chassis of the train
Check	daily check (DC): A-check, B-check, 24 hour-check
Demand	number of train units requested by NSR that have to be available for service
Maintenance Location	location where a train receives its planned structural maintenance
NedTrain	rolling stock maintenance company, full subsidiary of NS
NS Reizigers	NSR, passenger train operator, responsible for Dutch main railway lines
Pantograph	collector that takes the voltage from the overhead wires
Performance	number of realised (closed) work orders carried out during a shift on a single train unit
Process	all work orders carried out on a train unit during a shift
Production	number of work orders carried out during a shift on all train units at one Service Location
ProRail	rail infrastructure management company, responsible for the rail traffic control
Repair	corrective order (CO), preventive order (PO), direct repair (DHST), work from inspection (WUI)
Shunting	moving train units on short distance, not for revenue service (RANG)
Stabling	parking a train on dedicated track if it is not needed for revenue service
Supply	number of train units delivered to NSR available for service
Train	ride in revenue service for passenger transport
Train type	series of train units with the same characteristics, rolling stock type
Train unit	train set, fixed composition of coaches with driver cabins at both ends
Technical Centre	small maintenance building, including aerial platforms and working shaft
Throughput	number of train units serviced during a shift
Washing	external cleaning (PREIN): soap washing, head washing, oxalic washing
Working shaft	trench to work under the train or at the bogies

Appendix 2 Abbreviations

Abbreviation	**Service Location**
Ah	Arnhem
Amf	Amersfoort
Amr	Alkmaar
Aswpl	Amsterdam Zaanstraat
Bkh	Den Haag Binckhorst
Ddr	Dordrecht
Dv	Deventer
Ehv	Eindhoven
Ekz	Enkhuizen
Es	Enschede
Gn	Groningen
Gvc	Den Haag Centraal
Hdr	Den Helder
Hfd	Hoofddorp
Hgl	Hengelo
Hlm	Haarlem
Hn	Hoorn
Hrl	Heerlen
Ht	's Hertogenbosch
Ldd	Leidschendam
Lls	Lelystad
Lw	Leeuwarden
Mt	Maastricht
Nm	Nijmegen
Rsd	Roosendaal
Rtd	Rotterdam
Ut	Utrecht Centraal
Utctw	Utrecht Cartesiusweg
Utla	Utrecht Landstraat
Utoz	Utrecht OZ
Vl	Venlo
Vs	Vlissingen
Wgm	Watergraafsmeer
Zl	Zwolle
Zp	Zutphen

CO	Corrective order, if a crucial item is broken
DC	Daily check, A-, B- or 24 hour-check
DHST	Direct repair, to keep a train running
PO	Preventive order, to keep things working
PREIN	Washing order, soap washing, oxalic washing, head washing, in washing machine
RANG	Shunting order, for shunting within the Service Location
WUI	Work from inspection, repairs following from a check
WO	Work order of type CO, DC, DHST, PO, PREIN, RANG or WUI

Appendix 3 Service Location characteristics

A3.1 Layout characteristics

Data based on (ProRail, Perron- en spoorlengten, 2016) and (ProRail, Capaciteitsverdeling emplacementen , 2016)

Service Location	Layout type	Track length (m)	Capacity (number of coaches)	Number of tracks	Average track length (m)
Alkmaar	Shuffleboard	2.094	79	9	233
Amersfoort	Shuffleboard	5.033	190	12	419
Amsterdam Zaanstraat	Carousel	2.649	100	9	294
Arnhem	Shuffleboard	1.412	53	6	235
Den Haag Binckhorst	Shuffleboard	8.172	310	24	340
Den Helder	Shuffleboard	1.097	41	3	365
Deventer	Station	1.797	68	4	449
Dordrecht	Station	1.134	42	4	283
Eindhoven	Carousel	4.805	182	15	320
Enkhuizen	Shuffleboard	1.560	59	5	312
Enschede	Shuffleboard	1.994	75	6	332
Groningen	Shuffleboard	718	27	4	179
Haarlem	Shuffleboard	1.537	58	7	220
Heerlen	Station	4.978	188	9	553
's Hertogenbosch	Shuffleboard	855	32	5	171
Hengelo	Station	2.647	100	6	441
Hoofddorp	Carousel	3.466	131	9	385
Hoorn	Carousel	681	25	2	340
Leeuwarden	Station	1.869	70	5	374
Leidschendam	Shuffleboard	1.432	54	5	286
Lelystad	Carousel	3.001	113	6	500
Maastricht	Shuffleboard	5.214	197	13	401
Nijmegen	Shuffleboard	2.881	109	11	262
Roosendaal	Station	3.044	115	9	338
Rotterdam	Shuffleboard	4.247	160	16	265
Utrecht Cartesiusweg	Shuffleboard	5.069	192	21	241
Utrecht Landstraat	Shuffleboard	1.218	46	4	304
Utrecht OZ	Carousel	2.452	92	16	153
Venlo	Station	1.159	43	6	193
Vlissingen	Carousel	1.515	57	4	379
Watergraafsmeer	Carousel	5.774	218	14	412
Zutphen	Carousel	2952	111	6	492
Zwolle	Carousel	3.098	117	12	258

Capacity is expressed as number of coaches to be stabled, which is common use in Dutch railway business. This stabling capacity is calculated by dividing the total available track length by 26,4 meter, the UIC length of a coach.

A3.2 Equipment overview

Data based on (ProRail, Capaciteitsverdeling emplacementen , 2016) and (NedTrain, Overzicht van de NedTrain locaties, 2015).

Service Location	Washing machine	Technical Centre	Aerial platform	Working shaft
Alkmaar			X	X
Amersfoort			X	
Amsterdam Zaanstraat				
Arnhem	X		X	
Den Haag Binckhorst	X	X	X	X
Den Helder				
Deventer				
Dordrecht				
Eindhoven	X	X	X	
Enkhuizen				
Enschede	X			
Groningen			X	
Haarlem				
Heerlen				
's Hertogenbosch				
Hengelo		X	X	
Hoofddorp			X	
Hoorn				
Leeuwarden			X	X
Leidschendam				
Lelystad	X			
Maastricht	X		X	
Nijmegen	X		X	
Roosendaal			X	X
Rotterdam	X		X	X
Utrecht Cartesiusweg	X	X	X	
Utrecht Landstraat				
Utrecht OZ				
Venlo				
Vlissingen	X			
Watergraafsmeer	X	X	X	
Zutphen				
Zwolle	X		X	

Appendix 4 Standards for rolling stock handling

A4.1 Check standards

Check standards for all rolling stock types. The A-check for TRAXX+ICR is the 24 hour-check.

Type	A-check interval (days)	A-check standard time (min.)	B-check interval (days)	B-check standard time (min.)
SGM-II	12	38	2	9
SGM-III	12	44	2	10
SLT-IV	12	45	1	23
SLT-VI	12	51	1	27
MAT64	2	20	1	10
DM'90	2	60	1	22
DDAR	12	68	2	18
DDM	12	38	2	20
ICM-III	12	61	2	8
ICM-IV	12	64	2	11
VIRM-IV	12	54	2	12
VIRM-VI	12	60	2	14
DDZ-IV	12	76	2	15
DDZ-VI	12	90	2	18
TRAXX+ICR	1	180		

A4.2 Cleaning and washing standards

Cleaning and washing standards for all rolling stock types. Washing interval for soap is for both the sides and the head of the train, oxalic is the full train.

Type	Cleaning interval (days)	Cleaning standard time (min.)	Washing interval, soap (days)	Washing interval, oxalic (days)
SGM-II	1	10	7	63
SGM-III	1	15	7	63
SLT-IV	1	15	7	63
SLT-VI	1	20	7	63
MAT64	1	10	7	63
DM'90	1	10	7	63
DDAR	1	32	7	63
DDM	1	37	7	63
ICM-III	1	23	7	63
ICM-IV	1	30	7	63
VIRM-IV	1	37	7	63
VIRM-VI	1	56	7	63
DDZ-IV	1	49	7	63
DDZ-VI	1	56	7	63
TRAXX+ICR	1	80	7	63

Appendix 5 Work order variance

Average work order variance per Service Location, including standard deviations, and average number of train unit with at least one work unit, for all Mondays. The average ratio is the average number of work orders per train unit at that Service Location on a Monday.

Service Location	Average number of WO	Standard deviation in WO	Average number of train units	Standard deviation in train units	Average ratio
Ah	20,6	7,50	8,5	2,27	2,4
Amf	28,1	9,54	16,8	5,32	1,7
Amr	26,9	7,96	14,9	3,65	1,8
Aswpl	12,8	4,64	7,9	2,99	1,6
Bkh	106,2	19,73	49,2	5,65	2,2
Ddr	6,7	3,74	4,2	1,65	1,6
Dv	8,0	3,72	5,0	2,18	1,6
Ehv	68,0	17,23	31,4	5,12	2,2
Ekz	7,8	3,97	5,0	2,25	1,6
Es	23,7	8,23	14,2	4,36	1,7
Gn	22,3	5,95	12,3	3,10	1,8
Gvc	1,7	1,00	1,6	0,92	1,1
Hdr	7,2	2,82	4,4	1,50	1,6
Hfd	30,9	6,95	19,3	3,77	1,6
Hgl	33,5	12,76	12,2	2,74	2,8
Hlm	13,4	4,63	9,1	2,43	1,5
Hn	8,1	2,93	5,1	1,48	1,6
Hrl	11,4	5,14	8,0	3,21	1,4
Ht	6,2	3,22	4,2	1,70	1,5
Ldd	7,9	3,17	6,7	2,44	1,2
Lls	32,1	10,16	16,8	4,06	1,9
Lw	15,6	5,76	8,3	2,41	1,9
Mt	24,3	8,56	12,2	3,95	2,0
Nm	51,4	11,30	21,4	3,30	2,4
Rsd	42,8	10,02	22,9	3,20	1,9
Rtd	65,3	12,17	37,7	6,33	1,7
Ut	5,5	3,69	4,8	2,93	1,1
Utctw	44,2	13,70	20,4	5,10	2,2
Utla	13,8	5,87	8,0	2,34	1,7
Utoz	15,8	7,72	9,8	3,22	1,6
Vl	16,9	6,01	8,7	2,88	1,9
Vs	6,1	2,29	3,2	0,97	1,9
Wgm	92,2	16,90	57,1	8,17	1,6
Zl	88,2	12,91	35,3	4,33	2,5
Zp	18,1	5,37	9,4	2,67	1,9

Appendix 6 Example Work Order overview

Example of the visits of a train unit type ICM-III to Service Locations during a random month.

Work Type	Work Order Name	Location	Finish date	Finish time	Shift	Hours
DC	B-check	Hfd	01-05-2016	17:47	Late	0,5
CO	Cover missing trash bin	Es	01-05-2016	21:03	Late	0,0
DC	Washer fluid check	Es	01-05-2016	21:50	Late	0,2
WUI	Door damaged	Es	01-05-2016	21:50	Late	0,0
WUI	Information display defect	Es	01-05-2016	21:50	Late	0,1
CO	Rubber missing door	Es	01-05-2016	22:15	Late	0,0
PREIN	Soap washing	Es	02-05-2016	16:20	Late	0,0
PO	Typhoon check	Es	02-05-2016	17:31	Late	0,0
DC	B-check	Hfd	03-05-2016	14:00	Early	0,1
DC	A-check	Es	05-05-2016	20:16	Late	1,2
CO	Brakes stuck	Es	05-05-2016	20:16	Late	2,6
DC	B-check	Hfd	07-05-2016	14:14	Early	1,0
PREIN	Oxalic washing	Bkh	09-05-2016	04:45	Night	0,4
PO	Typhoon check	Hfd	09-05-2016	10:45	Early	0,0
DC	B-check	Hfd	09-05-2016	20:45	Late	0,6
CO	Safety system logout	Ehv	12-05-2016	01:05	Night	0,4
DC	B-check	Hfd	13-05-2016	01:23	Night	0,4
DC	Washer fluid check	Hfd	13-05-2016	01:23	Night	0,0
DC	B-check	Hrl	14-05-2016	09:02	Early	0,3
DC	B-check	Ehv	16-05-2016	21:56	Late	0,6
CO	Light defect cabin	Ehv	17-05-2016	13:56	Early	0,1
PO	Coupler cleaning	Ehv	17-05-2016	19:56	Late	0,1
DC	A-check	Ehv	17-05-2016	19:56	Late	1,1
DC	Washer fluid check	Ehv	17-05-2016	19:57	Late	0,0
WUI	Washer fluid shortage	Ehv	18-05-2016	13:34	Early	0,0
DHST	Brakes malfunction	Vl	19-05-2016	09:03	Early	0,2
DC	B-check	Vl	20-05-2016	02:11	Night	0,9
CO	Rubber missing door	Vl	20-05-2016	03:22	Night	0,0
WUI	Rubber loose door	Vl	20-05-2016	03:22	Night	0,2
CO	Rubber missing door	Vl	20-05-2016	04:23	Night	0,5
CO	Light defect interior	Vl	20-05-2016	05:20	Night	0,2
CO	Cover missing trash bin	Vl	20-05-2016	06:37	Night	1,2
DHST	Brake test	Vl	21-05-2016	17:08	Late	0,3
DC	B-check	Vl	21-05-2016	17:08	Late	0,7
DC	B-check	Hrl	24-05-2016	02:47	Night	0,3
DC	Washer fluid check	Hrl	24-05-2016	03:38	Night	0,1
DC	B-check	Hfd	25-05-2016	13:48	Early	0,2
CO	First aid kit missing	Ehv	25-05-2016	21:30	Late	0,6
DC	B-check	Bkh	27-05-2016	21:09	Late	0,5
DC	Washer fluid check	Bkh	29-05-2016	13:21	Early	0,0
DC	A-check	Bkh	29-05-2016	13:21	Early	1,0
DC	B-check	Vl	31-05-2016	14:00	Early	0,4

Appendix 7 Occupation night shift

Number of coaches present at the Service Locations on all weekdays in June and July 2016, based on realised work orders during the night shift.

SL=Service Location, Avg=average number of coaches present, st.dev.=standard deviation number of coaches present, plan=planned occupation, factor=plan/avg

Mondays

SL	04-07	06-06	11-07	13-06	18-07	20-06	25-07	27-06	Avg	St. dev.	stdev/ avg	Plan	Factor
Ah	22	34	21	20	35	27	28	24	26	5,4	0,20	36	1,36
Amf	42	51	25	35	49	26	35	19	35	10,8	0,31	87	2,47
Amr	32	23	29	10	29	18	10	9	20	8,9	0,45	53	2,65
Aswpl	30	23	24	17	24	38	11	18	23	7,8	0,34	41	1,77
Bkh	158	101	131	140	158	99	128	105	128	22,5	0,18	191	1,50
Ddr	20	24	18		18	28	22	22	22	3,3	0,15	30	1,38
Ehv	45	58	90	75	80	63	84	81	72	14,3	0,20	106	1,47
Ekz	6	22	10	14		4	12	10	11	5,4	0,49	51	4,58
Es	23	22	40	34	29	13	31	23	27	7,8	0,29	59	2,20
Gn	36	20	13	17	19	39	13	22	22	9,2	0,41	52	2,32
Hdr	17	20	17	18	22	29	12	23	20	4,7	0,24	43	2,18
Hfd	53	50	62	57	33	49	78	33	52	13,9	0,27	92	1,77
Hgl	10	16	20	18	27	17	18	28	19	5,5	0,29	45	2,34
Hlm	32	31	25	22	26	19	21	13	24	5,9	0,25	48	2,03
Hn	14	19	24	7	14	15	7	16	15	5,3	0,37	30	2,07
Hrl	11	11		18	14	4	12	7	11	4,2	0,38	66	6,00
Ht	12		20	18	19	28	22		20	4,8	0,24	35	1,76
Ldd	38	22	18	35	39	25	41	40	32	8,5	0,26	100	3,10
Lls	58	62	86	52	48	54	72	48	60	12,3	0,21	88	1,47
Lw	18	6	8	14		11	13	20	13	4,7	0,36	32	2,49
Mt	8	16	34	9	16	19	29	21	19	8,4	0,44	79	4,16
Nm	54	50	64	43	39	54	62	39	51	9,1	0,18	85	1,68
Rsd	72	50	72	56	66	44	48	54	58	10,2	0,18	138	2,39
Rtd	65	45	50	58	60	67	69	69	64	8,4	0,13	113	1,77
Utctw	48	13	82	62	64	38	80	80	58	22,7	0,39	100	1,71
Utla	26	26	29	34	15	26	29	37	28	6,1	0,22	40	1,44
Vl	17	24	17	14	23	17	23	17	19	3,5	0,18	27	1,42
Vs	12	16	10	20	10	26	16	16	16	5,0	0,32	20	1,27
Wgm	103	98	174	109	172	136	154	121	133	28,5	0,21	157	1,18
Zl	96	82	87	65	83	78	86	118	87	14,4	0,17	122	1,40
Zp	24	23	17	14	34	24	14	16	21	6,4	0,31	66	3,18
													2,23

Tuesdays

SL	05-07	07-06	12-07	14-06	19-07	21-06	26-07	28-06	Avg	St. dev.	stdev /avg	Plan	Factor
Ah	26	20	56	36	28	6	14	30	27	14,1	0,52	75	2,78
Amf	49	48	36	39	33	36	40	73	44	12,1	0,27	106	2,40
Amr	24	40	21	19	32	11	21	23	24	8,2	0,34	50	2,09
Aswpl	10	43	44	30	30	26	24	28	29	10,1	0,34	35	1,19
Bkh	132	122	176	172	149	138	180	147	152	20,3	0,13	228	1,50
Ddr	22	24	16	28	46	20	28	26	26	8,4	0,32	68	2,59
Ehv	92	57	90	45	69	67	98	62	73	17,6	0,24	123	1,70
Ekz	20	32	23	13	19	13	12	10	18	6,9	0,39	51	2,87
Es	53	26	39	34	34	39	34	24	35	8,4	0,24	58	1,64
Gn	53	22	23	21	37	41	34	23	32	10,8	0,34	52	1,64
Hdr	40	10	12		18	16	14	6	17	10,2	0,62	40	2,41
Hfd	70	57	54	54	46	47	40	52	53	8,4	0,16	90	1,71
Hgl	32	23	29	24	20	12	24	20	23	5,7	0,25	43	1,87
Hlm	28	27	21	15	30	17	25	22	23	5,0	0,22	46	1,99
Hn	6	8	23	24	17	20	22	19	17	6,4	0,37	34	1,96
Hrl	17	22	19	22	11	3	7	14	14	6,5	0,45	59	4,10
Ht	20	16	27	18	18	14	16	10	17	4,6	0,27	35	2,01
Ldd	53	46	35	34	25	28	30	52	38	10,3	0,27	103	2,72
Lls	58	66	60	58	42	54	62	68	59	7,5	0,13	106	1,81
Lw	11	13	21	23	31	11	32	14	20	8,1	0,41	42	2,15
Mt	59	41	25	26	32	12	13	40	31	14,6	0,47	79	2,55
Nm	96	66	46	61	30	66	40	57	58	18,8	0,33	84	1,45
Rsd	86	74	84	50	62	74	70	56	70	11,9	0,17	134	1,93
Rtd	60	90	38	62	49	69	75	55	62	15,0	0,24	99	1,59
Utctw	73	80	90	53	73	71	64	46	69	13,3	0,19	100	1,45
Utla	39	34	32	25	14	42	23	34	30	8,6	0,28	40	1,32
Vl	43	27	17	18	20	16	19	19	22	8,4	0,38	27	1,21
Vs	32	20	26	20	16	10	12	16	19	6,8	0,36	20	1,05
Wgm	151	140	190	133	161	104	88	113	135	30,9	0,23	145	1,07
Zl	132	91	86	90	96	74	102	85	95	16,1	0,17	123	1,30
Zp	38	34	34	25	26	29	37	32	32	4,5	0,14	56	1,76
													1,93

Wednesdays

SL	06-07	08-06	13-07	15-06	20-07	22-06	27-07	29-06	Avg	St. dev.	stdev / avg	Plan	Factor
Ah	32	30	35	18	32	30	24	28	29	5,0	0,18	40	1,40
Amf	35	44	52	42	38	33	43	38	41	5,6	0,14	79	1,94
Amr	22	27	34	6	21	20	11	47	24	12,0	0,51	50	2,13
Aswpl	26	22	17	11	8	17	9	12	15	6,0	0,39	36	2,36
Bkh	169	116	128	137	139	166	162	124	143	19,1	0,13	231	1,62
Ddr	14	24	18	16	22	8	16	24	18	5,1	0,29	26	1,46
Ehv	67	48	56	61	52	60	89	67	63	11,8	0,19	101	1,62
Ekz	12	24	30	23	22	23	9	12	19	6,9	0,36	51	2,63
Es	40	25	30	29	34	42	61	30	36	10,7	0,30	61	1,68
Gn	11	25	23	33	31	19	28	33	25	7,1	0,28	52	2,05
Hdr	12	6	26	14	6	24	22	14	16	7,3	0,47	40	2,58
Hfd	48	58	51	29	50	37	66	48	48	10,7	0,22	84	1,74
Hgl	12	27	38	40	12	26	8	14	22	11,6	0,53	43	1,94
Hlm	36	14	17	20	24	22	39	17	24	8,6	0,36	53	2,24
Hn	16	14	25	8	24	16	26	17	18	5,8	0,32	34	1,86
Hrl	12	14	5	23	15	13	12	28	15	6,7	0,44	62	4,07
Ht	10	15	20	4	10	20	12	18	14	5,3	0,39	35	2,57
Ldd	36	47	20	19	31	32	16	23	28	9,8	0,35	107	3,82
Lls	56	56	86	52	38	44	38	58	54	14,4	0,27	110	2,06
Lw	17	20	13	21	26	13	13	12	17	4,7	0,28	38	2,25
Mt	47	16	41	26	33	20	31	13	28	11,2	0,39	79	2,78
Nm	68	44	70	63	18	56	76	38	54	18,3	0,34	84	1,55
Rsd	50	82	76	94	62	70	47	60	68	15,0	0,22	138	2,04
Rtd	77	56	80	74	57	90	77	63	72	11,2	0,16	117	1,63
Utctw	47	73	60	77	72	71	80	96	72	13,4	0,19	100	1,39
Utla	38	31	24	22	16	37	32	26	28	7,1	0,25	50	1,77
Vl	17	18	15	22	23	17	21	28	20	4,0	0,20	31	1,54
Vs	18	18	20	6	14	10	22	10	15	5,3	0,36	20	1,36
Wgm	165	104	174	105	164	120	145	143	140	25,8	0,18	139	0,99
Zl	92	102	111	117	91	68	88	84	94	14,6	0,16	126	1,34
Zp	47	27	20	21	33	34	42	33	32	8,8	0,27	56	1,74
													2,05

Thursdays

SL	07-07	09-06	14-07	16-06	21-07	23-06	28-07	30-06	Avg	St. dev.	stdev /avg	Plan	Facto r
Ah	34	28	21	20	32	20	34	14	25	7,1	0,28	44	1,73
Amf	54	47	55	49	63	45	25	37	47	11,0	0,23	95	2,03
Amr	77	38	44	42	29	17	44	34	41	16,2	0,40	58	1,43
Aswpl	7	31	35	19	22	25	12	28	22	8,9	0,40	43	1,92
Bkh	178	168	199	129	121	135	145	91	146	32,4	0,22	225	1,54
Ddr	20	10	18	24	28	16	26	22	21	5,5	0,27	30	1,46
Ehv	96	82	64	88	110	43	76	77	80	18,9	0,24	112	1,41
Ekz	9	24	3	20	16	8	9	13	13	6,5	0,51	51	4,00
Es	29	24	34	43	16	30	26	36	30	7,7	0,26	58	1,95
Gn	45	33	29	31	22	27	32	13	29	8,6	0,30	54	1,86
Hdr	10	14	22	14	22	16	26	22	18	5,1	0,28	44	2,41
Hfd	47	30	39	43	52	49	60	67	48	10,9	0,23	90	1,86
Hgl	25	27	42	25	27	18	28	21	27	6,6	0,25	43	1,62
Hlm	7	19	22	8	24	16	22	19	17	6,0	0,35	46	2,69
Hn	12	16	19	21	11	16	7	11	14	4,4	0,31	30	2,12
Hrl	13	3	15	7	11	12	14	7	10	3,9	0,38	59	5,76
Ht	11	4	31	16	22	30	16	18	19	8,5	0,46	35	1,89
Ldd	39	32	42	21	32	31	19	36	32	7,5	0,24	107	3,40
Lls	72	50	62	44	24	20	60	62	49	17,6	0,36	88	1,79
Lw	30	18	10	12	21	12	16	9	16	6,5	0,41	36	2,25
Mt	69	25	19	25	33	23	36	23	32	15,1	0,48	79	2,50
Nm	86	84	65	44	34	55	24	44	55	21,0	0,39	80	1,47
Rsd	68	30	66	64	78	58	70	44	60	14,6	0,24	134	2,24
Rtd	91	84	94	109	70	103	64	30	81	23,9	0,30	100	1,24
Utctw	71	93	51	87	94	76	106	108	86	17,8	0,21	100	1,17
Utla	39	18	46	24	38	48	43	30	36	10,1	0,28	50	1,40
Vl	29	24	22	4	14	13	19	22	18	7,3	0,40	31	1,69
Vs	34	18	6	12	14	20	26	14	18	8,2	0,45	20	1,11
Wgm	150	143	172	131	181	119	196	115	151	27,7	0,18	154	1,02
Zl	104	73	98	57	113	89	116	106	95	19,2	0,20	125	1,32
Zp	52	31	24	49	47	14	41	34	37	12,4	0,34	57	1,56
													1,99

Fridays

SL	01-07	08-07	10-0	15-07	17-06	22-07	24-06	29-07	Avg	St. dev.	stdev /avg	plan	facto r
Ah	26	42	26	55	26	53	40	34	38	11,1	0,29	44	1,17
Amf	54	52	70	64	47	66	43	43	55	9,9	0,18	90	1,64
Amr	33	34	41	33	27	54	33	41	37	7,7	0,21	62	1,68
Aswpl	16	9	38	30	13	33	18	22	22	9,6	0,43	45	2,01
Bkh	130	163	140	121	128	99	156	132	134	18,7	0,14	219	1,64
Ddr	36	36	13	26	20	30	22	14	25	8,4	0,34	36	1,46
Ehv	81	91	44	86	44	66	79	90	73	18,1	0,25	115	1,58
Ekz	20	13		9	30	7	10	19	15	7,5	0,49	51	3,31
Es	45	31	17	40	26	28	35	16	30	9,6	0,32	58	1,95
Gn	25	11	31	13	40	31	30	25	26	9,1	0,35	52	2,02
Hdr	32	6	16	6	6	24	14	30	17	10,1	0,60	42	2,51
Hfd	39	64	49	46	53	60	62	57	54	8,1	0,15	89	1,66
Hgl	27	19	23	22	19	16	32	28	23	5,0	0,22	42	1,81
Hlm		23	16	11	8	10	17	19	15	5,0	0,34	46	3,10
Hn	18	10	12	13	13	12	8	4	11	3,8	0,34	30	2,67
Hrl	6	17	22	24	16	12	9		15	6,1	0,40	63	4,16
Ht	18	26	14	20	18	20	10	26	19	5,1	0,27	34	1,79
Ldd	22	36	38	14	20	26	16	10	23	9,4	0,41	101	4,44
Lls	38	48	56	50	70	58	20	46	48	13,9	0,29	86	1,78
Lw	8	4	23	24	16	26	15	24	18	7,6	0,44	40	2,29
Mt	13	36	21	30	47	36	30	20	29	10,2	0,35	73	2,51
Nm	68	78	34	70	59	32	62	26	54	18,7	0,35	78	1,45
Rsd	64	84	36	36	52	104	56	92	66	23,8	0,36	134	2,05
Rtd	68	55	123	61	77	69	49	36	67	24,3	0,36	92	1,37
Utctw	80	68	38	43	58	68	64	87	63	15,7	0,25	100	1,58
Utla	41	39	26	20	22	20	32	17	27	8,6	0,32	50	1,84
Vl	16	20	9	23	34	26	21	17	21	6,9	0,33	31	1,49
Vs	36	18	16	24		20	20	22	22	6,1	0,27	20	0,90
Wgm	183	184	164	208	144	134	147	133	162	25,5	0,16	144	0,89
Zl	100	93	84	101	96	68	73	109	91	13,4	0,15	122	1,35
Zp	33	40	20	19	17	29	27	38	28	8,2	0,29	57	2,04
													2,00

Appendix 8 Model results

A8.1 Model results

Model results from data set (all realised work orders August 2015 – July 2016).

Service Location	Number of tracks	Total number of realised WO	Expected number of WO	Difference expected WO vs. Total realised WO (%)
Ah	6	9.134	10.366	12
Amf	12	11.213	10.881	-3
Amr	9	10.219	8.142	-26
Aswpl	9	4.393	12.037	64
Bkh	24	37.441	35.732	-5
Ddr	4	2.718	3.747	27
Dv	4	2.066	3.747	45
Ehv	15	26.652	39.633	33
Ekz	5	3.634	3.862	6
Es	6	10.267	9.231	-11
Gn	4	8.699	3.077	-183
Hdr	3	2.416	2.292	-5
Hfd	9	11.436	12.037	5
Hgl	6	13.637	5.931	-130
Hlm	7	4.707	6.317	25
Hn	2	2.680	1.563	-71
Hrl	9	5.631	7.918	29
Ht	5	2.468	4.491	45
Ldd	5	2.933	3.862	24
Lls	6	11.807	11.376	-4
Lw	5	5.505	4.162	-32
Mt	13	10.628	17.474	39
Nm	11	18.324	17.212	-6
Rsd	9	15.505	9.207	-68
Rtd	16	24.907	24.778	-1
Utctw	21	16.037	30.904	48
Utla	4	5.641	3.578	-58
Utoz	16	5.069	27.062	81
Vl	6	6.261	5.101	-23
Vs	4	2.255	3.943	43
Wgm	14	32.786	37.133	12
Zl	12	31.375	30.694	-2
Zp	6	7.673	5.598	-37

A8.2 Basic and final model comparison

Model results from final model (paragraph 5.2.5) versus model results from basic regression model, including differences from realised number of work orders.

Service Location	Number of tracks	Total number of realised WO	Expected WO basic model	Difference basic model (%)	Expected WO final model	Difference final model (%)
Ah	6	9.134	6.755	-35	10.366	12
Amf	12	11.213	13.601	18	10.881	-3
Amr	9	10.219	10.178	0	8.142	-26
Aswpl	9	4.393	15.046	71	12.037	64
Bkh	24	37.441	27.293	-37	35.732	-5
Ddr	4	2.718	4.684	42	3.747	27
Dv	4	2.066	4.684	56	3.747	45
Ehv	15	26.652	31.144	14	39.633	33
Ekz	5	3.634	5.614	35	3.862	6
Es	6	10.267	6.755	-52	9.231	-11
Gn	4	8.699	4.473	-94	3.077	-183
Hdr	3	2.416	3.332	27	2.292	-5
Hfd	9	11.436	15.046	24	12.037	5
Hgl	6	13.637	7.414	-84	5.931	-130
Hlm	7	4.707	7.896	40	6.317	25
Hn	2	2.680	-3.735	172	1.563	-71
Hrl	9	5.631	11.509	51	7.918	29
Ht	5	2.468	5.614	56	4.491	45
Ldd	5	2.933	5.614	48	3.862	24
Lls	6	11.807	6.997	-69	11.376	-4
Lw	5	5.505	6.049	9	4.162	-32
Mt	13	10.628	14.742	28	17474	39
Nm	11	18.324	12.460	-47	17..212	-6
Rsd	9	15.505	11.509	-35	9.207	-68
Rtd	16	24.907	18.165	-37	24.778	-1
Utctw	21	16.037	23.870	33	30.904	48
Utla	4	5.641	4.473	-26	3.578	-58
Utoz	16	5.069	33.827	85	27.062	81
Vl	6	6.261	7.414	16	5.101	-23
Vs	4	2.255	1.631	-38	3.943	43
Wgm	14	32.786	28.461	-15	37.133	12
Zl	12	31.375	23.095	-36	30.694	-2
Zp	6	7.673	6.997	-10	5.598	-37

A8.3 Test model results

Model results from test data set (all realised work orders February 2015 – July 2015).

Service Location	Number of tracks	Total number of realised WO	Expected WO half year	Difference expected WO vs. Total realised WO (%)
Ah	6	4.330	5.183	16
Amf	12	5.501	5.441	-1
Amr	9	5.795	4.071	-42
Aswpl	9	2.764	6.019	54
Bkh	24	19.804	17.866	-11
Ddr	4	1.860	1.874	1
Dv	4	543	1.874	71
Ehv	15	14.739	19.817	26
Ekz	5	1.814	1.931	6
Es	6	4.696	4.616	-2
Gn	4	4.465	1.539	-190
Hdr	3	1.182	1.146	-3
Hfd	9	6.742	6.019	-12
Hgl	6	7.255	2.966	-145
Hlm	7	2.440	3.159	23
Hn	2	2.097	782	-168
Hrl	9	3.260	3.959	18
Ht	5	1.044	2.246	54
Ldd	5	2.099	1.931	-9
Lls	6	7.203	5.688	-27
Lw	5	2.945	2.081	-42
Mt	13	6.095	8.737	30
Nm	11	9.238	8.606	-7
Rsd	9	10.229	4.604	-122
Rtd	16	12.879	12.389	-4
Utctw	21	9.753	15.452	37
Utla	4	3.481	1.789	-95
Utoz	16	4.281	13.531	68
Vl	6	3.159	2.551	-24
Vs	4	1.140	1.972	42
Wgm	14	16.000	18.567	14
Zl	12	15.515	15.347	-1
Zp	6	4.137	2.799	-48

.

Appendix 9 Interview reports

Small reports of all interviews done during the project at multiple Service Locations are made. Those interview reports can be found in this appendix.

Kees, 23 May 2016
Kees (Kees, 2016) is long term planner for the Randstad Zuid region. He explained his task and shows the location Utrecht Cartesiusweg, where the interview was held.

Kees explained the long term planning process, from the beginning of a new schedule until the long term planning for a Service Location. He indicated the main points of interest for a Service Location: the stabling capacity and the infrastructure occupation. Infrastructural restrictions can be of major influence on the capacity of a location. Utrecht Cartesiusweg is a very good example. Although it is close to Utrecht Centraal, only 40% of the total capacity can be used. This is caused by the layout of the tracks around the station: the majority of the tracks simply is not connected to the switches towards Utrecht Cartesiusweg. Other locations within the Randstad Zuid region also have this kind of restrictions.

Kees also showed the location Cartesiusweg, where at the time one train unit was present. He pointed out that also within the location the layout has many restrictions. All stabling tracks are connected with only two switches. The succession times between two train movements take a lot of capacity. If a train is in the washing machine, one of the switches is occupied for a longer time, resulting in half of the location being locked. The interview with Kees was a good introduction to the Service Locations, also because Kees is working for NS since several decades.

Chris, 30 May 2016 and 25 August 2016
Chris (Chris, 2016) is team leader at the Service Location Rotterdam and works for NedTrain since 35 years. He explained the tasks at the Service Location and gave a tour around the location.

Rotterdam is one of the biggest Service Locations. The location was visited multiple times, amongst others during a full late shift and a full night shift. By doing this, good insight in the process was gathered. Chris is team leader and therefore he is part of all communication during a shift. During the shift, he is in contact with the mechanics in Rotterdam on the one hand and the planner (at Utrecht Cartesiusweg) on the other hand.

Chris gave a tour across the location Rotterdam, indicating the shortcomings of the layout. Rotterdam consists of four parts, all connected by the station. Staff is not allowed to cross the main tracks, resulting in long walking times. Especially if spare parts are needed, walking times can be very long, as these are located only at the northern part of the Service Location. For shunting, the location is well equipped, since all tracks are connected to each other at both ends. The washing machine is unique, since the train stands still while the machine runs along the train. NedTrain staff does not have any work on the washing machine, all work is done by the train drivers.

Rotterdam is one of the two Service Locations welcoming the IC Direct trains: every night, three train compositions are stabled and checked. A check mechanic also showed the 24 hour-check on the TRAXX-locomotive and the ICR-coaches. The late shift is quietly, while the night shift is more busy. Especially at the beginning of the train service in the morning, many trains leave the stabling tracks, around 5 a.m.. During the night, many checks are done. In order to do repairs, two mechanics are available. One of those mechanics even drove to the Maintenance Location in Leidschendam to pick up spare parts. This turned out to be more efficient than calling a courier. In the night shift, one mechanic drives to Dordrecht, to do the checks and small repairs on the train units which are stabled there. There is no regular staff located in Dordrecht during the day.

The rail traffic controller is in the same room as the team leader in Rotterdam. The participation of the rail traffic controller in the total process turned out to be very limited: he does not act proactively, only reactive. Most of the time, he only does the operations the planner planned. The requests for a shunting movement come automatically, the driver notifies what he wants and the traffic controller only checks whether that can be done safely at that moment. Only in exceptional cases, if sudden disturbances pop up, the traffic controller will change the track arrangement. By doing that, the right formation of the train units will be disturbed.

Peter, 6 June 2016

Peter (Peter, 2016) is manager washing of NedTrain. He explained the principle of washing and the challenges and recent developments for the washing process.

Peter is responsible for both washing and graffiti removal. He explained how washing basically works and what the constraints are. The washing production is currently far beyond standards, due to various causes. The final goal within regarding washing is a maximal deviation from the plan of 3 days. In the near future, train units will not be released for revenue service before the washing is done. Some train unit series need extra attention, especially ICM with the high positioned windows of the driver cabin. Those windows are washed separately. Getting a train to a washing machine sometimes really is a challenge. Due to the timetable, 14% of the train units does not even visit a Service Location with a washing machine within two months. To take these train units clean requires special attention.

Some washing machines perform much better than others. This partly depends on the location of the machine, for example if it is not centrally within the network. This holds for example for Vlissingen, that is on the far end of the network. However, this washing machine is the only one which can still be used when it is very cold, as temperatures are usually higher that close to the sea. Some other washing machines perform worse although they are well located, for example Rotterdam. The location within the Service Location of that machine is not optimal.

As a new development, the fast washing machines are mentioned. One is already installed in Enschede and a new one will be operational as from September at Grote Binckhorst. These machines wash in several minutes but can only do soap washings, not head and oxalic washings. Another development is season bounded washing. During summertime, trains become less dirty than during wintertime. As a result of that, they should be washed less in the summer. A program to set this up is currently being developed.

Jeroen, 21 June 2016

Jeroen (Jeroen, 2016) is planner for the Randstad Zuid region. He can plan for all Service Locations within the region. The interview was held at Utrecht Cartesiusweg, where all planners for this region are located. There is a constant occupation of planners for all big Service Locations, in Utrecht this is for Utrecht itself, as well as Rotterdam and Den Haag Binckhorst.

Jeroen showed his work in practice. He explains how the work comes in and is executed throughout a shift. He was planning for Utrecht Cartesiusweg during the visit. The planning is made for the upcoming shifts, based on information from the management system about the actual circulation planning for the train units. Besides that, the planner is constantly in contact with the team leader of the specific locations, to create new work orders or to fix other problems. New work orders are made if a mechanic ascertains a defect during the check, for example. Also work orders finished by mechanics are closed by the planner.

Planning the work includes the basic arrangement of the train units throughout the Service Location. All basic movements are considered on beforehand, and checked in Excel, what does not seem to be the optimal program to do so. Also all work is planned: all train units that will visit the Service Location are inserted in the management system (Maximo). By giving in a train units number, a long list of work is shown, also work not relevant for the Service Location. The tasks which have to be done have to be selected by hand, since the program does not indicate automatically which tasks are relevant or close to expire date. This results in an approach which is sensitive for mistakes, since tasks can be ignored by coincident. Ignoring the tasks can result in trains which are not available for revenue service. All tasks found are listed and filled in in the planning form. This form is send to the team leader at the beginning of each shift. The team leader is in the end responsible to assign the work to the mechanics.

Robin, 30 June 2016
Robin (Robin, 2016) is mechanic in the Randstad Noord region. He explains the tasks of a mechanic and demonstrates the work. He also shows the Service Location Hoofddorp.

Hoofddorp is not a particularly big Service Location, but it is located close to the station of Schiphol Airport, where some train series end. Therefore, a constant flow of train units comes into the location. The layout is simple, with several long tracks, which are connected to each other at both ends. Some parts of the location are recently rebuild, resulting in a bigger capacity.

Robin shows what has to be done for both an A- and a B-check. The start is at the office, with the tablet, that shows the work order, including possible defects. The B-check is a more simple check. The main points of interest are the bogies and the pantographs. A general walk along the train unit and a quick look in both cabins conclude this check. The A-check is more comprehensive. The A-check particularly also includes a functional test, including brakes check and a full test of all systems. Also the control panel is viewed and all safety equipment is checked. An inspection of the interior of the train is done by walking thru it. Outside, also the bogies and pantograph are expected, equal to the B-check.

After the check, the results have to be reported. All defects discovered are reported individually. These defects are checked by the planner and translated to a work order. Also the check is closed by the planner. By closing the work order for the check, a new work order is automatically generated, based on the closing time.

Theo, 4 July 2016
Theo (Theo, 2016) is team leader at the Service Location Den Haag Binckhorst, the biggest Service Location. He gave an elaborate tour throughout the location, shows the Technical Centre and explains his work including the challenges.

Binckhorst is a special Service Location, since it is split into two parts, divided by the mainline Den Haag Centraal-Zoetermeer. The northern part is called the Kleine Binckhorst, the southern part is the Grote Binckhorst. The Kleine Binckhorst is arranged as a carousel layout, while the Grote Binckhorst is a shuffleboard layout. The washing machine is located at the Kleine Binckhorst, although a new fast washing machine is being built during the visit at the Grote Binckhorst (and not taken into account for this study). The Grote Binckhorst has a newly developed Technical Centre since 2015.

Theo states that the capacity of Binckhorst is almost reached: during some nights, the logistic puzzle cannot be fully solved. This is mainly caused by the central switches, connecting all stabling tracks at the Grote Binckhorst. For the Kleine Binckhorst, the work order is the key element. After arrival, most train units are either checked first, then shunted to the cleaning platform, shunted to the washing machine and then reversely shunted to a stabling track. At that track, the train unit can also be checked if that was not done before cleaning. In this process, there are many dependencies, since train units sometimes have to wait before another train unit is finished at one of the service stations. In order to accelerate the process, sometimes the tasks are done in reverse direction. This may result in the opposite effect, that the process stagnates.

The size of Binckhorst is also challenging, especially for shunt drivers. Since there is a limited amount of drivers available and the walking distances are long. This results in loss of time. Also waiting during shunt movements, due to safety regulations, cause delays.

Veysel, 13 September 2016
Veysel (Veysel, 2016) is team leader at the Service Location Amsterdam Watergraafsmeer. He showed the location and explains what his task is.

Watergraafsmeer is the biggest Service Location in the Amsterdam area. It is special, since all IC Direct trains turn here and are cleaned here. Also the international trains, Thalys and ICE, are cleaned here, on a separate part of the Service Location, not taken into account in this study. Besides that, also the Maintenance Location Watergraafsmeer is located here, but somewhat separated from the Service Location. Watergraafsmeer is a very long location, with various separate parts, including a washing machine and a Technical Centre. Mechanics working at this location are also responsible for the Dijksgracht, a stabling area located near Amsterdam Centraal. During the day, it is common use at Watergraafsmeer to do as many repairs as possible on the present train units with a group of mechanics.

The washing machine is somewhat special, since train units cannot drive under own power into it. They have to be shunted with a dedicated shunt locomotive, causing extra work. The central switches connecting the multiple parts of the Service Location are highly occupied. At the western part of Watergraafsmeer, train are cleaned at a cleaning platform. If that is done, a train unit cannot undergo a check, since one side of the train cannot be inspected. This is a challenge, since the train units have to be shunted to do the checks.

Ramon, 3 October 2016
Ramon (Ramon, 2016) is manager of the Service Locations in the South region. His location is Eindhoven, where he also showed the Technical Centre and the rest of the location.

Service Locations within the South region are Eindhoven, Roosendaal, Heerlen, Maastricht, Venlo and ’s Hertogenbosch, of which the latter two are not continuously occupied. Eindhoven is already the most important location in terms of number of work orders and number of stabled train units. The location will become more important in the near future, since all Sprinter train series will terminate in Eindhoven, as Arriva will take over service in the Limburg province from December 2016.

The Technical Centre is frequently used, for example for retrofitting wheels at the wheel lathe. This tasks sometimes disturb regular tasks, since the trains sometimes occupy parts of the tracks. This is also the case if the washing machine is used. This unplanned work causes extra logistic work. Most of the planning staff does not have a specific logistic background, what may result in making wrong decisions. Besides that, the NSR planner has other interests than NedTrain, what can influence the process. Eindhoven has also special regulations for working: only during a limited time frame, mechanics are permitted to work at certain parts of the area. This strongly influences the shunt planning: train units with dedicated work are forced to go to areas without regulations.

Eindhoven has some very long tracks, which are partly managed by ProRail and partly by NedTrain. The possibilities to work on train units which are stabled between other train units at those tracks are very much limited. This also require accurate planning. This also holds for the staff: they work for the regular tasks at the Service Location, but if needed also in the Technical Centre. Therefore, Eindhoven is a very interesting Service Location.