# Master Thesis

*Developing a model for estimating road capacity values of weaving sections*


dat. mobility

UNIVERSITEIT TWENTE.

*01-10-2018*
Jurre Janssen, s1504002
Supervisor: Pr. Dr. Ir. E.C. van Berkum
Supervisor: Dr. Ir. L. Wismans
Company: DAT. Mobility
Supervisor Company: Ir. L. Brederode

| | |
|---|---|
| TITLE | Developing a model for estimating road capacity values for weaving sections. |
| DESCRIPTION | This research focuses on the development of a model for estimating road capacity values for weaving sections. It also elaborates on the process and variables of the (development of the) model. |
| TYPE | Master Thesis |
| VERSION | Final version |
| DATE OF PUBLICATION | 1 October 2018 |
| LOCATION | Deventer, the Netherlands |
| PAGES | 90 |


| | |
|---|---|
| AUTHOR | Jurre Janssen |
| STUDENT NUMBER | s1504002 |
| CONTACT | j.c.janssen@student.utwente.nl |
| EDUCATIONAL INSTITUTION | University of Twente |
| 1st SUPERVISOR | Dr. Ir. L. Wismans |
| 2nd SUPERVISOR | Prof. dr. ir. E.C. van Berkum |
| COMPANY | *DAT. Mobility* |
| LOCATION | Deventer, The Netherlands |
| SUPERVISOR | Ir. L. Brederode |
| TIME PERIOD | 19 March – 1 October 2018 |

# Preface

*This preface is written in Dutch*

Voor u ligt de Master Thesis *"Developing a model for estimating road capacity values for weaving sections".* Gedurende de afgelopen zeven maanden, heb ik een neuraal netwerk ontwikkeld welke capaciteitswaarden voor weefvakken schat. Deze thesis is geschreven ten behoeve van de afronding van de Master *Civil Engineering and Management,* met gecombineerde specialisaties *Transport Engineering and Management* en *Transport and Logisitics,* aan de Universiteit Twente en is uitgevoerd bij DAT.Mobility te Deventer.

Dit voorwoord wil ik tevens gebruiken om een aantal personen te bedanken. Allereerst wil ik graag mijn begeleiders namens de universiteit - Luc Wismans en Eric van Berkum - van harte bedanken voor hun goede begeleiding en feedback gedurende dit onderzoek.
Bovendien wil ook graag mijn dagelijkse begeleider van DAT.Mobility, Luuk Brederode, bedanken voor zijn goede ondersteuning, begeleiding, feedback en het aandragen van nieuwe ideeën voor het onderzoek.

Ten slotte wil ik alle collega's van DAT.Mobility bedanken voor de goede samenwerking, ondersteuning en werksfeer gedurende de afgelopen zeven maanden.

Ik wens u veel leesplezier toe.

Jurre Janssen

Deventer, 1 oktober 2018

# Abstract

Macroscopic traffic models are used for long-term projections on road networks and deliver travel times and congestion on road sections. Therefore, truthful road capacity values are essential parameters within traffic assignment models. Especially correct capacity values for weaving sections are of great importance because of their 'node' function within traffic models. However, exactly for these type of road sections, insufficient capacity values are currently used within the traffic model. Since, currently used capacity estimation methods are too time-consuming (microscopic simulation) or insufficient, the purpose of this research was to develop a model that estimates capacity values for weaving sections relatively fast, delivers truthful capacity values and is an improvement of the currently available capacity estimation models.

Since it is stated that capacity estimation problem cannot be solved by closed-form solutions (i.e. mathematical expressions), this research developed a data-analytic model, in the form of a neural network, which estimates capacity values by means of a set of independent variables. To do so, a dataset including a set of significant variables was gathered first. The model development process is an iterative process, which was divided in a pre-training phase, training phase and post-training phase. To identify if the neural network meets de demands of the research goal, real-case configurations and currently used capacity estimation models have been used to validate the model.

In the pre-training phase a suitable dataset is found in the appendices of the Dutch *Handboek Capaciteitswaarden Infrastructuur* (CIA). However, during supplementing the dataset, it was found that the dataset was not able to reproduce, whereby the complete data grid of CIA is re-simulated. The set of significant variables, that were able to implement in the model, where found to be the weaving configuration, configuration length, ratio of heavy traffic, ratio of weaving traffic and the division of traffic flows. During the training phase optimal structures and settings of the neural networks are iteratively found. A network with 10 neurons in the hidden layer delivered the most sufficient results. Finally, the estimated capacity values are analysed by means of the neural network output (i.e. regression plots, error histograms), a comparison of the estimated capacity values on real-case configurations and by comparing the estimation performance of the neural network and the currently used methods. It was found, during the post-training phase, that the neural network can estimate capacity values for weaving sections with only small errors and outscores the currently used methods for capacity estimations. Therefore, this research concluded that the use of a neural network delivers (relatively) fast and truthful capacity values for weaving sections. Moreover, the developed neural network is found to be an improvement of the currently used capacity estimation methods.

It should however be noted that the use of the neural network is restricted for extraordinary weaving configurations. This can be solved by gathering supplementary data for this kind of configurations. Furthermore, all capacity values used in this research are simulated with FOSIM, which is stated to be the 'ground-truth' for this research. It should however be mentioned, that uncertainties related to the used simulation settings and validity are present. Moreover, configurations where no capacity value can be simulated for, are not included in the neural network. Finally, the performance of the neural network can be improved, in future research, by implementing variables as speed limit and speed reductions in macroscopic traffic models.

# Table of Contents

# 1 Problem context and description

*The representation of congestion within macroscopic traffic simulation models is generally executed within the traffic assignment model of the macroscopic simulation model (see Paragraph 1.1). Congestion is the result of a too high traffic demand related to the road capacity (see Paragraph 1.2). Primarily for discontinuities in a road network the importance of correct capacity values is large (see Paragraph 1.3). However exactly for these type of road sections, insufficient capacity values are currently estimated or too time-consuming processes for estimating truthful capacity values are present (see Paragraph 1.4). Therefore, this research has the purpose to develop a solution, which is found to be a data analytic model, which can estimate capacity values for weaving sections relatively fast, delivers truthful capacity values and is an improvement of the currently available capacity estimation models (see Paragraph 1.5). In this chapter the problem context and solution approach is broadly discussed.*

## 1.1 Traffic Assignment models

Nowadays, different categories of traffic simulation models are used for evaluating traffic operations for, amongst others, freeway traffic. In general, three different categories of conventional traffic simulation models for freeway traffic can be distinguished, namely: microscopic, mesoscopic and macroscopic traffic simulation models. Microscopic traffic simulation models focus on individual vehicle movements. These models describe the reaction of every driver depending on the surrounding traffic of a certain driver. These reactions are accelerating, braking and lane-changing and are based on car-following models (Treiber & Kesting, 2013). Mesoscopic traffic models consist of both microscopic and macroscopic aspects of traffic simulation models (Payne, 1979). These models maintain individual vehicle representation (microscopic aspect) but with a more aggregate representation (macroscopic aspect) of traffic dynamics (Burghout et al., 2006). Macroscopic models are often used for strategic transport planning purposes and represents traffic flow on road segments in terms of aggregate measures as traffic density, speed and traffic flow. Macroscopic models are stated to be able to describe collective phenomena such as the evolution of congested regions or the propagation of velocity waves (Treiber & Kesting, 2013). Thus, the level of individual driving behaviour is not described by macroscopic models, however it is aggregated to a physical relationship through a cost function or the fundamental diagram of traffic flow (Bliemer, et al., 2012). Regarding the connection to cost functions and/or fundamental diagrams, it should be noted that macroscopic traffic models use capacity values as input, where microscopic traffic models can deliver capacity values as output. In Figure 1.1 a comparison of the above-mentioned categories of models is visually presented.

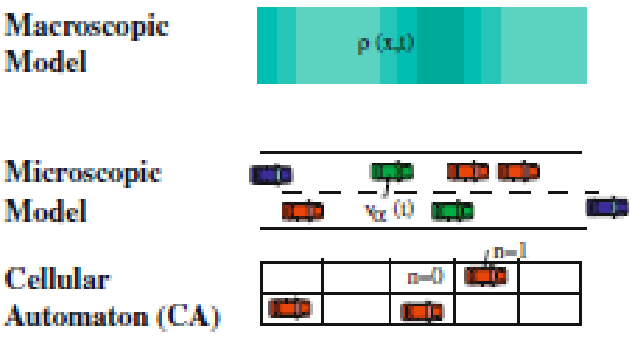



**Figure 1.1: Representation of the different traffic simulation models (Treiber & Kesting, 2013).**

Since macroscopic models are capable of modelling route choices and predicting traffic states on certain road networks, macroscopic models are extended with traffic assignment/propagation models. Traffic assignment models describe the interaction between road travel demand and the road infrastructure supply (Bliemer, et al., 2017). More generally, these models describe the interaction between flow and capacity. Traffic assignment models do consist of two main models,

between which interaction exist, namely: a route choice sub-model and a network loading sub-model.

Where the route choice sub-model determines traffic flows on road segments by means of travel demand, the network loading sub-model virtually propagates the traffic through the network and yields travel times given the infrastructure supply. Since travel times are, in many models, the most determinant attribute for route choice behaviour, the route choice is iteratively adapted to these travel times (depending on the equilibrium which is solved). Consequently, due to changing route choices, travel times per road segment change due to altered traffic flows.
Many different traffic assignment models do exist, whereby the above described interaction between network loading and route choices do coincide. *Bliemer et al. (2017)* described that the different model types and their capabilities could be classified on the basis of three different type of assumptions, which are: spatial, temporal and behavioural assumptions. The following model types can be distinguished based on spatial assumptions: unrestrained models, capacity-restrained models, capacity-constrained models and capacity- and storage-constrained models. Unrestrained models do not consider capacity values and roughly represent free flow traffic. In capacity-restrained models it is still possible that flows can exceed road capacity, however travel times increase when traffic flows increase. For capacity constrained models it holds that traffic flow cannot exceed capacity. Moreover, in capacity- and storage-constrained models it holds that traffic flow will not exceed capacity and ensures that queues will spillback to upstream road segments (Bliemer et al., 2017). In the latter case, travel times for the road segment, where traffic flow exceeds capacity, and adjacent road segments will increase due to the modelled spillback given that traffic demand exceeds infrastructure supply.

Temporal assumptions do concern the choice of time periods regarding travel demand (flow over time) and route choice loading. Where, dynamic models consider time-varying travel demand and multiply periods for route choice within each time period, static models consider a stationary travel demand and only a single time period for route choice and network loading. Finally, behavioural assumptions do concern allocation models for route choice. Where, all-or-nothing models allocate all travellers to follow the fastest route on given travel times (Bliemer et al., 2017), equilibrium models allocate travellers, which are stated to be non-cooperative, based on a certain user equilibrium. This user equilibrium holds that no individual driver can unilaterally reduce its travel times by shifting to another route (Wardrop, 1952).

Furthermore, temporal interactions on road networks are described by wave speeds and vehicle propagation speeds. Wave speeds are used to propagate traffic states through the road network. Wave speeds in the hypocritical branch (i.e. traffic densities lower than the critical density) are considered as forward waves. The wave speeds in the hypocritical branch are equal to the slope of the hypocritical branch of the fundamental diagram, which will be discussed in Chapter 3.1. When traffic densities are higher than the critical density, or when traffic flow exceeds capacity, backward wave speeds arise which propagate backwards on road segments. In this manner, queues and potential spillback arise on upstream road segments (Bliemer et al., 2017). These aspects of queue build-op and spillback only plays a role for capacity constrained models, since these models include hypocritical branches of the fundamental diagram. This means that actual flow rates on a link segment also depend on the upstream and downstream links. In other words, an interaction between traffic states on adjacent road section is present. Hence, output variables as (link) travel time are derived from cumulative inflows and outflows from that link segment (Bliemer et al., 2012). In this manner, it is feasible to calculate dynamic features as queuing, spillback and shockwaves.

In case of the software package OmniTRANS the traffic propagation or network loading sub-models MaDAM (dynamic) and STAQ (quasi-dynamic) are offered. MaDAM is a macroscopic and dynamic model used in OmniTRANS on behalf of realistic modeling of congestion dynamics. Several types of fundamental diagrams can be assumed within MaDAM (Bliemer, et al., 2017), which has four input parameters related to the fundamental diagram: the roadway free-flow speed, the speed-at-capacity, the capacity and the jam density (Rakha & Crowther, 2002). These input parameters are set for every link in then network before the simulation starts and are fixed during the simulation.
On the other hand, STAQ is a quasi-dynamic model, which has more realistic outcomes than static models and less calculation time than dynamic traffic assignment models. Since the computational efficiency of the STAQ model is relatively high in combination with its realism, it is recommended to use STAQ for calculations on larger road networks (Bliemer et al., 2012). Since both MaDAM and STAQ are models which are consistent with traffic flow theory, displayed in a well-known fundamental diagram, the outcome highly depends on the shape of the used fundamental diagram and the explicit capacity constraints of road sections (Bliemer et al., 2017).

## 1.2 Importance of road capacity parameter within assignment models

In case the (fixed) parameter of the capacity of a certain road section does not correspond with the actual capacity of the road section, implausible computations of the traffic conditions on certain road sections are made. Moreover, due to the interaction between traffic states on adjacent road sections in combination with capacity constraints which are taken into account in these macroscopic models, incorrect capacity values could negatively influence the plausibility of the complete network performance.

After all, correct capacity values for road segments are an important input parameter for the network loading sub-model. Since, interactions exist between the route choice sub-model and network loading sub-model, the output of traffic assignment models strongly depend on the capacity value for road segments. Capacity values can be fixed or dynamic, where fixed capacity values are constant over time and dynamic values do vary over time (per iteration or simulation) due to changing dynamic variables (i.e. traffic flow and composition). When the (fixed or dynamic) parameter of the capacity of a certain road section does not correspond with the actual capacity of the road section, implausible computation of travel times on the road section itself and the up- and downstream road sections are made. As a result, implausible input for the route choice sub-model is delivered, which results in a worse performance for a larger share of the road network. Hence, the road section capacity is a weighty parameter which is a determinant for the plausibility of the resulting congestion pattern (e.g. location and extent of the traffic jam) and the derived road network effectiveness indicators (e.g. speed, (link) travel times and vehicle-loss-hours). Therefore, the output of the traffic assignment model could become implausible due to the interaction between the two sub-models in combination with inaccurate values for the earlier mentioned four fundamental diagram related parameters.

## 1.3 The role of weaving section capacity

Discontinuities in a road network are prominent sources of the emergence of congestion. For freeway networks these discontinuities are weaving sections, merging lanes, lane drops and to a lesser extend diverges. Weaving sections are specific points of interest in macroscopic models, since here interaction between weaving and non-weaving vehicles plays a role, which reduces the ‘base capacity’ of these road sections. The base capacity can be seen as the theoretical maximum number of vehicles that can pass a road section without any discontinuities. In case of regular lane drops, capacity reductions are currently present in traffic assignment models. However, for merging lanes and weaving sections this is not the case. Moreover, it could be stated that weaving sections and

merging lanes roughly represent nodes in a road network. Therefore, the network performance highly depends on the performance of these specific sections and lanes. This is due to the fact that sending or outflow rates are decomposed in different directions at weaving sections and merging lanes. Therefore, is stated that intersections and its adjacent merging lanes or weaving sections are decisive for the operation of the complete road network (Rijkswaterstaat, 2015). Moreover, at the above described road sections mandatory lane changes are necessary for drivers to reach their desired destination. The obligatory merging manoeuvres cause turbulence in the traffic stream, which explains that capacity values for weaving sections and merging sections are lower than capacity values for upstream located road sections. *Rakha and Zhang, 2006* state that the introduction of lane changes within a traffic streams reduces the capacity in a fashion that is proportional to the level of turbulence within a weaving section. So called third order (quasi) dynamic traffic assignment models do already model the turbulence effects in their traffic propagation model. However, the most used models are first order implementations of these models, including STAQ, which not model the effects of turbulence. For these models, a reduced capacity value due to turbulence should be implemented. Furthermore, it is stated that capacity values for these essential sections are largely influenced by dynamic variables. These explanatory variables concern, amongst others, weaving flow rate, traffic flow composition and (entering) speed (Vermijs R. , 1998). Depending on the type of macroscopic model, the values for these explanatory variables are varying over time, while capacity values in many macroscopic traffic assignment models assumed to be fixed values and do not vary over time.

## 1.4 Current Practice

Currently, in the Dutch state of practice of transport modelling, capacity values for highway road segments, including weaving sections, are derived from the *"Handboek Capaciteitswaarden Infrastructuur Autosnelwegen"* (abbreviated as *CIA*), which is a manual for deriving capacity values for highway road segments published by the Dutch highway road authority. However, the CIA does not sufficiently provide unambiguous capacity values. This is due to the fact that capacity values in the CIA are point estimations (non-continuous) concerning road and environmental configurations under standard conditions (e.g. dry weather, good pavement conditions, no objects near road). With point-estimations it is meant that CIA provides capacity values for road configurations with a certain combination of values for critical variables. Critical variables according to CIA are the weaving section length, weaving ratio and ratio heavy traffic. For certain weaving configurations (i.e. 2+2 weaving sections), *CIA* provides capacity values with, roughly three varying values per category of critical variables. This results in $3^3 \equiv 27$ capacity values per weaving configuration. A deviation of the standard conditions, which will be discussed in Chapter 3.1, or values for critical variables will lead to different capacity values and could force the implementation of reduction factors to derive a more plausible capacity value. It is however not conclusive to multiply several reduction factors leading to a capacity value under more than one divergent condition.
Furthermore, finding an effective capacity value by interpolation between capacity values under standard conditions (e.g. weaving section length) is stated to be doubtful (Rijkswaterstaat, 2015). In addition, previous research found that parameters as heavy-traffic ratio do not have a linear relation with road capacity (Semeida, 2013). Moreover, the CIA states that the positive contribution of the length of a weaving segment to the road capacity relatively decreases as the length increases (Rijkswaterstaat, 2015). This also suggests that the relation between some variables (in this example weaving section length) is non-linear with road capacity, which means linear interpolation will be inadequate. Due to a combination of non-linear patterns and a deficient number of point-estimations of capacity, simple non-linear interpolation between capacity values under certain conditions will, most probably, not provide sufficient results.

Moreover, due to a growing road network and dynamic traffic conditions in macroscopic traffic models, it is time-consuming and ineffective to derive capacity values manually per road section by a derivation of capacity values in CIA or by microscopic traffic simulation. In other words, due to a theoretical infinite number of combinations of values for independent variables, simply performing microscopic traffic simulations is not plausible due to its time-consuming characteristics and the iteratively changing values of the independent variables. Furthermore, it is stated that a generation of a simple closed-form solution for capacity values given the specification of the other variables is not possible (Transportation Research Board, 2000). Solutions for solving these sorts of systems which are too complex for closed-form solutions, are stated to be mathematical modelling or computer simulation.

## 1.5 Solution Approach

In short, a necessity does exist for implementing capacity values, which are reduced compared to the 'base capacity' due to turbulence on weaving sections. The currently used or developed methods for generating capacity values for weaving sections are insufficient or are too time-consuming to implement within traffic assignment models. Therefore, a solution should be developed which can estimate capacity values relatively fast, which delivers truthful capacity values (i.e. corresponding with currently used capacity values) and is an improvement of the current capacity estimation models.
Previous research already estimated and expressed road capacities by means of linear regression and neural network techniques (Awad, 2004; Semeida, 2013). However, these estimations/expressions of capacity values are based on capacity values from the HCM and field data respectively. Since road conditions and traffic behaviour do vary per country, these capacity expressions are not applicable for the Dutch situation. Nevertheless, results from other previous research showed that it is possible to train neural networks, which is a form of mathematical modelling, able to approximate outcomes of traffic simulations with a high accuracy. Moreover, it was found that the method can be applied broadly within the field of traffic analysis and transport planning. The application of these kind of meta-models are less time consuming than executing traffic simulations (Gora & Bardoński, 2017). Furthermore, a meta-model, or in this case a model based on the outcomes of the microscopic traffic simulation model, simplifies the simulation model in two ways: the response of the meta-model is stated to be more deterministic than the original simulation model and the run times of the meta-model are generally much shorter than the original simulation (Barton & Meckesheimer, 2006). For these reasons, designing a meta-model to estimate road capacity values is preferred over implementing microscopic traffic simulations to estimate capacity values.

Since capacity values are important parameters in traffic assignment models and inappropriate interpolations between fixed capacity values are currently used, it is desired to develop a data-driven meta model which is able to estimate dynamic capacity values for weaving sections given a set of both static (e.g. weaving section length) and dynamic (e.g. flow rate, flow composition) explanatory variables influencing road capacity. For computational and model efficiency it is essential for the meta model that it could directly be connected to traffic assignment models within macroscopic traffic simulation models as OmniTRANS. Since, the meta model should be suitable for macroscopic traffic simulation models, certain boundary conditions regarding the explanatory variables are necessary. In other words, the meta model should preferably only contain input parameters for estimating capacity values which are already available as an input or can be derived from other input variables in traffic assignment models and/or microscopic traffic simulation models.

# 2 Research goal, strategy and questions

*In this section the purpose and the desired results of the research are defined. Furthermore, the research questions and the research strategy elaborated.*

## 2.1 Research goal

The aim of this research, derived from the problem context, is:

> *The aim of this research is to develop a model, applicable within macroscopic capacity constrained traffic assignment models, which estimates the capacity value for weaving sections given a set of explanatory variables*.*

* The set of variables that significantly influences the capacity at given road segments is described later in this research.

## 2.2 Research questions

To attain the research goal, the research is set up with three main research question, which are divided into several sub questions. The research strategy to answer these research questions are discussed in section 2.3.

1. **Which variables do significantly influence the value of road capacity for weaving sections?**
    a. Which type of road capacity is most suitable for the meta model and how is it defined?
    b. Which variables influence the road capacity according to the literature?
    c. Which variables do significantly influence the road capacity for weaving sections and merging lanes?
    d. Which of the explanatory variables are available in traffic assignment models and/or microscopic simulation models or could easily be derived or added?

2. **What is the most suitable meta model and how should it be developed?**
    a. According to the literature, which kind of data analytic models are available?
    b. Which type of data analytic model is the most suitable for the desired meta model?
    c. Which data is available and how can the dataset be supplemented?
    d. How can discrepancies or omissions between the available data and the significant variables (question *1c*) be solved?
    e. In which manner should the available or simulated data be implemented in the meta model?
    f. In which manner the data should be structured within the meta model?

3. **In which manner should the meta model be validated and/or applied in a test case?**
    a. What is/are (a) suitable method(s) for training and testing the meta model?
    b. Which performance indexes are suitable for describing model performance?
    c. Which options of model validations are available and can reasonably be used?
    d. How does the model perform when applying it to NRM configurations and how can this performance be defined?

## 2.3 Research strategy

*In this paragraph the challenges to reach the goal of the research are presented. First, a proper definition of road capacity is found. Then the capacity-influencing variables for weaving sections are elaborated. Since it is found in Chapter 0 that a dataset of capacity values is a necessity for estimating 'other' capacity values, a dataset with varying values for the significant variables is constructed. With a dataset consisting of capacity values, a data-analytic model is developed for estimating capacity values for real-case configurations. Before applying the model on real-case configurations, the model is firstly trained and tested on the dataset with currently present values of independent and dependent variables. In Figure 2.1 a summary of the research strategy is presented which will be extensively discussed in this paragraph.*

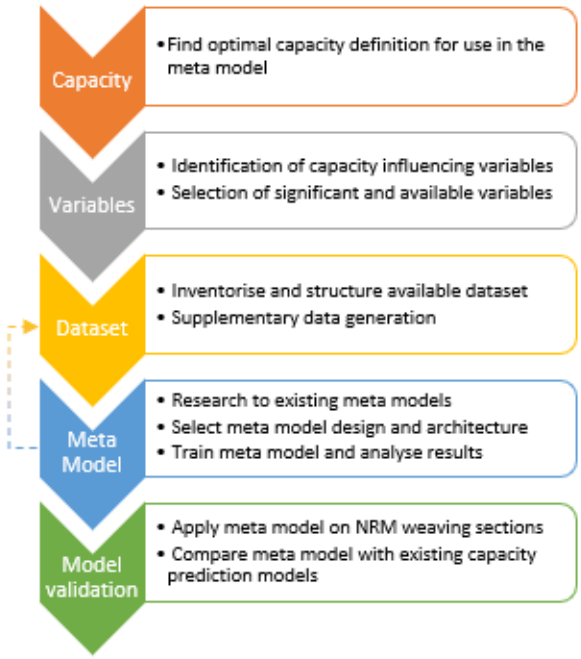



**Figure 2.1: Summarized overview of the research strategy represented by a flow chart.**

First of all, several definitions of road capacity do exist, of which a corresponding definition with traffic assignment models is selected. CIA considers more than one type of capacity value, namely the *free capacity* and the *capacity under congestion* (Rijkswaterstaat, 2015), an unambiguous capacity definition does not exist. Hence, it is essential to firstly define a proper capacity value for using in the desired model. The desired capacity value should correspond with the assumptions and definitions within/made by the traffic assignment model, since the desired model could be possibly applied within traffic assignment models. Therefore, the most suitable capacity definition depends on the traffic assignment model. Since, the capacity definition is already essential for the remaining content of the research, the currently available capacity definitions and the most suitable capacity definition for application within the model will be discussed in Chapter 3.1.

Secondly, the capacity values are dependent on certain variables concerning road geometry, traffic composition and other environmental factors. Therefore, a literature and/or simulation study examines which variables do significantly influence road capacity. According to the CIA and previous research, some variables influencing road capacity are (amongst others): lane width, heavy-vehicles ratio, number of lanes, free flow speed or maximum allowed speed (Awad, 2004; Rijkswaterstaat, 2015; Semeida, 2013). These capacity-influencing variables are elaborated in Chapter 3.2. Furthermore, a quantitative analysis on the influence of these variables on capacity is presented in Chapter 4.1. The quantitative analysis on the influence of the variables is made for two reasons, namely: it provides an insight in the capabilities of the model for pattern recognition and it indicates the insignificant variables.

Since some variables do not have a large contribution to the capacity values, not all variables could be considered in the model. Furthermore, some variables are (strongly) mutually correlated, for example the total lanes and median with. In such cases, it is inefficient to implement both mutual correlated variables. Hereby, it becomes inefficient to implement both mutual correlated variables, for example total lanes and median with. Therefore, the variables which are taken into account in the meta model are elaborated. Moreover, the explanatory variables which are applied in the meta model should, preferably, be available in or could directly be derived from the traffic assignment model. For example, the distance between the driving lane and crush-barrier could theoretically influence road capacity. However, when this variable is not available in traffic assignment models or

is not easily to derive or add, the explanatory variable cannot be used within traffic assignment models and the meta model.

In Chapter 0 it is stated that a model will be developed that can estimate capacity values relatively fast, which delivers truthful capacity values and is an improvement of the current capacity estimation models. Since this model is stated to be trained by means of capacity values belonging to certain weaving configurations, data is firstly gathered and structured to create a data grid/-set before developing the data-driven model. Several options do exist to derive capacity values, namely: by means of microscopic simulation and/or measured capacity values on the Dutch freeways. Measured capacity values are hard to derive due to the fact that it is costly and time consuming to determine the extent of weaving traffic on weaving sections. Furthermore, micro simulation makes it possible to supplement the available data by varying in values for explanatory variables. A disadvantage of data gathering by means of measured capacity is the lack of this possibilities in varying in values for explanatory variables. For these reasons, it is preferred to gather data by means of micro simulation. For this research the micro simulation software of FOSIM is used. Furthermore, FOSIM is validated for Dutch freeway traffic (Vermijs R. , 1998). Therefore, it is not necessary to re-calibrate the model-parameters, as driver characteristics, within FOSIM for all other configurations (Dijker & Knoppers, 2006). However, by using FOSIM some restrictions do apply, amongst others: the number of independent variables which can be implemented. Fortunately, a dataset for weaving sections derived from FOSIM simulations is already present in *CIA*. Therefore, the use of FOSIM will prevent a bias between the available and created data. Moreover, to prevent a bias with the CIA data, the simulation design should correspond with the design used in CIA. FOSIM will be further introduced in Section 3.4, in the theoretical framework. In Section 4.4 the used settings and simulation design will be discussed.

As stated above, a dataset of capacity values for weaving sections is already present in CIA, which is described in Chapter 3. Roughly, it contains capacity values for varying values for the explanatory variables. For estimating a meta-model, the dataset should match with real case configurations, which should be indicated. However, the dataset is most likely to be incomplete for the purpose of estimating a meta-model. In other words, the range of varying values for the explanatory is not sufficient to match with real-case configurations. For this reason, supplementary data is gathered with FOSIM, as described in previous paragraph. Since the design of the dataset is a component of the iterative process of model design, it is hard to define the efficient size of the dataset at this moment. For weaving sections, it holds that many configurations do or could exist. The weaving configuration has a marked effect on traffic operations and thus influences capacity values for these road sections (Transportation Research Board, 2000). Therefore, it would be time-consuming to simulate a broad range of values for explanatory variables for all configurations. Since the weaving configurations can be classified in different types of classes (e.g. symmetric or nonsymmetric), a broad range of values for explanatory variables will initially only be simulated for one type of weaving configuration, assuming a meta model can predict beyond several classes. As stated before, both the already available dataset and the desired ranges of values for explanatory variables are described in the theoretical framework in Chapter 3.

When both the most suitable capacity definition and its (most) influencing variables are determined and a suitable dataset is gathered, it is desired to process these into a model. The model does estimate the desired road capacity values based on a dataset with capacity values and the explanatory variables. According to previous research, regression models for predicting road capacity do give less confident results than neural network models (Awad, 2004; Kadari et al., 2015; Semeida, 2012; Semeida, 2013; Yap et al., 2015). Moreover, it is stated that neural network

technique models are able to predict capacity values for all weaving configurations together, and is easily able to identify the different hidden patterns across configurations (Awad, 2004). Therefore, on first sight it is preferred to use a neural network technique to predict capacity values for weaving sections. However, the precise type, design and architecture of the meta model should be defined and elaborated. More literature and arguments for the model choice and architecture is elaborated upon in the next Chapter 3.5.

Since, several types of weaving configurations are modelled, it will be discovered if more than one model is desired. In other words, it will be investigated if separate model types of weaving sections (i.e. symmetric and non-symmetric) do improve the total model results. Moreover, it could be that the model performance will improve if some weaving section configurations are separated into another sub-model. Firstly, this can be discovered by means of an analysis of patterns in the data. Furthermore, this could be discovered during the iterative process of the model development. In other words, if the model results improve when the model is split in two structures, a model architecture with two sub-models will be preferred. However, one can state that the meta model should be able to categorize this eventual segregation itself since the models are stated to be able to find hidden patterns in (large) datasets.

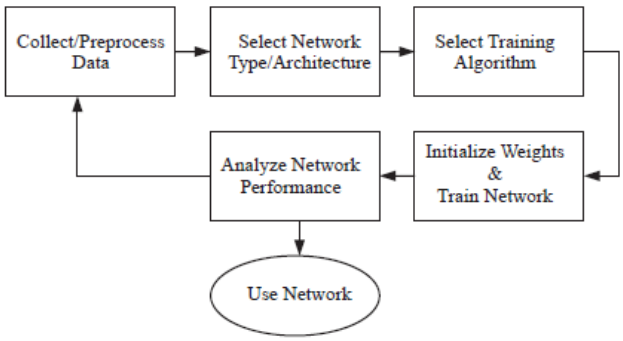



**Figure 2.2: Overview of the iterative model development process (Hagan et al., 2014).**

Once the model architecture is defined, the model is trained and tested with the available dataset. For this reason, the data will be separated into a training, validation and testing dataset. Generally, the training dataset make up approximately 70% of the full data set, where both the validation and testing dataset make up 15% each of the full data set (Hagan et al., 2014). Where the training set will be used to fit the meta model(s) to the data, the validation dataset is used to estimate the prediction error for the model selection. Finally, the test dataset is used for the assessment of the generalization error of the final chosen model (Hastie et al., 2001). It should be noted that it is of high importance that all the datasets represent the total dataset well. In other words: all the three datasets should contain approximately sufficiently similar scatter plots of value ranges for explanatory variables. For example: it is prevented that the training set only contains data points for short weaving sections and the test set only contains data points representing long weaving sections. Moreover, it is analysed if the model is performing worse for specific weaving configurations or combinations of values of explanatory variables. For example: the overall error of the model can be sufficient, however for some configurations the model could perform worse. Then the overall error could be improved by increasing model performance for that region of these configurations. For this purpose, regression plots of model errors could provide insights in both the competence of the model to interpolate or extrapolate and miscalculations for specific configurations (outliers). Regression plots display the trained output versus the target values and provide in that manner an insight in model errors for every specific combination of explanatory variables in the dataset. For example: running a separate dataset with only a specific type of weaving configurations can identify potential malfunctions of the model. Since the model architecture could be changed due to the results of the model training and validation, the model development processes is an iterative process, which is shown in Figure 2.2. For example, when the model does not perform on certain types of weaving sections, the dataset will be extended with more data points for these specific types to improve predictions for these configurations. Obviously, this also means that the performance is again evaluated for the other configurations as well. Besides, the

model architecture itself could be adapted to improve predictions of the meta model (Kadari et al., 2015). More literature on model architecture and the training process is described in Chapter 3.5. The training process itself is described in Chapter 5, which is completely dedicated to the training process.

Finally, the outputs of the trained model are analysed which process is better known as post-training analysis. For this case, the post-training analysis consists of three different elements. The first analysis is made of the model performance on the available dataset, which is divided in a training, testing and validation set. The second component is the analysis of the model performance on the real-case configurations to evaluate the ability of the model to interpolate and extrapolate as well. For this research it is assumed that the configurations which are available in the NRM-West model are representative for all Dutch weaving sections. The NRM models are strategic macroscopic models from Rijkswaterstaat used for long-term projection on the Dutch freeways (Rijkswaterstaat, 2018). Therefore, real-case configurations are derived from the NRM-West model because other capacity estimation methods have been applied on these configurations and it is safe to assume that it consists of a representative set of cases for the rest of the Netherlands. For this analysis, the weaving sections and traffic conditions which are currently present in real-case are the input of the data analytic model. The capacity values which are subsequently estimated by the data analytic model are compared to the simulated capacity values of the same configurations.

The last component of the post-training analysis is an analysis on the improvement (or deterioration) of the data analytic model. This is discussed by means of a comparison between the estimated capacity values by the data analytic model, the simulated capacity values (ground truth) and the previous estimated capacity values for the available real-case configurations. Capacity values are currently predicted by a calculation method incorporated in QBLOK, which is the standard network loading model of the Dutch NRM models. Furthermore, a nearest neighbour-based method has been developed for estimating capacity values. The nearest neighbour-based method compares the NRM configuration with the present configurations in the CIA manual and selects the capacity value of the most similar configuration. In this manner, it can be assessed if an implementation of the meta model will result in more accurate capacity values. The post-training phase of the model estimation is elaborated in Chapter 6.

Optionally, the meta model can be applied within macroscopic traffic simulation models, for example: the NRM-West model within OmniTRANS. Several options for implementing the meta model within traffic simulation models do exist. The main directions of implementing the model is by implementing it within the dynamic network loading sub-model or by implementing it after every iteration. However, the implementation om the meta model itself is not a component of this research. Although, by implementing the meta model, resulting traffic states, queues and spillback under the predicted capacity values can be compared to the output with the currently used capacity values. Moreover, resulting congestion images and locations can be compared to the output of the model using current capacity estimation methods and real-life congestion images using *Google Traffic* (Possel, 2017).

To summarize, in Figure 2.3 the model development process, as shown in Figure 2.2, is merged with the research strategy flow chart. The three phases of the model development are also the fundament for the structure of this research. In Chapter 3 a theoretical framework is presented which discusses the components of this research in a theoretical manner. In Chapter 4 the pre-training phase of the model development is elaborated. The pre-training phase comprise the definition of a suitable capacity definition, the inventory of the significant variables and the gathering of capacity values for the dataset. Subsequently, the training process will be described in Chapter 5 which includes the selection of the structure of the data analytic model and settings for the model training. In Chapter 6 **,** where the post-training phase is described, the output of the data analytic model is analysed. Furthermore, the data analytic model will be applied on real-case configurations, by doing so, a comparison in model performance can be made for different capacity estimation models.

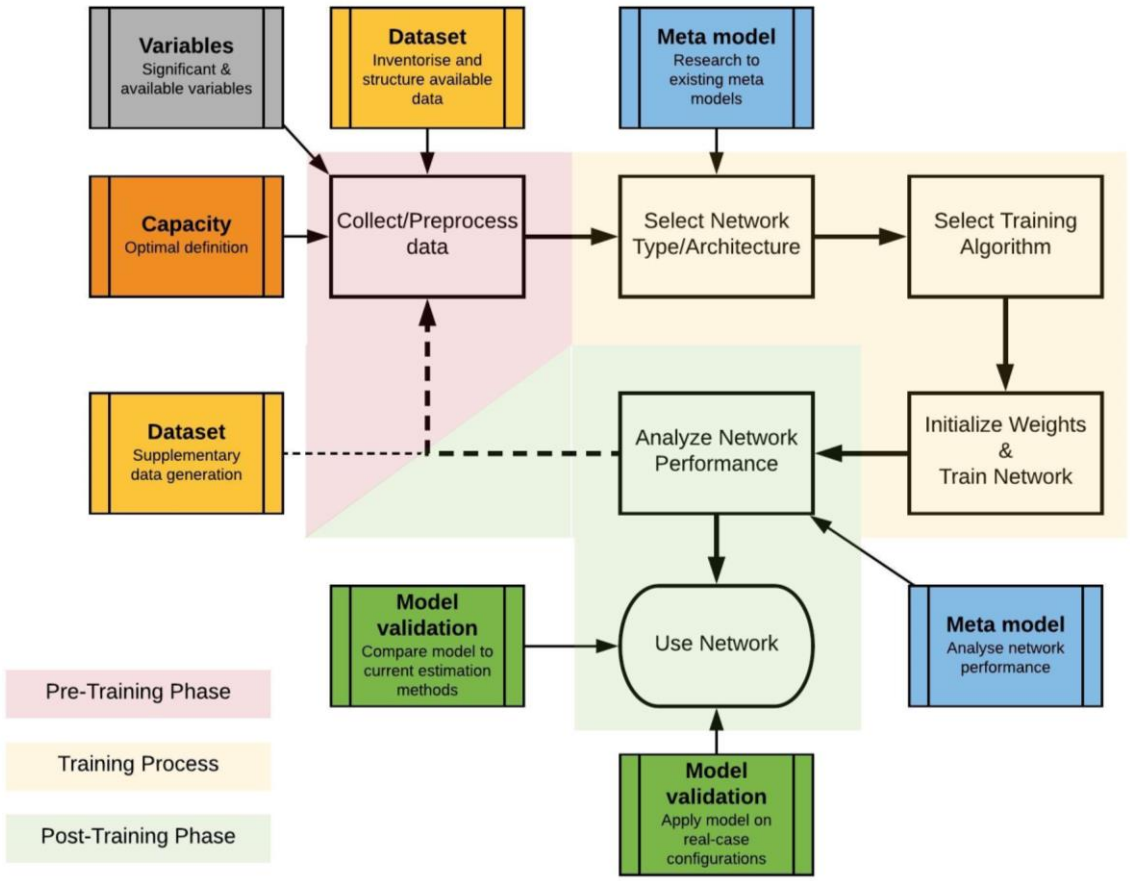



**Figure 2.3: The model development process (to: Hagan et al., 2014) merged with the activities of the research strategy. Where the total model development process can be divided in the pre-training phase, training process and post-training phase.**

# 3 Theoretical Framework

*Before discussing the model estimation phases (pre-training, training and post-training phase), this chapter elaborates on the available literature and previous research on the aspects concerning this research. The chapter is divided in five sections, which all concern a specific component of the research, corresponding with the flow chart in* Figure 2.1 *, namely: capacity definition, variables influencing capacity, (available) dataset(s), model selection and structure(s) and finally model training and testing. Besides, this chapter endeavours to answer all descriptive research questions by means of literature review and summarizes these answers (or answer approach) at the end of this chapter.*

## 3.1 Introduction to Road Capacity

Road capacity, often abbreviated with the capital letter C, is roughly defined as the maximum flow per lane (Treiber & Kesting, 2013). However, some different interpretations on this definition have been arising. Therefore, several different definitions of road capacity and road capacity values do currently exist. For this reason, the most 'suitable' definition of capacity applicable for this research should be chosen. To do so, a current set of commonly used capacity definitions will be discussed in this chapter. Finally, the most 'suitable' capacity definition for application within macroscopic traffic simulation models, from the elaborated set of capacity definitions, for this research is chosen.

### 3.1.1 Road Capacity Definitions

In the *CIA* road capacity is defined as: "The maximum number of motor vehicles per unit of time (mostly expressed as mv/h) of which it can be reasonably assumed that it can traverse a section or uniform segment of a road lane or road way during a certain period of time under prevailing road, traffic, control and environmental conditions." (Rijkswaterstaat, 2015). Obviously, the *CIA* has derived its capacity definition from the well-known Highway Capacity Manual of the Transportation Research Board. The *HCM* defines road capacity as: "The maximum hourly rate at which persons or vehicles reasonably can be expected to traverse a point or uniform section of a lane or roadway during a given time period under prevailing roadway, traffic, and control conditions." (Transportation Research Board, 2000).

Since traffic flow and thus road capacity is dependent of the prevailing road, traffic and control conditions, capacity values do alternate for different road, traffic and control conditions. Moreover, due to constantly changing composition of different drivers, their driving behaviour and their made choices, road capacity values are not fixed quantities but stochastic variables. Nevertheless, it is found that many road authorities, including *Rijkswaterstaat/CIA,* do work with static values for road capacities (Rijkswaterstaat, 2015; Calvert et al., 2016).

As stated above, the capacity values are dependent of prevailing road, traffic and control conditions. Both the *Highway Capacity Manual (HCM)* and the *Handboek Capaciteitswaarden Infrastructuur Autosnelwegen* (CIA) do provide capacity values for certain road configurations under base conditions. Therefore, the given (standard) road capacity values do only hold in case the prevailing road, traffic and control conditions correspond with the base conditions. The base conditions defined by CIA are, amongst others, that: the road is designed according current guidelines for designing highways; no large objects are situated directly next to the highway (for example: acoustic shielding); the road is not on a slope; in daylight under dry weather; the pavement is in good condition and without any form of traffic management measures (Rijkswaterstaat, 2015).

Above the prevailing base conditions, the definition of the road capacity has other implications, namely the 'dependency on traffic composition' and 'during a given period of time'. As earlier

described, the road capacity is not a fixed value due to the stochastic element in traffic composition. Amongst other things, the number of heavy vehicles (trucks etc.) do have an influence on the capacity value and do change over time. Therefore, the CIA provides standard capacity values with a share of fifteen percent of heavy traffic of the total traffic flow. This value is seen as a frequent ratio of heavy traffic. Furthermore, the 'during a given period of time' component of the capacity definition plays a significant role. The capacity value depends on the used definition of the aggregation level of time. In case traffic flow will be measured with an aggregation level of one minute, the (median) measured capacity value, based on the maximal traffic intensity, will be higher than capacity values measured with an aggregation level of fifteen minutes.

According to *CIA* three methods exist to assess capacity values for a given road section. First of all, the capacity values can be assessed by means of microscopic traffic simulations. Since the dynamic traffic model FOSIM is specially calibrated and validated for the Dutch highways, *Rijkswaterstaat* (as author of the CIA) uses this model to assess capacity values. Therefore, capacity values in *CIA* are all derived from FOSIM simulations. These capacity values are currently used in Dutch practice.
Secondly, the capacity value can be assessed by means of measurement data. According to *CIA*, the Brilon method or fundamental diagram method can be used assessing capacity values. The Brilon method provides a capacity observation as the downstream intensity during a time interval in the time interval before the interval that congestion occurs. Here, congestion is defined as a speed detection below a certain threshold at an upstream detector (Leferink, 2013).
Furthermore, the capacity can also be assessed by means of the fundamental diagram method. By using this method, the measured traffic intensity will be displayed on the y-axis and the corresponding traffic density will be displayed on the x-axis of the fundamental diagram or a derivate of the fundamental diagram. Afterwards, the measured data will be fitted to a fundamental diagram. Because of this, the capacity can be derived from the fitted fundamental diagram. An example of this fundamental diagram and the fitting of the data is shown in Figure 3.1.
Finally, *Minderhoud et al., 1997* made a broader inventory of available capacity assessment methods and concluded that the product limit method, of which the Brilon method is a derivative, is preferred above the empirical distribution method and fundamental diagram method respectively.

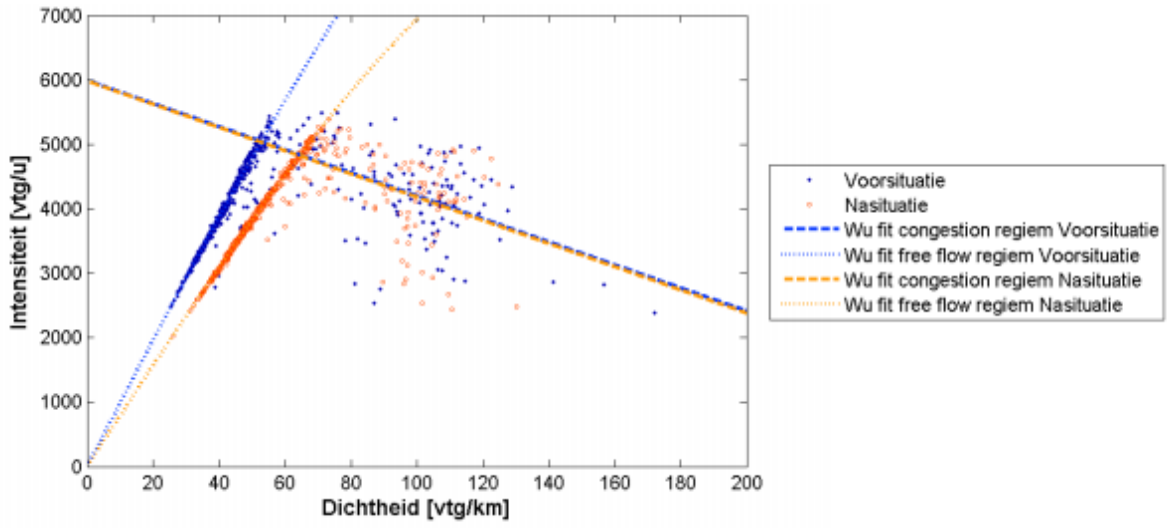



**Figure 3.1: Example of a fit from measured data to a fundamental diagram. Here, the traffic intensity (veh/h) is displayed on the y-axis and the traffic density (veh/km) is displayed on the x-axis (to: Rijkswaterstaat, 2015).**

Next to the above described free capacity definition (or: inflow capacity) also a congestion capacity (or: outflow capacity) definition exists. The CIA defines the outflow capacity as the capacity when congestion has occurred. It could be measured through the downstream traffic intensities once congestion has occurred upstream (Rijkswaterstaat, 2015). It is found that the difference between the free capacity and outflow capacity, which is known as capacity drop, is strongly dependent on the local situation. However, research found that the difference between the free and outflow capacity is mostly between ten to fifteen percent (Grontmij, 2009).

### 3.1.2 Most Suitable Capacity Definition

As stated above, it is needed to choose the most suitable (set of) capacity value(s) to be assessed with the meta model. The terminology 'most suitable' means in this case the set of capacity value(s) which correspond, preferably seamlessly, with the capacity definition within traffic assignment models. To find the 'most suitable' capacity definition for this research, a closer look to fundamental diagrams and their corresponding capacity values is desired. All traffic assignment models, including capacity constrained traffic assignment models, assume a fundamental diagram. The fundamental diagram is defined by a hypocritical branch and hypercritical branch. Here the hypocritical branch represents traffic conditions with densities lower than the critical density or flows lower than the critical flow. The hypercritical branch represents traffic conditions for densities higher than the critical density or traffic flows higher than the critical flow. In Figure 3.2 it is illustrated that the hypocritical (blue part) and hypercritical (red part) branch of the fundamental diagram are separated by the value for capacity (C). It therefore should be defined if the desired capacity definition represents traffic flows at the hypocritical or hypercritical branch. Since traffic flows at the hypercritical branch are mostly instable, traffic flows at the hypocritical branch are more stable to calculate average traffic conditions. For the concerning traffic assignment models, this value is defined as the maximum flow through any part of a road section also known as the physical road capacity (Bliemer et al., 2017). This definition of the capacity value at the hypocritical branch lies in line with the definition of the *free capacity* by *Rijkswaterstaat*. Moreover, capacity values which are already used within OmniTRANS are derived from the highway capacity manual published by *Rijkswaterstaat*. For these reasons, the most suitable capacity definition of this research is the definition of the *free capacity* of *Rijkswaterstaat*, namely: "The maximum number of motor vehicles per unit of time (mostly expressed as mv/h) of which it can be reasonably assumed that it can traverse a section or uniform segment of an road lane or road way during a certain period of time under prevailing road, traffic, control and environmental conditions" (Rijkswaterstaat, 2015). Since capacity values are and will be derived by means of FOSIM simulations, the capacity is made operational by the maximum number of vehicles per hour that can pass a certain road section (detector downstream of the diverge) within a time interval of 15 minutes.

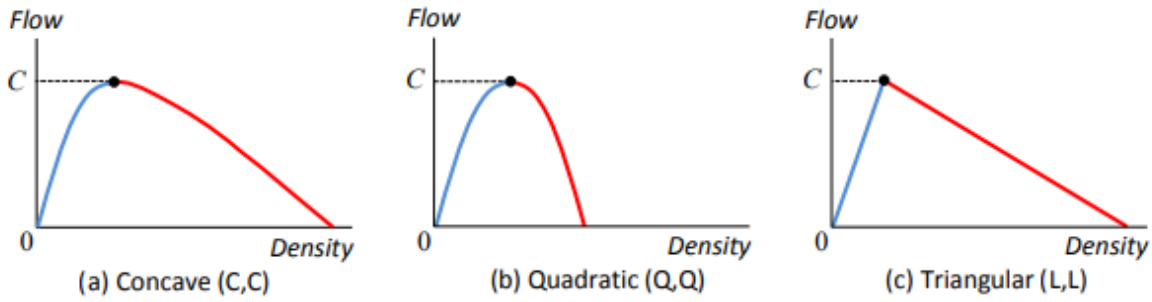



**Figure 3.2: Several examples of fundamental diagrams that can be assumed in traffic assignment model. Where the blue part of the fundamental diagram represents the hypocritical branch and the red section the hypercritical branch (Bliemer, et al., 2017).**

## 3.2 Variables Influencing Road Capacity

Since the provided capacity values in CIA and HCM do only hold if the road configuration and environmental context correspond to the base conditions, any deviation of these conditions will hold that the provided capacity value is not legitimate anymore. Therefore, reduction factors are needed to correct the capacity values. In this chapter the general and weaving section and merging lane specific variables influencing capacity values will be distinguished. Furthermore, restrictions and limitations of using these explanatory variables in the final meta model will be discussed.

### 3.2.1 General variables

As stated in the previous paragraph on road capacity, capacity values only hold if the road configurations and environmental context correspond to the so-to-say standard conditions. According to the CIA the deviating circumstances could be distinguished in six main categories, namely: infrastructural factors; environmental factors; traffic factors; traffic management factors; developments in in-car systems and ITS and incidental factors. These factors are derived from available research and HCM (Rijkswaterstaat, 2015). For almost every deviation a reduction factor is given. This reduction factor should be multiplied with the capacity value under base conditions to assess the actual road capacity value. However, actual capacity values become untrustworthy if more than two reduction factors are applied (Rijkswaterstaat, 2015). In Figure 3.3 the general variables derived from the CIA are summarized. In addition, more research on influencing variables for road capacity is available. *Calvert et al., 2016* do also recognise the need for accurate stochastic capacity values as input for traffic models. They found that capacity values on workdays are 4% higher than on weekend days and a decrease of the capacity drop of 8% in comparison with workdays. Moreover, it is found that variable speed limits on freeways reduce the chance of traffic break down (Geistefeldt, 2011). However, it should be remarked, as already stated, that not all variables can be taken into account in the meta model, since not all variables are present or taken into account in traffic assignment models. Moreover, influencing variables could only be implemented in the meta model if they are present in the available data or could be added (by simulation or measured data) to the currently available data. Clearly, some restrictions do hold for implementing these general variables in the meta model. These restrictions will be discussed in Paragraph 3.2.4.

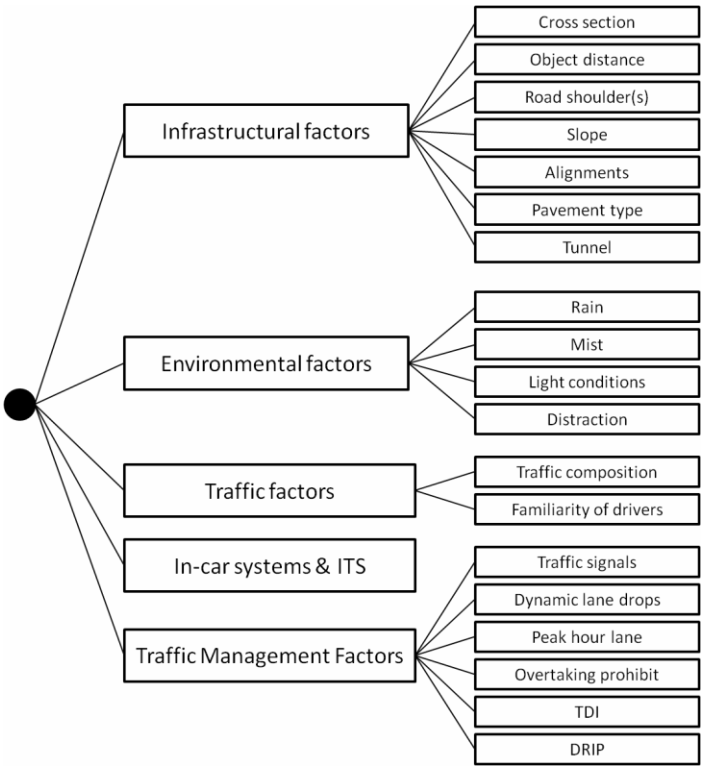



**Figure 3.3: Overview of the general capacity influencing variables according to CIA.**

### 3.2.2 Driving behaviour on Weaving sections

In this section the specific variables that influence road capacity for weave sections will be elaborated. Since weaving sections are acknowledged as concentrated/high turbulence areas within the freeway system (Zhang, 2005; Al-Jameel, 2011; Roess, 1988) and capacity is negatively influenced by an increased level of turbulence (Rakha & Zhang, 2006; Van Beinum et al., 2018), the driving behaviour on weaving sections is analysed first.

On weaving sections and directly up- and downstream at weaving sections, drivers are required to perform driving manoeuvres to enter or exit the mainline of the freeway, or to cooperate and/or anticipate on the weaving traffic. These manoeuvres include (desired or) required lane changes and change in speed and headways (Van Beinum et al., 2018). The individual driving manoeuvres are visualised and are shown in Figure 3.4. When summarized, the individual driving manoeuvres are: gap search, adjust speed, execute lane changes and adjust lead headway (Vermijs R. , 1998).

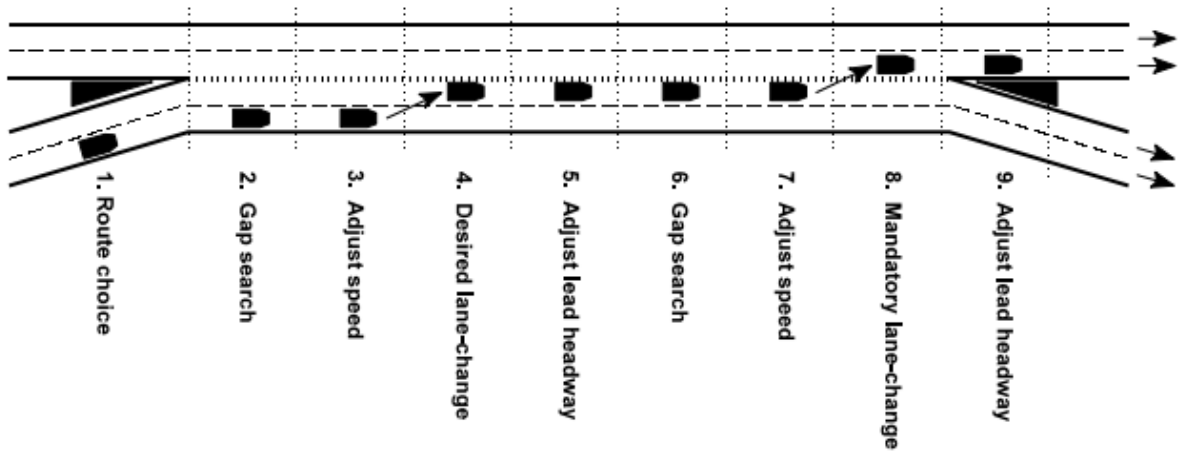



**Figure 3.4: Schematic overview of the individual driving manoeuvres on weaving sections (Vermijs R. , 1998).**

Due to these manoeuvres, changes in lane flow distribution, greater speed variability and changes in headway distribution (on the different lanes) will arise. These phenomena are referred to as turbulence, which negatively influence highway capacity (Van Beinum et al., 2018). In other words, when lane changes are mandatory in high density traffic, the procedures of gap-searching and lane changing lead to braking manoeuvres which cause lower traffic speeds and as a consequence traffic-breakdowns.

*Van Beinum et al., 2018* have developed a more extensive theoretical framework for turbulence compared to *Vermijs, 1998* (Figure 3.4). In this framework, shown in Figure 3.5, the phenomena turbulence is structured considering three parts, namely: driving manoeuvres, microscopic behaviour and macroscopic effects. Here, the individual driving manoeuvres are coupled to microscopic behaviour of drivers. The lateral and longitudinal behaviour, which are both microscopic, do result in macroscopic effects as a change in density (vehicles per kilometre per lane), speed and headway. These macroscopic changes are directly influencing (the degree of) turbulence since turbulence is defined as the joint effort of changes in lane flow distribution, greater speed variability and changes in headway distributions.

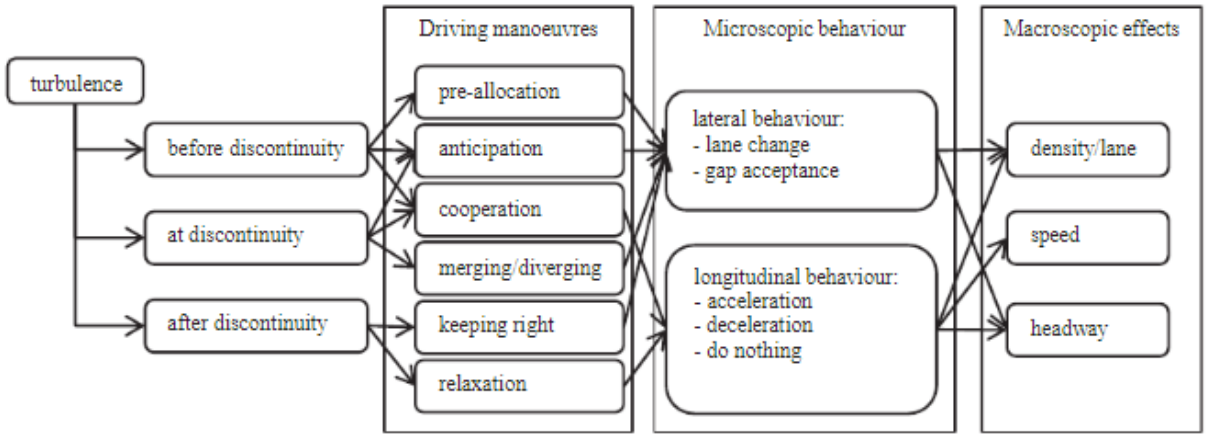



**Figure 3.5: Theoretical framework for turbulence near freeway discontinuities (including weave sections) (van Beinum et al., 2016).**

In the next paragraph the variables influencing road capacity are described qualitatively by means of a literature review. These variables could, theoretically, all be traced back to the components of turbulence as shown in Figure 3.4 and Figure 3.5.

### 3.2.3 Weaving section specific variables

Elaborating on the obtained knowledge about the driving behaviour and adjacent turbulence on weaving sections, the variables influencing capacity for weaving sections specific are investigated based on a literature study. A more quantitative analysis, based on a data-analysis, of the influence of the separate variables on capacity is made in Chapter 4.

The impact of lane changes on freeway operations is stated to be a function of traffic composition, weaving/merging demand level, site geometry, information and signage system, and driver characteristics (Al-Kaisy et al., 1999). Based on this statement and research on capacity values of merging, weaving and diverging sections, the following influencing variables are summed up (Awad, 2004; Rakha & Zhang, 2006; Transportation Research Board, 2000; Rijkswaterstaat, 2015; Vermijs, 1998; Zhang, 2005):

- Weaving section configuration;
- Weaving section length;
- Traffic composition;
- Ratio of weaving traffic;
- Allowed speed;
- Speed differences between vehicles.

#### *3.2.3.1 Weaving Section Configuration*

A well-formulated definition for the weaving configuration is given by the *Transportation Research Board* in the *Highway Capacity Manual*. The weaving configuration is namely defined as: 'The organization and continuity of lanes in a weaving segment, which determines lane-changing characteristics.' (Transportation Research Board, 2000). Since it is found that the weaving configuration influences lane-changing characteristics, which has impact on the turbulence in the traffic flow (Rakha & Zhang, 2006), it can be concluded that the weaving configuration do have a certain influence on road capacity.

Hence, the weaving or merging configuration do influence road capacity. Actually, the configuration of a weaving section can have several different components which influence the road capacity.

Three main components of the configuration can be identified, namely: the number of available lanes, the availability of a taper merge or diverge and especially for weaving sections the arrangement of lanes (i.e. the difference between a symmetric 2+1 and 1+2 configuration).

For this reason, the *HCM* classify the weaving configuration in several weaving types. The Weave type is defined as a classification scheme that categorizes weaving configuration into one of the tree types, namely type A, B and C (Transportation Research Board, 2000). These three types of weaving configurations are separated by means of the (minimum) required lane changes for the weaving traffic to reach their destination. In Figure 3.6 three examples for the different weaving types are separated.

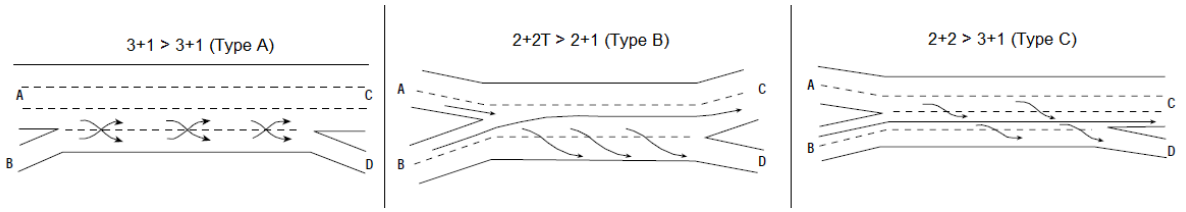



**Figure 3.6: Examples of the three different weaving types according to the *Highway Capacity Manual* (To: Transportation Research Board, 2000).**

From Figure 3.6 the three different weaving types and the (minimum) required lane changes per configuration type are shown. It can be seen that at least one lane change per direction is needed for configuration type A. For weaving type B (taper configurations) only one direction is required to make a singular lane change. Weaving type C involves at least two lane changes for one direction while the other direction can change destination without any required lane change.

The reason for this segmentation of the weaving configurations is made because of the different impact of the different lane-change characteristics. For example: a weaving flow of 500 vehicles per hour on configuration A and C implies 500 required lane changes where a weaving flow of 500 vehicles per hour on configuration B could imply 250 required lane changes.

Moreover, the configuration effects the proportional use of lanes by weaving and non-weaving vehicles (Transportation Research Board, 2000). In other words, drivers do have less flexibility in lane use for configuration type A, since the weaving vehicles must primary occupy the two lanes adjacent to the crown lane. For type B segments this effect is stated to be the least severe, since this configuration type require the fewest required lane changes, namely [0,1] or [1,0].

#### *3.2.3.2 Weaving Section Length*

In the literature it is found that the length of a weaving section theoretically will influence road capacity. The length of a weaving sections namely constrains the time and space in which the driver must execute the driving manoeuvres described in Figure 3.4 to execute a (required) lane change (Transportation Research Board, 2000). Regarding the driving manoeuvres required for lane changing, an increased length of a weaving section provides favourable opportunities for drivers to search a sufficient gap and adjusting their speed. Moreover, an increased length implies more available time for cooperative behaviour by performing cooperative lane changing or yielding to create gaps (Marczak et al., 2013).

However, it is found that the impact of the configuration length on capacity decreases as the length of the weaving sections increases (Rakha & Zhang, 2006). In other words, adding extra length to a relative short weaving section has more impact on road capacity than adding length to a relative long weaving section. CIA does also discuss this relation, it is namely stated that: “in case the length

of a weaving section increases, capacity will increase as well. However, at an increasing length the increase of capacity will increase relatively less. This can be explained by the fact that most lane changes are made at the beginning of a weaving section. Therefore, the supplementary length of a long weaving section has a relatively limited effect on road capacity" (Rijkswaterstaat, 2015).

*Van Beinum et al., 2018* and *Marczak et al., 2013* confirm in their researches that most of the lane changes are performed at the beginning of the weaving section. Moreover, *Al-Jameel, 2011* did also found that capacity values do increase at an expansion of configuration length, primarily for short weaving configurations. At a certain length, an increase in length do not have a significant influence on road capacity. Furthermore, it is found that the influence of configuration length is not an isolated variable. *Al-Jameel* namely found that the influence of configuration length is larger for weaving sections with four lanes than for weaving sections with two lanes.

Regarding the patterns found in this section, the configuration length does significantly influence road capacity. It is stated that the greater impact of increasing configuration length lies primarily at short configuration lengths. At longer configurations, an increase of length has relatively less or no impact on road capacity.

#### *3.2.3.3 Traffic Composition*

A component of the traffic composition is the ratio of heavy traffic, the ratio of heavy traffic is defined by the relative share of trucks on the total traffic flow. According to the guidelines for estimating capacity values, road capacity is influenced by the ratio of heavy traffic. This can be explained by the smaller accelerating performance of trucks, the reduction of sight for other drivers and because trucks cannot easily change lanes (Rijkswaterstaat, 2015). For these reasons, capacity values for traffic conditions with deviating ratios of heavy traffic are (or could be) rectified by means of reduction factors. These reduction factors are derived from the representation of number of cars by trucks. In other words: a truck could be counted as two regular vehicles. This linear relationship of corrections factors is also utilized by *Awad, 2004* in his research on estimating road capacity by means of multiple linear regression and neural network modelling. *Al-Jameel, 2011* does also endorse that the ratio of heavy traffic negatively influences road capacity. Moreover, it is found that the influence of the ratio heavy traffic differs according to the number of total lanes in the weaving section.

However, for merging sections it is stated that these reduction factors cannot simply be applied for traffic situations with a traffic flow of more than 750 trucks per hour. Then the most right lane is namely occupied with a large number of trucks, whereby lane changing is extremely hindered (Rijkswaterstaat, 2015). When assuming for weaving sections that non-weaving tucks will use the two most right lanes, this restriction do also hold for weaving sections. Furthermore, previous research found that the relationship between road capacity and the ratio heavy traffic is not linear (Semeida, 2013; Jiang & Adeli, 2004). This founding implicates that the often used (linear) reduction factors based on CIA are not legitimate.

Other aspects of traffic composition are the familiarity of drivers with the specific road conditions or weaving section, which is already shown in Figure 3.3. In the literature it is found that the familiarity of a driver with the route clearly influences the driving process (Intini, 2016). Furthermore, the driving behaviour in certain geographical regions can differ in relation to driving behaviour in another geographical region. This is found in a study on driving attitudes and behaviour in rural and urban areas in Norway. Here it was concluded that differences in attitudes and self-reported behaviour were significant due to type of geographical area (Nordfjærn et al., 2010). For the Dutch

practice it is only stated that personal driving style can deviate a lot from the average driving style (Ligterink, 2016).

#### *3.2.3.4 Ratio of weaving traffic*

Next to the ratio of heavy traffic, the ratio of weaving traffic also is a capacity influencing variable. The ratio of weaving traffic is a measure which represents the number of lane changes on weaving sections. Lane changes increase the level of turbulence which negatively influence capacity. The ratio of weaving traffic can be expressed in three components, namely: the ratio of weaving traffic from origin 1 (WR1), the ratio of weaving traffic from origin 2 (WR2) and the volume ratio (VR). In Figure 3.7 a schematic overview is shown of a 2+1 weaving sections with its origins and weaving traffic. The weaving ratio is equal to the amount of weaving traffic from an origin compared to the continuous traffic from the same origin. For example, when 1000 vehicles come from origin 1, where 300 vehicles go to destination 2 and 700 to destination 1, the weaving ratio will be 300/1000 = 30%. The volume ratio is defined as the total weaving vehicles compared to the total weaving and non-weaving traffic.

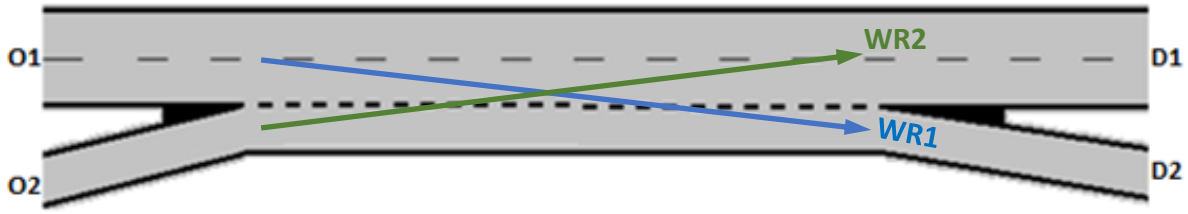



**Figure 3.7: Schematic overview of a weaving section with O1 (Origin 1), O2 (Origin 2), WR1 (Weaving Ratio 1) and WR2 (Weaving Ratio 2).**

In the capacity-value tables of *CIA*, no volume ratio is calculated. On contrary, *CIA* provides the weaving ratios for both origins. When both weaving flows are known, the volume ratio can be calculated by: $VR = \frac{WeavingFlow_1 + WeavingFlow_2}{Flow_1 + Flow_2}$. Since *CIA* assumes that traffic flows are proportional to the division of lanes and when the weaving ratios are known, the volume ratio can be calculated by: $VR = \frac{WeavingFlow_1 + WeavingFlow_2}{Flow_1 + Flow_2} = \frac{(WR1 * Lanes_{H1}) + (WR2 * Lanes_{H2})}{Lanes_{H1} + Lanes_{H2}}$.

In the CIA dataset it can be seen that capacity decrease consequently when the volume ratio increases. Moreover, according to *Awad*, *2004* the volume ratio has a relatively large influence on capacity values of weaving sections. In his research a correlation matrix for the capacity of weaving sections, as shown in Figure 3.8, and the six most influencing variables showed that the volume ratio has, together with the number of lanes, the strongest correlation with the capacity of a weaving section (Awad, 2004). Moreover, in the linear regression model of *Awad* the volume ratio has the highest beta and relative increase. This relation is also seen in the highway capacity manual.

| Variable | $C_i$ | $N_{w\,max}$ | $N_{w\,min}$ | $S$ | $L$ | $N$ | VR |
|---|---|---|---|---|---|---|---|
| $C_i$ | 1.00 | | | | | | |
| $N_{w\,max}$ | 0.18 | 1.00 | | | | | |
| $N_{w\,min}$ | −0.05 | −0.35 | 1.00 | | | | |
| $S$ | 0.17 | 0.00 | 0.00 | 1.00 | | | |
| $L$ | 0.17 | 0.00 | 0.00 | 0.00 | 1.00 | | |
| $N$ | 0.66 | 0.05 | −0.14 | 0.00 | 0.00 | 1.00 | |
| VR | −0.62 | −0.19 | −0.30 | 0.00 | 0.00 | 0.00 | 1.00 |

**Figure 3.8: Correlation matrix for the road capacity data af weaving sections in HCM. Where $C_i$ is the capacity (pc/h), $N_{w,max}$ and $N_{w,min}$ represent a variable for the minimum and maximum required lane changes, S is the free flow speed (km/h), L is the length of the weaving segment (m) and N represents the number of lanes and VR is the volume ratio (Awad, 2004).**

### 3.2.3.5 *Allowed Speed*

The allowed speed on a road section also influences road capacity. It is found that a lower speed limits homogenizes traffic and reduce (desired) lane changing activity (Soriguera et al., 2017). Therefore, less turbulence and higher capacity values are expected. Moreover, *Cremer, 1979* did also found that capacity increase occurred as a result of traffic homogenization, which also is supported by a created flow-occupancy diagram, shown in Figure 3.9. It is shown that the maximum flow, which is equal to the capacity, increases when the speed limit (expressed as *b*) decreases.

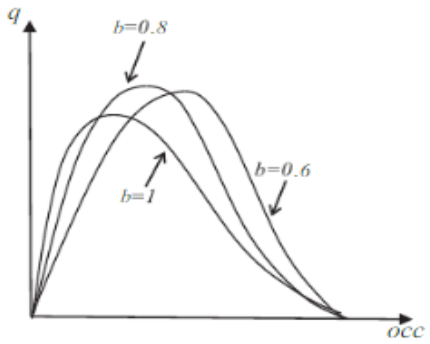



**Figure 3.9: Flow - Occupancy diagram, where *b* stands for the ratio between the speed limit and free flow speed (Soriguera et al., 2017)**

Anyhow, the relationship above holds for a continuous freeway sections without disturbances (i.e. merges or weaving sections). *CIA* has investigated the influence of the speed limit on capacity values for weaving sections. In Figure 3.10, capacity values with varying speed limits are shown. In this figure it can be seen that, for these configurations, the optimal capacity lies at a speed limit of 100 km/h. Moreover, from this figure it can be seen that the influence of the speed limit differs according to the varying weaving section lengths and ratios of heavy traffic.

**Influence of the speed limit on a 2+2 symmetric weaving section**

| Speed limit (km/h): | | 80 | 100 | 120 |
|---|---|---|---|---|
| %HT | Length (m) | | | |
| 5 | 600 | 6730 | 7430 | 7070 |
| 15 | 600 | 6080 | 6470 | 6420 |
| 25 | 600 | 5580 | 5720 | 5720 |
| 15 | 900 | 6500 | 6800 | 6800 |

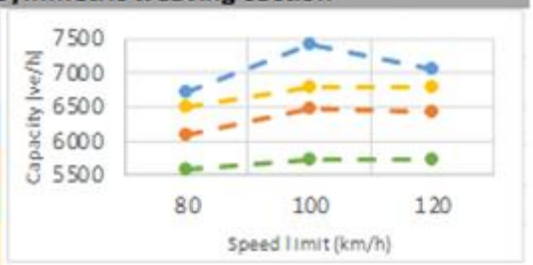



**Figure 3.10: Capacity values from the "*Handboek Capaciteitswaarden Infrastructuur*" for a 2+2 symmetric weaving section with varying speed limits. Left: Capacity values in a table form with varying ratios of heavy traffic (%HT) and weaving length. Right: Capacity values in a graph form. Data Source: Rijkswaterstaat, 2015.**

### *3.2.3.6 Speed differences*

As already introduced in Paragraph 3.2.3.1, speed differences between the weaving and non-weaving traffic do negatively influence road capacity. This is also supported by previous research conducted by *Hidas, 2005* which concluded that the merging process becomes increasingly difficult as the speed difference between the merging and non-merging vehicles increases. This is due to the fact that speed differences between vehicles lead to an increased gap acceptance. *Marczak et al., 2013* confirm in their research that the difference in speed between the putative follower and the merging vehicle do significantly influence the gap acceptance of the merging vehicle. In their logistic regression model it is found that the probability of gap acceptance decreases when the speed difference between the merging vehicle and putative follower increases.

At weaving sections a speed difference between the weaving and continious traffic can also arise due to sharp turns on the single weaving/merging lane. Hereby, the merging vehicles enter the weaving sections, as depicted in Figure 3.11, with a lower speed.

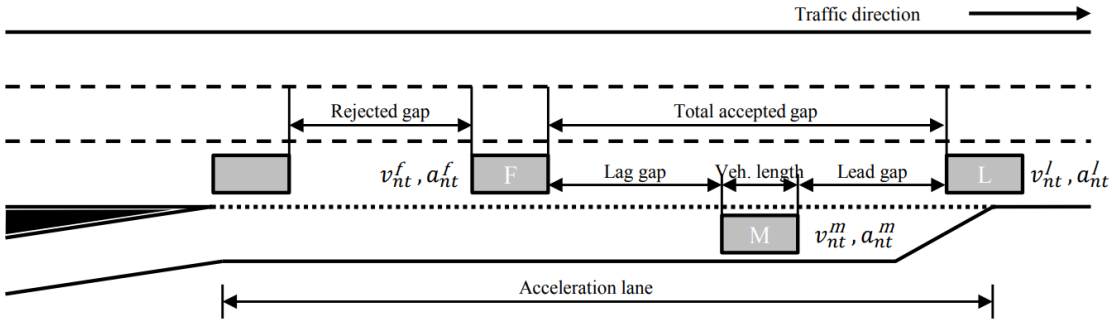



**Figure 3.11: Overview of the merge process for merging/weaving vehicles. Here, the merging vehicle do accept or reject a gap based on the gap, speed, speed differences and accelerations (Marczak et al., 2013).**

Furthermore, in Paragraph 3.2.3.1 it is shortly discussed that the weaving configuration influences speed differences between weaving and non-weaving vehicles. Since faster vehicles do have to overtake slower traffic on the left lane, mean speeds on the left lane will be higher than on the rightest lane. On a single lane, faster traffic is required to follow the slower traffic, which results in a relatively low mean speed on the single lane. This phenomenon is shown schematically in Figure 3.12.

Here, it is shown that an 1+2 configuration forces the slower weaving traffic to change lanes with fast continuous traffic. Due to higher speed differences, capacity values for a 1+2 configuration will be lower than for 2+1 configurations. This can be explained by a lower acceptance of gaps when speed differences increase. Consequently, more space for merging/weaving is required which negatively influences capacity.

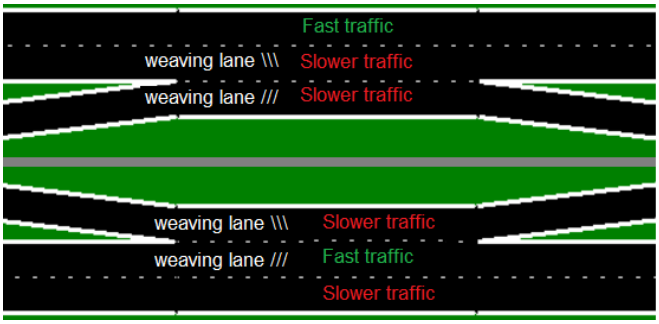



**Figure 3.12: Overview of the symmetric weaving configurations 2+1 (up) and 1+2 (down). Moreover, an indication of the weaving lanes (lanes where the required lane changes need to be executed) and slow or fast traffic is made.**

### 3.2.4 Restrictions on using variables

As already described, some restrictions on using the variables influencing capacity in the meta model do hold. Since values for the explanatory variables should both be present within or can be derived from the traffic assignment model and microscopic traffic simulation model, the restrictions can be derived from FOSIM and macroscopic traffic simulation models, in this case OmniTRANS. Furthermore, the respective variable should be implemented in FOSIM on an accurate way within behavioural models. In this paragraph a short indication/description of these variables will be made. However, in Chapter 4 the final set of variables which can be considered within the meta model will be described.

First of all, from the FOSIM user manual, which will be described in Paragraph 3.4, it follows that all the weaving section specific variables (as described in Chapter 3.2.3) can directly be implemented in FOSIM. For the general variables (as described in Chapter 3.2.1) it holds that a large part can directly be implemented in FOSIM. The effect of other variables on capacity values as slope, alignments, pavement type, tunnel, rain, mist and light conditions can indirectly be approached by adapting driver characteristics or speed reductions. However, implementing this method of changing driver characteristics could lead to suspicious outcomes (Dijker & Knoppers, 2006). The familiarity of drivers could theoretically be approached by an adaption of the lane change areas. However, the familiarity of drivers is not easily to determine. Moreover, insufficient knowledge is present regarding the relationship between driving behaviour and familiarity (Intini et al., 2016).

On the other site, the macroscopic traffic simulation model does also restrict the explanatory variables that can be used. First of all, all weaving section specific variables, which include weaving configuration, weaving section length, traffic composition, weaving ratio and allowed speed, can directly be derived and implemented from/within OmniTRANS. However, infrastructural factors as lane width or slope, are not structurally implemented in OmniTRANS.
The environmental factors (i.e. rain, mist) are not implemented in OmniTRANS and can therefore not be used. Moreover, the purpose of most strategic transport models is to model traffic for average working days. Therefore, these variables are not relevant, since they do not include/occur during an average working day. Furthermore, infrastructural factors are not regularly implemented in OmniTRANS. Theoretically, these variables can be added to road links separately. However, it can be questioned if implementing these factors is worth the effort regarding the, relatively small influence of these factors on the road capacity (Rijkswaterstaat, 2015). For the reasons above, these variables are not taken into account in the final meta model. This results that only the traffic composition from all variables in Figure 3.3 is taken into account.

Regarding these *'two-sided'* restrictions, all the weaving section specific variables can directly be implemented within the meta model. These variables include: weaving configuration, weaving length, traffic composition, ratio of weaving traffic and speed limits. Furthermore, other variables as overtaking prohibitions and road configurations (i.e. peak lanes or slopes) can only be implemented after adding values for these variables within the OmniTRANS model first. However, it can be questioned if these variables will significantly influence road capacity and/or it is worth implementing these variables in OmniTRANS.

The final set of considered variables will be determined after a data-analysis on significant variables is executed and an inventory of the available data(sets) is made. These two processes are part of the pre-training phase, as shown in Figure 2.3, of the model development and will be discussed Chapter 4.

## 3.3 Available Dataset

Since the solution approach is focussed on a data analytic (or meta) model, data is necessary to train the model. The Highway Capacity Manual of *Rijkswaterstaat*, CIA, provides capacity values for weaving sections which are, mostly, derived from simulations in FOSIM (Rijkswaterstaat, 2015). As stated before, the advantage of microscopic traffic simulation in FOSIM is that FOSIM is validated for Dutch freeways (FOSIM, 2018). Moreover, the usage of FOSIM provides the opportunity to vary with values of explanatory variables, which is not possible when using loop detector data. In this section the available data derived from CIA for weaving sections is described. All these capacity values lie in line with the already chosen capacity definition in Paragraph 3.1.2. Furthermore, some statements about potential data supplements will be discussed in this chapter.

First of all, CIA distinguishes symmetric and asymmetric weaving sections. The configuration of a symmetric weaving section does not vary for the complete length of the weaving section. This does not hold for asymmetric weaving sections. For asymmetric weaving sections the configuration can vary along the length of the weaving section. For example, an asymmetric weaving section could have two lanes origin from the north and one from the south and end with one lane heading for the north and two heading for the south (configuration 2+1>1+2). In a more mathematical expression this will be expressed as:

$$\boldsymbol{if}\ (Lanes_{O1} = Lanes_{D1}\ \&\ Lanes_{O2} = Lanes_{D2}\ )\ \boldsymbol{then}\ "Symmetric"\boldsymbol{else}\ "Assymetric"$$

For both type of weaving sections, capacity values are provided by CIA. These capacity values are provided with several combinations of values for explanatory variables. A sample of capacity values is shown in Table 3.1. In this table it can be seen that capacity values do differ for combinations of explanatory variables. The explanatory variables considered by CIA are the weaving configuration, ratio of heavy traffic, weaving length and ratios of weaving traffic. For example, it is shown that capacity decreases from 8690 to 6840 vehicles per hour when the heavy traffic ratio increases from 5 to 25%. Other such tables are presented for 10 symmetric weaving section configurations and for 24 asymmetric configurations. This results in 727 datapoints. A single datapoint represents a capacity value for a weaving configuration with a combination of weaving traffic ratio, length and ratio of heavy traffic.

**Weefvak type 2+2 -> 2+2**

| % wevend verkeer * | | | 5 % vrachtverkeer | | | 15 % vrachtverkeer | | | 25 % vrachtverkeer | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Weefvak-lengte (m) | 650 | 750 | 850 | 650 | 750 | 850 | 650 | 750 | 850 |
| H2->B1 | H1->B2 | max. snelheid (km/h) | 120 | 120 | 120 | 120 | 120 | 120 | 120 | 120 | 120 |
| 25% | 25% | | 8.690 | 8.710 | 8.830 | 7.670 | 7.690 | 7.660 | 6.840 | 6.830 | 6.850 |
| 50% | 50% | | 7.240 | 7.510 | 7.430 | 6.550 | 6.640 | 6.700 | 5.800 | 5.870 | 6.010 |
| 75% | 75% | | 6.040 | 6.290 | 6.180 | 5.420 | 5.620 | 5.740 | 4.810 | 5.060 | 5.150 |

**Table 3.1: Capacity values, according to the free capacity definition of CIA, for a symmetric weaving section configuration 2+2 (Rijkswaterstaat, 2015).**

Since a data-analytic model is not able to extrapolate and the validity of large interpolations is doubtful, these 727 datapoints represent, on first sight, a too small range of values for explanatory variables. Therefore, it will most probably be required to simulate capacity values for more data points. For example, CIA has capacity values for symmetric weaving sections 3+1 for a length of 600, 700 and 800 meters, while a length of 1680 meter is observed at real-case weaving sections (within the NRM-West network). Furthermore, weaving sections 3+1 with weaving ratio 25/25 are observed in a real case, while the current dataset of CIA only provides values for 50/17, 75/25 and 100/33. Moreover, truck ratios higher than 25% are observed at real-case configurations, with a maximum of 55%.

Despite the extensive dataset for weaving sections, more data points will be gathered by means of simulation, since the meta model should predict capacity values for real-case configurations, which will be represented by the weaving sections from the NRM-West model. Moreover, it is stated that data analytic models perform relatively worse in case of extrapolation (Hagan et al., 2014). Furthermore, previous research also showed that 'extreme values' (e.g. very low weaving ratio's or very low truck ratio's) are lacking in the dataset (Vermijs R. , 1998). This is also shown in Figure 3.13 where no unambiguous capacity values can be derived for (extremely) low and high weaving flow rates.

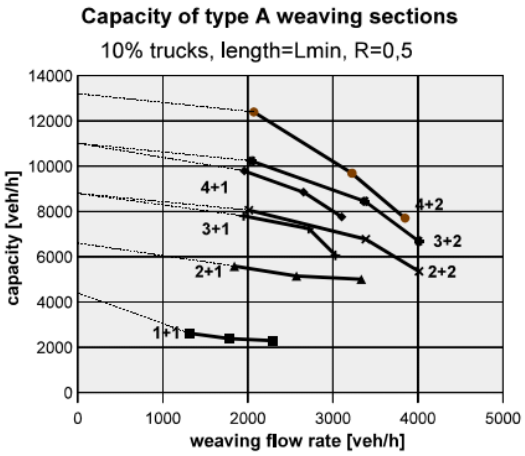



**Figure 3.13: Graphical overview of capacity values for symmetric weaving sections with varying weaving flow rates (Vermijs R. , 1998).**

Concluding, the dataset for weaving sections consist of 727 datapoints with varying values for the explanatory variables. However, supplementary data should be created with a broader range of values of explanatory variables.

## 3.4 FOSIM Simulations

The available dataset of CIA will be supplemented by means of the microscopic traffic simulation software of FOSIM, as indicated in previous paragraph. In the previous chapters the arguments in favour of using FOSIM are already elaborated. In this paragraph the process of gathering data with FOSIM is elaborated by means of the manual guide of FOSIM.

As already stated, FOSIM is a microscopic traffic simulation model which is calibrated and validated for Dutch freeway traffic. Therefore, the parameters concerning driving behaviour can be used without any adjustment (FOSIM, 2018). FOSIM simulates the driving behaviour of individual drivers. The behaviour of all drivers together results in a certain traffic flow. Therefore, traffic flow characteristics as road capacity is an output of the model, which is desired as an input for the traffic assignment model in this case. Essential for executing a simulation is the translation of the real road and traffic conditions to input where FOSIM can deal with. Since FOSIM can deal with all commonly road configurations as symmetric and asymmetric weaving sections, taper merges and taper diverges, all the configurations concerning this research can be implemented in FOSIM.

For executing a simulation and generating a capacity value for a certain road section, the first step to be taken is the implementation of the weaving configuration. When the road contours of a certain configuration are implemented, characteristics as speed limits can be added to separate road sections. In Figure 3.14 an example of a weaving configuration implemented in FOSIM is shown.

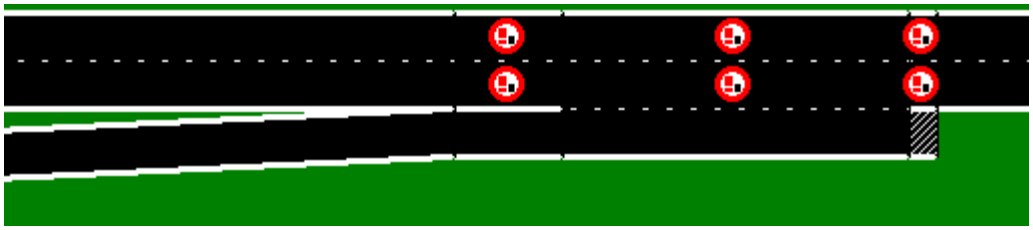


**Figure 3.14: Overview of a road configuration of a merging lane with overtaking prohibitions for heavy traffic implemented in FOSIM.**

The local characteristics that are the most relevant for this research are: speed reduction factors (for the purpose of curves and speed differences between traffic flows) and speed limits. After implementing the road configuration and local characteristics, lane change behaviour and so-called lane change areas should be added. Lane change areas are created in case vehicles are required to change lane. These lane change areas consist of two parts, namely: a desired lane change part and a required lane change part. FOSIM decides, by means of the road geometry, on which location lane change areas are required. Depending on the weaving configuration, the length, location and division of lane change areas should be defined.

Furthermore, detectors should be added to the road geometry. These detectors gather measured variables of all passing vehicles. For the purpose of measuring capacity values for certain road configurations, at least two detectors should be added. One detector is placed upstream of the expected bottleneck to register if congestion has occurred. The detector downstream then measures the capacity which is defined as the highest measured intensity. A schematization of the location of the detectors is shown in Figure 3.15.

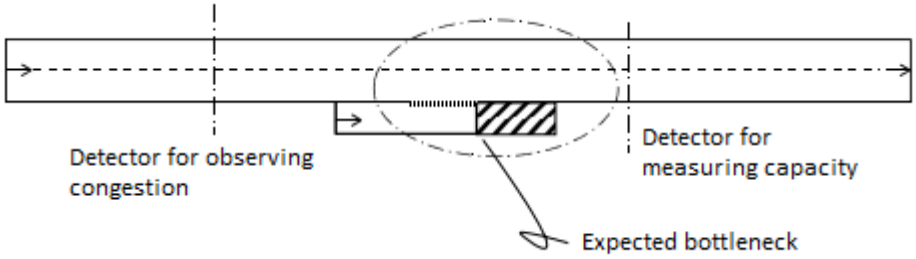



**Figure 3.15: Location of the detectors for the purpose of measuring capacity values (Dijker & Knoppers, 2006).**

Now, one can define the traffic composition and traffic intensities. For the traffic composition, the ratio heavy traffic can be set for each origin. The traffic intensities will be increasing over time, with time steps of five minutes (which is a commonly used aggregation level). Since the traffic intensities are increasing, at some point the intensity will reach the capacity. Then congestion occurs which will be registered by the upstream detector. At that moment the downstream detector will write-in the highest measured traffic intensity.

Since capacity values do have a stochastic character, more simulation runs will be executed for certain road configurations. More detailed: the results of a simulation depend on a draw by lot during the simulation. This draw by lot determines the departure time and driving behaviour of individual drivers within the simulation. Therefore, the results of only one simulation cannot give an unambiguous judgement about the capacity values. A sufficient number of simulations should be executed with different values for the random generator. The required number of simulations per combination of variables can be approached by a statistical formula, namely: $n > \frac{z^2}{d^2}\sigma^2$. Here, the number of simulations is represented by *n*, the level of confidence by *z*, the standard deviation of the measures by *σ* and the desired accuracy of the capacity value by *d*. A distribution graph is the output of a series of simulations, an example is shown in Figure 3.16. Within the distribution graph the mean and median capacity values are shown, where the median capacity value is the final capacity value for the certain road configuration (Dijker & Knoppers, 2006).

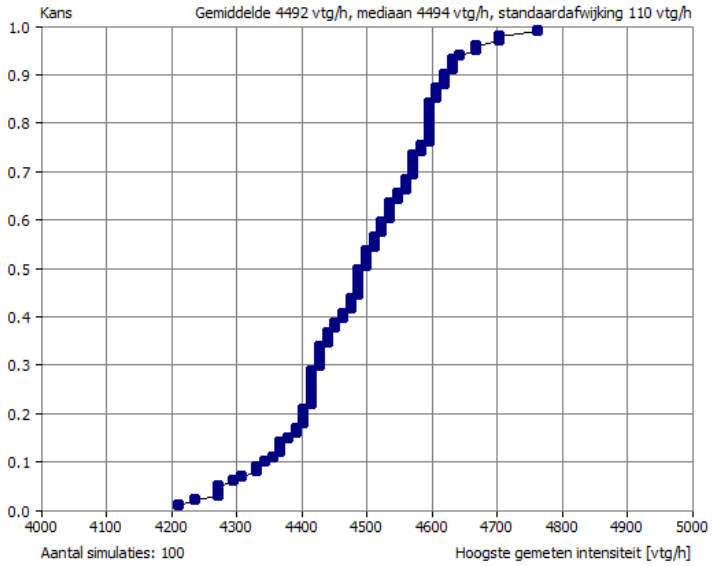



**Figure 3.16: Capacity value distribution for a merging section (length: 650m; %trucks: 15%; merging ratio: 10%; max. speed: 120 kpu). For this configuration, which is shown in Figure 3.14, the final capacity (median value) is equal to 4494 veh/hour.**

## 3.5 Data-analytic Models

In this paragraph the available and used data analytic models in the literature and previous research are elaborated. First of all, an overview of used data analytic models in the literature will be provided. Then, advantages and limitations of the available models will be discussed. The choice of the data-analytic model will be assessed regarding the predictive power, pattern recognition and model simplicity. Furthermore, in the next two paragraphs more literature on model structure and model training will be discussed.

### 3.5.1 Available Data-analytic Models

First of all, a well-known category of data analytic models are regression models. For this research multiple linear regression (MLR) and multiple nonlinear regression (MNLR) models could theoretically be used. These types of models are often used to model linear or nonlinear relationships between dependent variables and independent variables (Kadari et al., 2015). The assumption of these models is that the relationship between the dependent variable and the vector of regressors is linear (MLR) or nonlinear (MNLR). For predicting or forecasting the dependent variable (in this case road capacity) the regression equation will fit a model to the observed dataset of independent and dependent variables (Adamowski et al., 2012). Examples of the use of regression models in the field of traffic engineering are given by *Miauo*, *1994; Kadari et al.*, 2015; *Yap et al., 2015* and *Semeida, 2013*. *Miauo, 1994* researched the relationship between truck accidents and geometric design of road sections with several regression models. *Kadari et al., 2015* developed a model for pedestrian gap acceptance behaviour at unprotected mid-block crosswalks under mix traffic conditions by means of multiple linear regression. Furthermore, *Yap et al., 2015* developed a model for predicting roundabout lane capacity by means of a set of explanatory variables. Here, a multiple nonlinear regression model is used. Finally, the most related research to this topic is the research of *Semeida, 2013* In his research capacity for road sections is assessed by means of multiple nonlinear regression (Semeida, 2013). In the developed model, road section capacity is predicted on the basis of a set of explanatory variables, including, amongst others, number of lanes, lane width and the ratio heavy traffic.

Another method closely related to regression analysis is kriging, which is also called spatial correlation modelling. Kriging is used to predict a value or function for a given set of values by computing a weighted average of the known values in the neighbourhood of the point (Kleijnen, 2007). This method is, however, not used often in the field of transportation research. Moreover, this method is relatively similar to the nearest neighbour method, which is currently used for capacity estimation. It is therefore questionable if the implementation of the kriging method will lead to a significant improvement of capacity estimations.

Next to the regression-based models, artificial neural networks are well-known data analytical models to predict dependent variables by means of a set of independent variables. It is found that various transportation studies have used artificial neural networks (ANN), which are an important alternative to statistical regression models (both linear and nonlinear) for data analysis and pattern recognition in large datasets (Yap et al., 2015). Artificial neural network techniques are computational tools with different possible structures (Awad, 2004). In other words, several network paradigms do exist (Smith & Demetsky, 1997).The fundamental structure of neural networks consists of a series of nodes and weight factors that link the various nodes together in hierarchical manner. This hierarchical manner is represented by three types of layers, namely: an input layer, hidden layers and an output layer (Lord & Mannering, 2010). In the input layer, the independent or explanatory variables are situated. The output layer consists of the output variables of what is being modelled, in this case the weaving section capacity. The nodes in the hidden layers between the input and output layers are connected by links each carrying a weight that quantitatively describes the importance of those connections, thus denoting the strength of one node to affect the other node (Semeida, 2013). Examples of the use of neural networks in the field of traffic engineering are given by *Dharia and Adeli*, 2003; *Jiang and Adeli, 2004* and *Awad, 2004*. *Dharia and Adeli, 2003* developed a neural network model for rapid forecasting of freeway link travel time. Despite the fact that they stated that backpropagation is the most widely used neural network model, *Dharia and Adeli, 2003* used a counterpropagation network, which are both network paradigms. This choice is made because backpropagation networks have, among others, a very slow rate of convergence. *Jiang and Adeli, 2004* used a clustering-neural network model for freeway work zone capacity estimation. In these clustering networks, work zone patterns are first grouped into similar clusters using a data clustering approach. The clustering of similar work zone patterns could also be used for similar weaving section configurations. The most related research to this topic is the research of *Awad*. *Awad, 2004* used neural network techniques for estimating traffic capacity for weaving sections. In the developed neural network, the capacity value of weaving sections is estimated by means of six explanatory variables. In this case, the explanatory variables are the maximum and minimum number of lane changes by weaving movements, the number of lanes, the length, the free flow speed and the volume ratio.

*Barton and Meckesheimer, 2006* inventoried other meta modeling methods for optimizing simulation models. However, these models are not used, for the best of the authors knowledge, in the field of transport engineering. Only two of their described models, neural networks and regression analysis are used in the field of transport engineering, which are already elaborated in this section.

### 3.5.2 Model Selection

Now the two most-used data analytic models are presented, the advantages and limitations of the models will be discussed. A number of the above described researches both used regression models and neural network models for predicting and/or estimation. It was found that neural network models have a better prediction and estimation capability than regression models (Adamowski et al.,

2012; Awad, 2004; Kadari et al., 2015; Semeida, 2013; Yap, Gibson, & Waterson, 2015). This conclusion is also derived by those researchers who focussed on capacity estimation using both neural network and regression techniques. Furthermore, it is stated that neural networks are more capable of recognise paterns in complex and nonlinear problems and large datasets than regression models (Smith & Demetsky, 1997). Moreover, it is stated by *Yap et al.* that neural networks are able to approximate much more complex relationships between dependent and explanatory variables, including nonlinearity and interaction between explanatory variables (Yap et al., 2015). This brings us to a limitation of regression models. It is namely stated that regression models are not able to handle multicollinearity among independent variables (Semeida, 2013). This problem is also described by *Kadali et al.* who noted that the main drawback of regression models is the multicollinearity problem due to which several explanatory variables are eliminated from the model due to insignificance of statistics. Moreover, it is stated that this problem can be rectified with artificial neural network techniques due to its massive structure of input over the output connected architecture (Kadari et al., 2015). Besides, neural network learning algorithms adapt connection weights to improve performance based on current results. Traditional statistical techniques are not adapte but typically process all training data simultaniuosly (Duliba, 1991). Finally, it is stated that the use of regression techniques requires prior knowledge of patterns in the dataset.

Based on the above stated arguments in favor of neural networks, one could state that the use of neural network techniques is preferred for this research. However, the use of neural networks attended with several drawbacks and limitations. First of all, neural networks are seen as a black-box procedure and may not have interprentable parameters (Lord & Mannering, 2010). This in contrast to regression models, where relations between input and output variables are clearly shown (Kadari et al., 2015). Furthermore, four main problems in the use of neural networks in data modelling are discussed by *Livingstone et al.,* they found that overfitting, change effects, overtraining and interpretation (as stated above) are the four main problems in the use of neural networks (Livingstone et al., 1997). The problem of overtraining can be reduced by a proper combination of training and testing the model. Here, it is important that also testing data is used which is not present in the training dataset. Furthermore, the problem of the lack of overview of the most influencing variables can be tackled by a sensitivy analysis, which is elaborated in the research of *Kadari et al.* (Kadari et al., 2015).

Since the use of neural network techniques is, on first sight, preferred based on the above stated arguments, the fundamental architecture of neural networks will be described in the next paragraph. Furthermore, the network development and training will be discussed in paragraph 3.5.2. Since, more network paradigms (e.g. backpropagation, clustering networks) are available, the model type depends on the model development and training results which is an iterative process (see Figure 2.2).

## 3.6 Introduction to Neural Networks

Since the use of neural network models is preferred on first sight, due to its predictive performance, a literature study on neural networks will be executed. First of all, an introduction on neural network structures will be given in Paragraph 3.6.1. In accordance with regression-based models, neural networks should be trained, and the performance should be tested. Since several training methods and training settings are available next to several output-analysis methods, an inventory of these components will be made in Paragraph 3.6.2.

### 3.6.1 Neural Network Structures

The structure of a neural network model will contain several components. The three main components of a neural network are the input layer, the hidden layer(s) and the output layer. Furthermore, for every layer several choices regarding components or specifications of the layer should be made. These choices and specifications do depend on the external problem specification In short, a neural network is not an unambiguous model. In other words, several choices need to be made within a neural network structure. The book of *Hagan et al.* does provide a good overview of the current neural networks and its structures (Hagan et al., 2014).

In Figure 3.17 a neural network is schematically displayed. It can be seen that a neural network contains an input layer and a layer of *S* neurons. Moreover, for complex problems it is stated that more neurons in the hidden layer and/or the number of hidden layers will increase model performance. The input layer is a vector or static values from a vector denoted as respectively **p** or $p_r$. In the case of this research, the input vector would contain values of the significant variables. For example: **p** = [*Weaving section length, Ratio Heavy Traffic, Ratio merging traffic*].

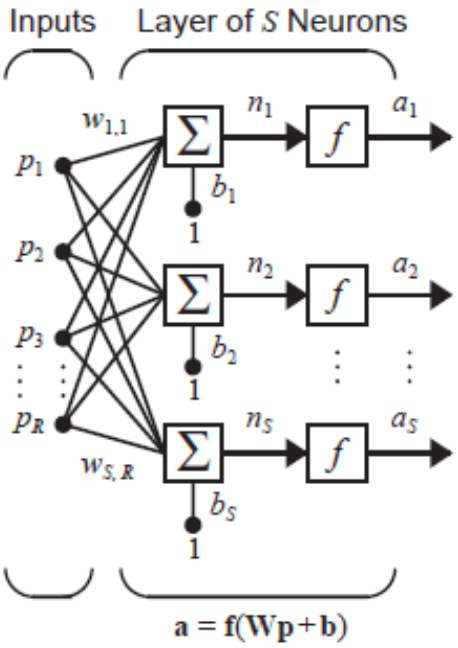
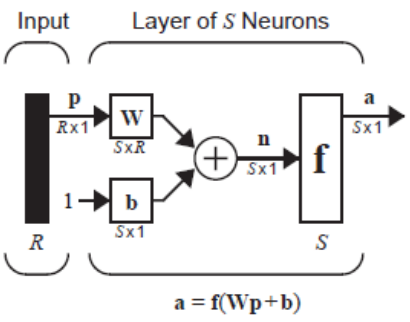



**Figure 3.17: Structure of a neuron model with an input layer and a layer of S neurons (to: Hagan et al., 2014).**

The neuron layer, also known as hidden layer, consist of a summer and transfer function. The input of the summer (Σ) is equal to the scalar input *p* multiplied by weight *w*, which is thus transferred to a scalar with a value equal to *w*p*. Further, the summer adds a bias *b,* hereby the output of the summer will be equal to: **W**p+b (see Figure 3.17 for an overview). The output of the summer, *n*, will become the input of the transfer function (*f*). The transfer function is a particular linear or nonlinear function of *n*. Since, many different transfer functions exist, a function is chosen to satisfy some specification of the problem that the neuron is attempting to solve (Hagan et al., 2014).

According to *Hagan et el.*, the suitable architecture of a neural network depends on the external problem specifications. In this case, the external problem specifications are related to the capacity assessment problem for weaving sections and merging lanes. First of all, the number of input and output values in the input and output layer are defined by the external problem specification. For example, if it is found that the capacity is influenced by four parameters and only a single (scalar) capacity value is desired as output, the input layer exists of input vector **p** with four columns and the output layer must provide one scalar value *a*.

Three other design variables are the number of hidden layers and number of neurons in the hidden layers, the use of biases and the desired transfer function. Since it is most probably that more than two layers will be used (Awad, 2004; Jiang & Adeli, 2004; Semeida, 2013), external problem specifications do not directly tell the number of required neurons and layers. Furthermore, neurons without biases could be used. However, the biases give the network an extra variable and makes the network more powerful (Hagan et al., 2014).




**(to: Hagan et al., 2014).**

Since neural network structures may become very complex, the structure is often displayed in an abbreviated notation. This notation is shown in Figure 3.18. Here the input layer is represented as a bar with input vector **p** with R columns. Further, the weight matrix, **W**, with *s* columns and *r* rows is displayed in one box, just like the bias. In the next paragraph the model development and model training will be discussed.

### 3.6.2 Model Development and Training

As stated before, the model training and development is an iterative process, which is also shown in Figure 2.3. Since the model architecture relies on the training results, no unambiguous model structure can be presented at this moment. Therefore, the model development and training process are described in this paragraph. Furthermore, the possible choices within this process will be elaborated. This will be done by means of the book *Neural Network Design* from *Hagan et al.*, 2014.

Obviously, the neural network development process is an iterative process. This process begins by the collection of data and pre-processing it to make training more efficient. The data and its eventual supplements concerning this research are already presented in previous chapter. After the data is selected, an appropriate network type and architecture should be chosen. Then a training algorithm for the network should be selected. The training algorithm should be appropriate for the network and the problem which is tried to be solved. After, or even during, the training process problems with the data, network architecture or training algorithm may be discovered. The above described process will be iterated until the network performance is satisfactory for its purpose. Hence, the model development process as shown in Figure 2.3 can be subdivided in three main aspects, namely: pre-training steps, network training and post-training analysis (Hagan et al., 2014).

#### *3.6.2.1 Pre-training Phase*

The pre-training steps do concern the selection of data, the data pre-processing and the choice of the network type and architecture. The available data and potential supplements are already described in previous chapter. However, during post-training analysis, it can be indicated if the training data was insufficient. After collecting the data, the dataset should be separated into a training, validation and testing datasets. The most common method for dividing the data is to randomly select a training, validation and test set out of the total dataset. Overall, this produces good results, however it is recommended to check the division for major differences between the sets. Major differences between the three sets could cause worse training results or overfitting.

Furthermore, the data should be pre-processed for an efficient training process. It is common to normalize the input data before applying them to the network. Three standard methods for data normalization are studied. The first method normalizes the data in a manner that the values for the variables fall into a standard range, mostly a range of -1 to 1 is chosen. The second normalization procedure is created to adjust the data so that it has a specified mean and variance. In a rare case, non-linear transformations are performed as a part of pre-processing stage.
The last step of the pre-training process is the selection of the basis network architecture and the selection of network architecture specifications.

First of all, the choice of a basic network architecture depends on the type of problem which is desired to solve. Four types of problems are identified, namely: fitting, pattern recognition, clustering and prediction. Since the problem of this research concerns fitting, the standard network architecture is the multilayer perceptron network. The fundament of this type of network is shown in Figure 3.17 and Figure 3.18. Moreover, these fitting network structure represents a feedforward neural network, which is used for this research. In most cases, *tansig* functions are used as transfer functions in the hidden layer for these types of neural networks. The *tansig* transfer function provides normalized output value between -1 and 1, which is preferred over *logsig* functions with an output range between 0 and 1. In the output layers, linear transfer functions are used.

When the basic network architecture is developed, the specifications of the architecture should be defined. The specifications of the architecture concern, among others, the number of neurons and the number of layers. The number of layers is not determined by the problem for fitting or pattern recognition purposes. It is stated that the standard procedure is to begin with a network with one hidden layer. It is also needed to select the number of neurons in each layer. The number of neurons in the output layer is equal to the size of the target vector. In case of this research, the size of the target vector is equal to one, namely: road capacity. The number of neurons in the hidden layer is not known prior to the training of the network. In this case, the standard procedure is to begin with more neurons than necessary and to use early stopping to prevent overfitting, which will be described in the next paragraph (Hagan et al., 2014).

#### *3.6.2.2 Training Phase*

The second phase of the model development and training procedure is the network training. This process includes initializing the weights, a selection of the training algorithm, the performance index and the criterion for stopping the training of the network. Since it is most probably that a multilayer network will be used, the initialization of the weights and biases will be done by setting the weights and biases to small random values. These are not set on zero to prevent that the network training will start at a saddle point on the performance surface (Hagan et al., 2014).

Then a training algorithm should be selected. For multilayer networks, gradient- or Jacobian-based algorithms are generally used. These algorithms can be implemented in both a sequential or batch mode. In sequential form the weight will be updated after each input is presented to the network. In batch form all inputs are presented to the network, and the total gradient is computed by summing the gradients for each input, before the weights and biases are updated. It is stated that the Levenberg-Marquardt algorithm is usually the fastest training method for multilayer networks. For example, this algorithm is also used in the research of *Adamowski et al.* (Adamowskiet al., 2012). Along with the training algorithm, the stopping criteria should be defined. This is required because the training error never converges to zero. For this reason, other criteria for stopping the training of the network are required. Stopping criterions are also used to prevent overfitting of the network. Several stopping criteria are suggested by *Hagan et al*., *2014*. The first stopping criteria is also the simplest one, namely stopping the training after a fixed number of iterations. Another stopping criterion is the norm of the gradient of the performance index. When this norm reaches a sufficiently small threshold, the training could be stopped. However, it is found that multilayer networks can have many flat regions in the training results, where the norm of the gradient will be small and therefore training results become inaccurate. The last discussed stop criterion is the point where reduction in the performance index per iteration becomes small. Performance indexes could be represented by the mean square error or the mean absolute error. For this stopping criterion it is also found that training could stop too early. Well elaborated examples of the usage of stopping are given by the research of *Yap et al.* and *Kadari et al.* (Kadariet al., 2015; Yap et al. , 2015).

However, even after the training algorithm has converged, post-training analysis could lead to the conclusion that the model needs to be modified and retrained. Moreover, it could happen that a single training run may not produce optimal performances. This could be caused by the possibility of reaching a local minimum of the performance surface. Therefore, it is recommended to restart the training several times with different initial conditions and select the best performing network (Hagan et al., 2014). Since training a neural network is involved with two stochastic elements, every training run the model will perform different. These stochastic elements are the initialisation of weights and the division of training, testing and validation set (Beale et al., 2018). Every training run, different initial weights will be generated next to varying compositions of the training, test and validation sets. Nonteheless, these initial weights and composition of training, testing and validation sets can be set captive in case of sufficient performance.

#### *3.6.2.3 Post-training Phase*

The final phase of the model development and training process is the post-training analysis. This analysis is recommended to determine if the training was indeed successful. Many techniques are available for post-training analysis. *Hagan et al.* only discussed the most common ones.
First of all, an useful tool for analysing neural networks, which are concerning fitting problems, is by means of a regression analysis. First the trained output and corresponding target value will be plotted in a graph. From this graph it could be seen how well the training results represent the corresponding target values. Then data points that fall far from the regression line could be investigated. It could then be discovered if these so-called outliers did arise due to problems with the data. For example, these data points could be located far from the training data points. When this is the case, gathering more data in that specific region will improve model performance. Therefore, separate datasets with only a certain range of values can be tested. For example, the model can be tested supplementary with datasets containing the weaving configuration types separately. In this manner, it can be discovered if the model is performing worse for a certain configuration type.

The regression analysis should be performed on each subset (training, validation and testing dataset) individually, as well as the full data set. If differences between model performance on the subsets arise, overfitting or extrapolation could be the case. As stated above, the full dataset is separated in a training set (70%), validation set (15%) and testing set (15%). The training set will be used to calculate gradients and determine weight and bias matrices. The validation set is used to stop training before overfitting occurs. Finally, the test set is used to predict the future performance of the trained network. Since the post-training analysis should be executed for all sets separately, differences between network performance on subsets can arise. Moreover, one or more supplementary datasets can be created to evaluate model performance on datapoints which are not present in one of the three above described sets. In other words: the model can be tested on configurations which are not (logically) present in the main dataset which is created with FOSIM.

Four possible causes for malperformance or differences in model performance on subsets are presented by *Hagan et al.* First of all, one of the problems cannot be identified by means of the differences between the datasets, namely the problem that the network could have reached a local minimum. This problem can be solved by retraining the network with other sets of random initial weights. The other three problems can be identified by means of the earlier described differences. First of all, overfitting most probably occurred when the validation error is much larger than the training error. This can be solved by selecting another (often slower) training algorithm to retrain the network. Secondly, if the errors for all three sets are similar in size, but too large, it can be concluded that the network is not powerful enough to fit the data. This can be solved by increasing the number of neurons in the hidden layer. When this measure is also ineffective, the number of hidden layers should be increased. The last identified problem could occur if test errors are large in comparison to the training and validation errors. In this case extrapolation could have occurred. If the inputs of the testing set are outside the range of the training data, extrapolation will occur. This could be solved by supplementing the dataset with a broader range of values for independent variables.

Finally, it is stated that it is often useful to assess the importance of each element in the input vector. In other words, if an element of the input vector is not important for the output vector, it can be eliminated to improve the computation speed and preventing overfitting. *Hagan et al.* explained that a sensitivity analysis computes the derivatives of the network response with respect to each element in the input vector. For this research this could be weaving section length, weaving ratio et cetera. If the derivative with respect to a certain input element is small, the input can be removed from the input vector. However, in this research the most influencing variables for road capacity will be elaborated prior to the model development. Therefore, it is expected that input removal will not occur that quickly for this research.

The paragraphs above illustrated that the model development and training process is iterative. Therefore, it is ineffective to present all the possibilities and potential methods within the choices that are required to make. Nevertheless, the steps for model development and training are elaborated to present the fundamentals of neural networks and its development. Moreover, some of the potential structures, training algorithms and other choices that should be made are identified and elaborated. In Chapter 4 to 6 the model development, training process and post-training analysis will be elaborated upon.

## 3.7 Conclusions within Theoretic Framework

In the theoretic framework current available literature related to the five pillars, as shown in Figure 3.19, of this research is elaborated. Moreover, the theoretic based research questions are answered in this chapter. In Figure 3.19, a summary of the literature study is present and a reference to the answered research questions is given.

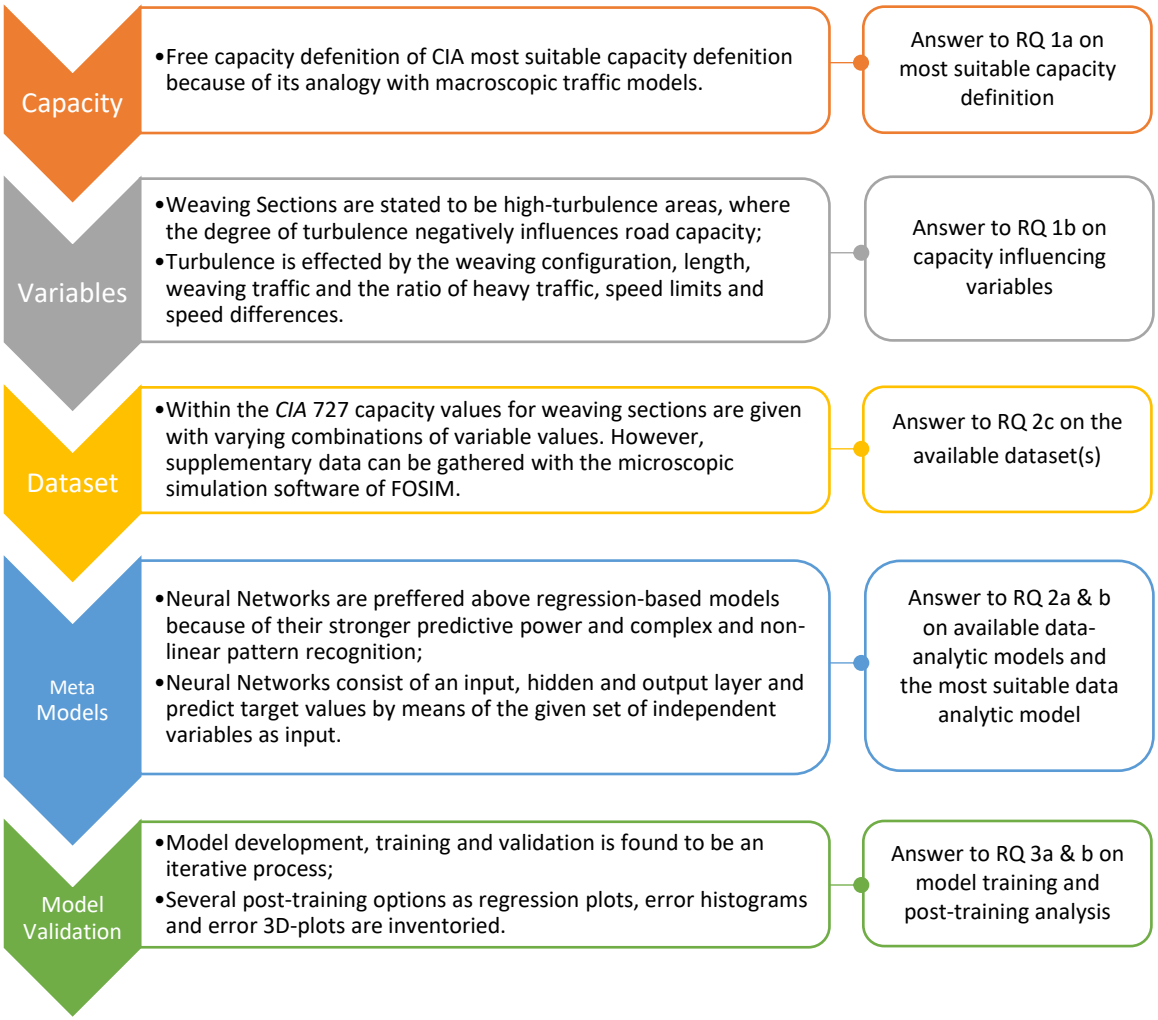



**Figure 3.19: Summary of the theoretic framework for the five pillars of this research and a reference to the answered research questions (RQ).**

In the next chapter the pre-training process of the model estimation is discussed. Since the capacity definition is already selected in this chapter, this component of the pre-training process is not discussed in the next chapter. Then, the pre-training process comprise, in this case, the collection and pre-processing of the dataset which is used to train the meta model. To do so, the significant variables influencing capacity will be analysed quantitatively. Then a comparison is made between the currently available dataset (*CIA*) and the real-case weaving section (*NRM-West*), which indicates if and how supplementary data is needed to be gathered.

# 4 Pre-training Phase

In the pre-training phase of the model development, the collection and pre-processing of the data used to train and validate the neural network will be presented. Figure 4.1 shows that the pre-training phase comprise a quantitative research to the optimal capacity definition and significant variables. Moreover, the available datasets - one for training the network and one containing real-case configurations to validate the network - are inventoried and the process of gathering supplementary data is discussed.

Since the optimal capacity definition for use in the meta model is already defined, the next paragraph discusses the significant variables. After this chapter on the pre-training phase, the final dataset consisting of dependent (capacity values) and independent (inventoried significant variable) variables, which are used in the training and post-training phase is known.

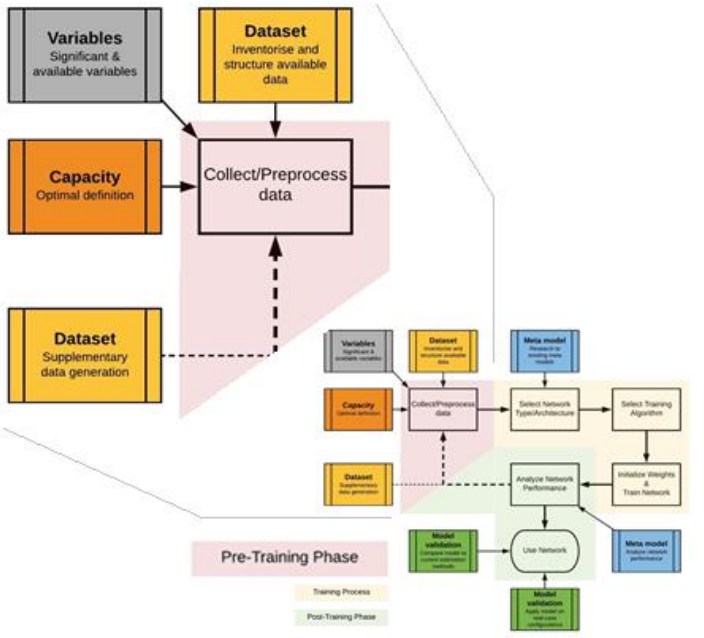



**Figure 4.1: (Iterative) Model development process as already shown in Figure 2.3, with the components included in the pre-training phase enlarged.**

## 4.1 Significance of Independent Variables

In the theoretical framework (Chapter 3), the phenomena turbulence, traffic-breakdowns and independent variables which influences capacity (dependent variable) are discussed. This inventory was made using the method of literature review. In this paragraph, a more quantitatively research on the capacity influencing variables is made using the currently available dataset and FOSIM simulations. First, the weaving specific variables will be discussed. Finally, the final set of independent variables which could be considered within the meta model will be elaborated.

### 4.1.1 Weaving Section Configuration

The weaving section configuration includes the geometry and lane division of the weaving section. A more delimited definition was given in the *Highway Capacity Manual*: 'The organization and continuity of lanes in a weaving segment, which determines lane-changing characteristics.' (Transportation Research Board, 2000).

First of all, the number of total lanes of the weaving section has a positive relation with the capacity. This makes sense, as the more lanes a weaving section has, the more space is available for vehicles to pass the weaving section per hour. In Figure 4.2 the relation between the number of total lanes and capacity is shown. In this figure the capacity is depicted on the y-axis, where the number of lanes of *origin 1* is shown on the x-axis. Here it can be seen that the capacity indeed increases as the number of lanes increases.

Furthermore, the configuration type was stated to have an influence on road capacity. The configuration type categorized weaving sections based on the number of required lane changes that are needed to change direction (see Figure 3.6). In Figure 4.8, where the influence of the volume ratio on road capacity is shown, it can be seen that the capacity decreases less under an increasing volume ratio for a type B configuration than for a type A or C configuration. This figure endorses the statement in the *HCM* that capacity is more sensitive to changes in the volume ratio for configuration types A and C than for configuration type B (Transportation Research Board, 2000). The reason for this is that less lane changes are required to change direction for vehicles. Therefore, it can be concluded that the configuration type does influence capacity.

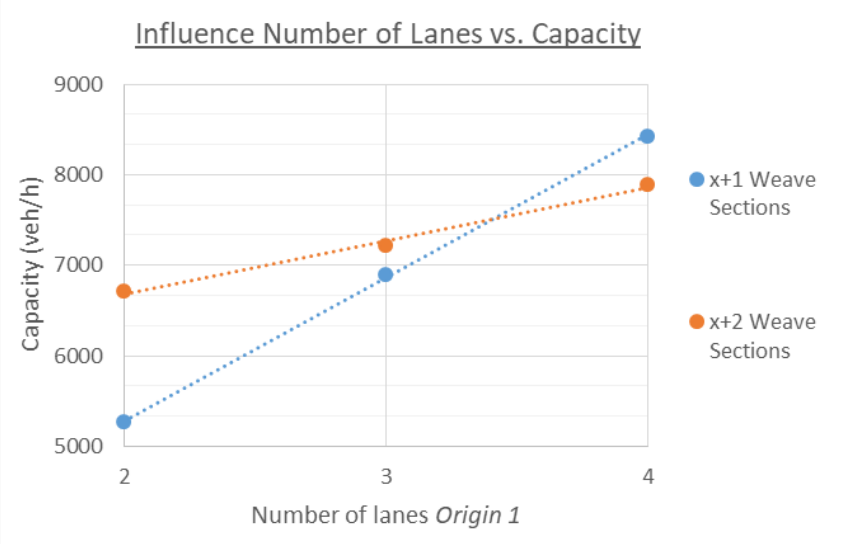



**Figure 4.2: Capacity values for weaving sections with a varying number of total lanes. The orange datapoints represent x+2 weaving sections (2+2, 3+2 and 4+2 configurations) and blue datapoints represent x+1 weaving sections (2+1, 3+1 and 4+1), with the number of lanes of origin 1 (x) on the x-axis.**

Finally, the weaving configuration and the division of lanes do influence the speed difference between the weaving and non-weaving vehicles on the weaving lanes. In Figure 3.12 a visual overview of this theoretical influence was already shown. With the available data and FOSIM simulations this effect is made quantitative. The speed differences on the weaving lanes determined using OFSIM simulations are displayed in Figure 4.3. The speed differences on the weaving lanes for the 1+2 configuration are approximately equal to 8 kilometers per hour (blue and orange line), while the speed differences on the weaving lanes for the 2+1 configurations are negligible (yellow and grey line).

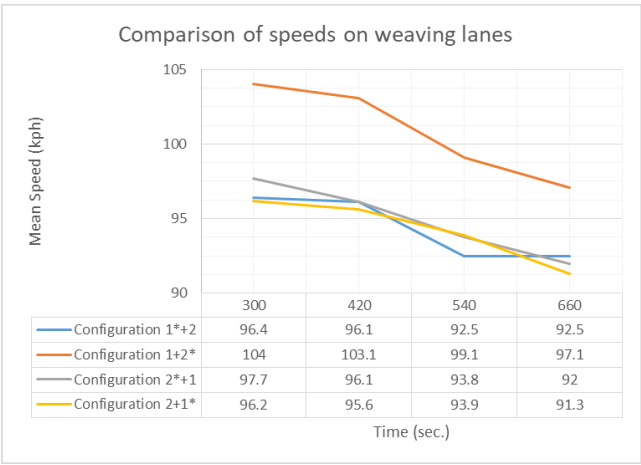



**Figure 4.3: Comparison of mean speeds on weaving lanes derived from FOSIM simulations. The configurations are shown in *Figure 3.12*.**

These speed differences implicate that the capacity values for 1+2 configurations are lower than for 2+1 configurations, since it was found that speed differences hinder the lane-changing process (Hidas, 2005). From the FOSIM simulations these statements are confirmed since the capacity values for the 1+2 configuration structurally lie lower than capacity values for 2+1 configurations, which is shown in Table 4.1. Furthermore, it can be noticed that the effect of the speed differences decreases when the ratio of heavy

| VR | %HT | Capacity | |
|---|---|---|---|
| | | *2+1 > 2+1* | *1+2 > 1+2* |
| 0.33 | 5 | 6594 | 6138 |
| 0.33 | 15 | 5712 | 5424 |
| 0.33 | 25 | 4974 | 4932 |
| 0.67 | 5 | 5802 | 5364 |
| 0.67 | 15 | 4998 | 4830 |
| 0.67 | 25 | 4590 | 4386 |

**Table 4.1: Capacity values for the above described 2+1 and 1+2 configurations with several values for the Volume Ratio (VR) and ratio of heavy traffic (%HT).**

traffic increases. This can be explained by the fact that trucks do slow down the traffic, which homogenize traffic flow.

To sum up, the weaving configuration do significantly influence weaving section capacity in three different manners, namely: the total number of lanes, the configuration type (which is determined by the lanes from the origins and destinations) and the division of lanes.

### 4.1.2 Weaving Section Length

In Paragraph 3.2.3.2 the influence of the weaving section length on capacity is already elaborated theoretically. In the literature, it was found that an increasing length has a positive influence on weaving section capacity. However, with an increasing length the impact on capacity itself will decrease. First, the theoretical statements that drivers tend to change lanes as quickly after the gore as possible will be verified. This will imply that increasing the length for an already long weaving section, will not have that large influence anymore.

Within FOSIM a merging configuration is constructed with a length of at least 1000 meter whereby a detector is placed every 50 meters. It should be noted that lane change areas are implemented directly after the gore. Furthermore, the traffic flow on the merging lane is 15% of the total traffic flow on the freeway mainline. In Figure 4.4, the above described configuration is shown.

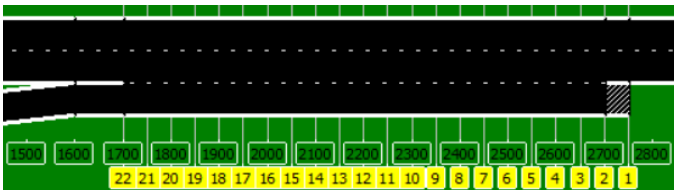



**Figure 4.4: Overview of the FOSIM configuration of a 1000-meter merge with detectors every 50 meters.**

Every time period within the simulation the detectors give the traffic flow per lane as output. In this manner, traffic intensities on the merging lane can be analysed per 50-meters segment. After several simulation runs the results of the detectors are extracted.

In Figure 4.5 the results of the simulation are shown. In this figure, the percentage of merged traffic over the distance from the gore is shown. It can be seen that approximately 60% of the traffic does perform the required lane change to the freeway mainline at the first 50 meters of the merging lane. After 550 meters all the traffic has merged. Moreover, at lower traffic intensities (t=300), the traffic on the merging lane do perform their lane changes earlier. On the other hand, at higher traffic intensities (t=1140), traffic on the merging lane uses more (length) of the merging lane before performing a lane change.

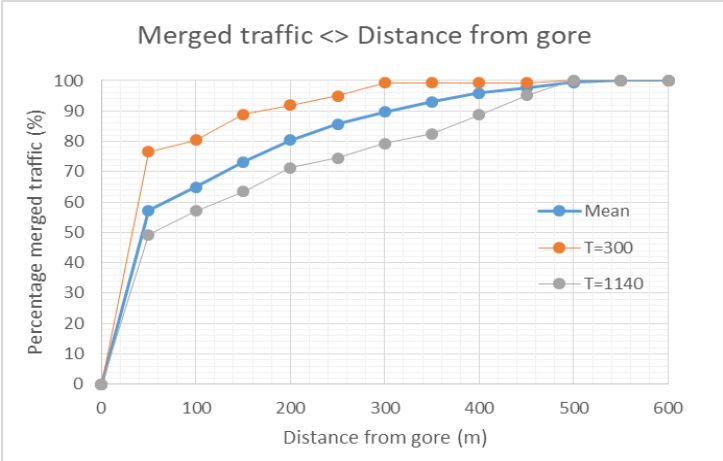



**Figure 4.5: Overview of the percentage merged traffic versus the distance from the gore. Where the orange line (simulation time = 300) represents low traffic intensities, the grey line (simulation time = 1140) represents high traffic intensities just before congestion occur and the blue line represents the mean (all time periods within the simulation).**

The results of the simulation support the two main statements from the literature, namely: an increased configuration length has less impact on the capacity. Since all traffic on the merging lane of 1000 meter has merged before 550 meter, a longer merging lane than 550m will not improve capacity that much. Moreover, from Figure 4.5 it can be concluded that a longer merging lane has an higher impact on capacity when traffic intensities are high. This can be declared by the fact that

traffic is denser and less sufficient gaps are initially available for a lane change. Once more, it should be noted that the simulation outputs depend on the implemented lane change behaviour in the simulation software. Nevertheless, the simulation study shows that the driving behaviour coincides with assumptions in literature.

Next to the simulation study in FOSIM, the relation between capacity and configuration length is analysed by means of the available and gathered data. In Figure 4.6 the relation between capacity and configuration length is shown for three different configurations, which were all derived from FOSIM simulations. This figure shows that capacity, in fact, increases when the weaving length increases. Moreover, it is shown that the capacity increases less quickly when the configuration length increases. For that reason, adding a logarithmic trend line to the data points results in a higher $R^2$-value than adding a linear trend line.

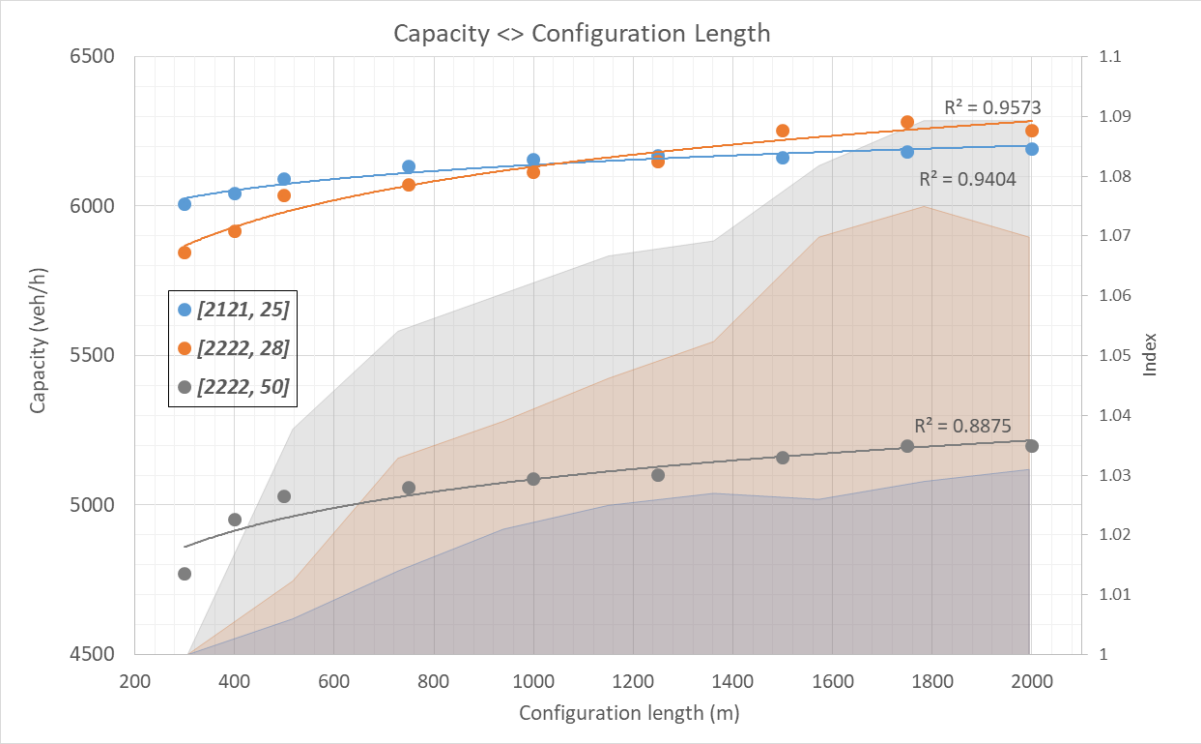



**Figure 4.6: Overview of the relationship between capacity and weaving length for three different (symmetric) configurations with varying volume ratios. Namely configuration 2+1 (VR = 25), configuration 2+2 (VR = 28 and 50). On the left axis the capacity values are shown. On the right axis (stacked graphs) the index of capacity values is shown per configuration, where the lowest capacity at length = 300m has index 1. Hence, when the index of a configuration is 1.03 at a length of 2000m, the capacity increased with 3% compared to the configuration with length of 300m.**

From this figure it can also be seen that the length has more positive impact on capacity values for more complex configurations than for less complex configurations. The blue data point of the 2+1 configuration and its corresponding (logarithmic) trend line is less steep than the two trendlines for the 2+2 configurations. From the percental capacity increase (indices) on the right axis, it can also be seen that the capacity increase for the 2+1 configuration is less high than for the 2+2 configuration with the approximately the same volume ratio. Namely, the capacity increase for 2+2 configurations is approximately seven to nine percent while the capacity increase for 2+1 configuration is 3 percent. Moreover, it is found that an increase of length has more positive impact for configurations with a higher volume ratio (grey data points) than for configurations with a lower volume ratio (orange data points).

This corresponds with the statements that traffic occupies more space to perform lane changes at higher traffic intensities (higher volume ratio implies more traffic on the weaving lanes). Since more lane changes (desired) are needed to execute, an increased length gives the drivers more time for gap searching an adjusting speed.

The quantitative analysis verifies the three main statements from the literature, namely: the configuration length has a positive impact on road capacity, an increasing length decreases the impact of configuration length on capacity and at higher traffic intensities the length has more influence on capacity than at low traffic intensities.
In short, the configuration length significantly influences road capacity. Moreover, from Figure 4.6 it can be concluded that no identical relation for length and capacity do hold for all configurations together. For that reason, it will be ineffective to develop a mathematical expression between capacity and configuration length.

### 4.1.3 Traffic Composition

In Paragraph 3.2.3.3 it is found that the ratio of heavy traffic, the only variable concerning traffic composition that can be reasonably taken into account, negatively influences capacity. Figure 4.7 shows that the ratio of heavy traffic is indeed a significant variable which influences capacity. In this figure capacity values for several configurations with a varying ratio of heavy traffic are shown. The figure shows that, as expected, road capacity decreases when the ratio of heavy traffic increases. As assumed in the theoretic framework, a perfect linear relation between the ratio of heavy traffic and capacity does not hold. Moreover, the trend lines are not linear which is implicitly assumed when using the reduction factor methods from CIA (and HCM). However, it should be noted that, when excluding the data points with the ratio heavy traffic of zero and thirty-five, the trend line seems quite linear.

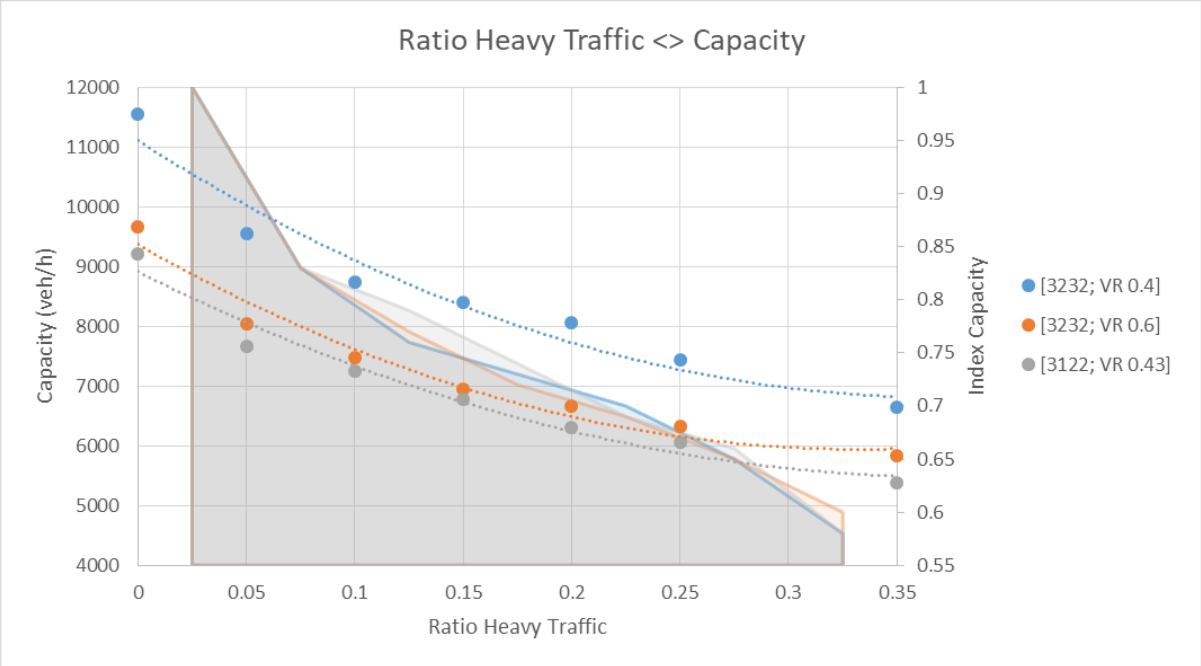



**Figure 4.7: Overview of capacity values for several weaving configurations (3232 and 3122) with varying ratios of heavy traffic. On the left axis the capacity values are expressed as an index (percent) where the capacity values with index 1 representing configurations with the 0% heavy traffic.**

Since the ratio of heavy traffic influences road capacity significantly and no general equation or reduction factor holds for all configurations jointly, the ratio of heavy traffic is included in the meta model.

### 4.1.4 Ratio of Weaving Traffic

In Paragraph 3.2.3.4 it was concluded that weaving traffic increases turbulence which negatively influence road capacity. In this section simulation studies and data analysis will be executed to quantify the effect of the weaving traffic on the road capacity.

The simulated CIA-dataset reveals that capacity decreases when the volume ratio increases (Figure 4.8). In literature, this is explained as follows: a higher volume ratio implies more weaving traffic and thus more turbulence which results in a decrease of capacity. In the HCM it is stated that the capacity is more sensitive to changes in the volume ratio for configuration types A and C than for configuration type B (Transportation Research Board, 2000). This can also be seen in this figure, since the (yellow) trendline through the yellow datapoints (type B configuration) decreases less than the other trendlines (A&C configurations). The yellow, orange and blue data points and trend lines are based on the weaving ratios (and thus volume ratios) included in CIA, which assumes that the weaving ratios are proportional to the ratio of the number of lanes. For the black data points, the weaving ratios are not proportional to the ratio of the number of lanes. Here, the volume ratio is increased due to another combination of weave ratios (as shown in Figure 4.9). However, one remarkable data point is shown in Figure 4.8 which do not coincide with the general assumption that the capacity decreases when the volume ratio increases. The capacity of this black data point with a volume ratio of 0,2 is lower than the capacity of the yellow data point with volume ratio 0,3.

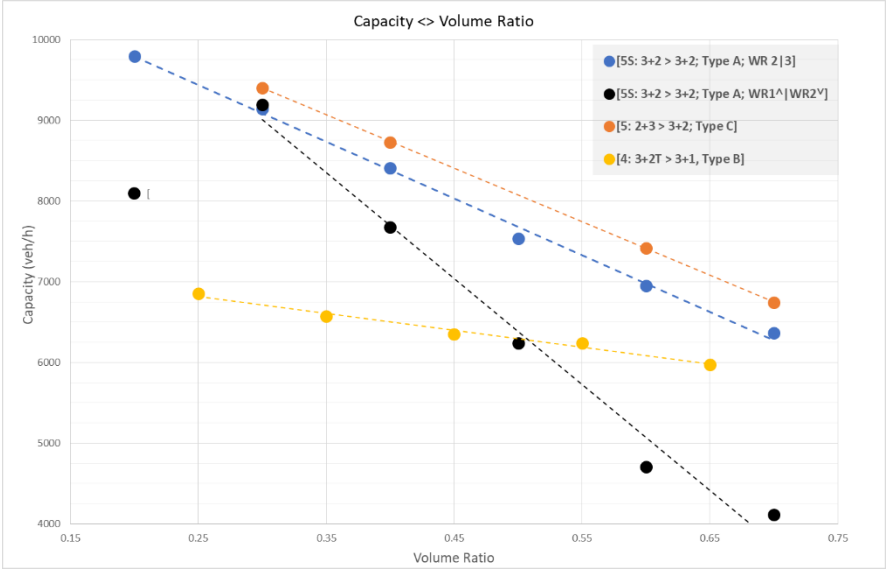



**Figure 4.8: Overview of the influence of the volume ratio on road capacity for weaving sections. Only variations in the volume ratio are made (i.e. length and ratio heavy traffic are constant).**

To further investigate this phenomenon, in Figure 4.9, all combinations of weaving ratios for configuration *[5S: 3+2 -> 3+2]* and their corresponding capacity values are shown. The two arrows in this figure represent the increase of volume ratio as shown in Figure 4.8.

Capacity is lower for this combination of weaving ratios ($0 < WR1 < 0.2$; $0.4 < WR2 < 0.5$) because the demand for destination 1 is too high: 50% percent of the traffic from origin 2 weaves to destination 1 whereas all the traffic from origin 1 has destination 1 as their destination. Figure 4.9 reveals that this phenomenon primarily occurs in situations with high weaving rations for one direction in combination with low weaving ratios for the other direction, which implicates that the demand for a destination is, in these cases, more determinant for the capacity than the turbulence caused by the lane changes.

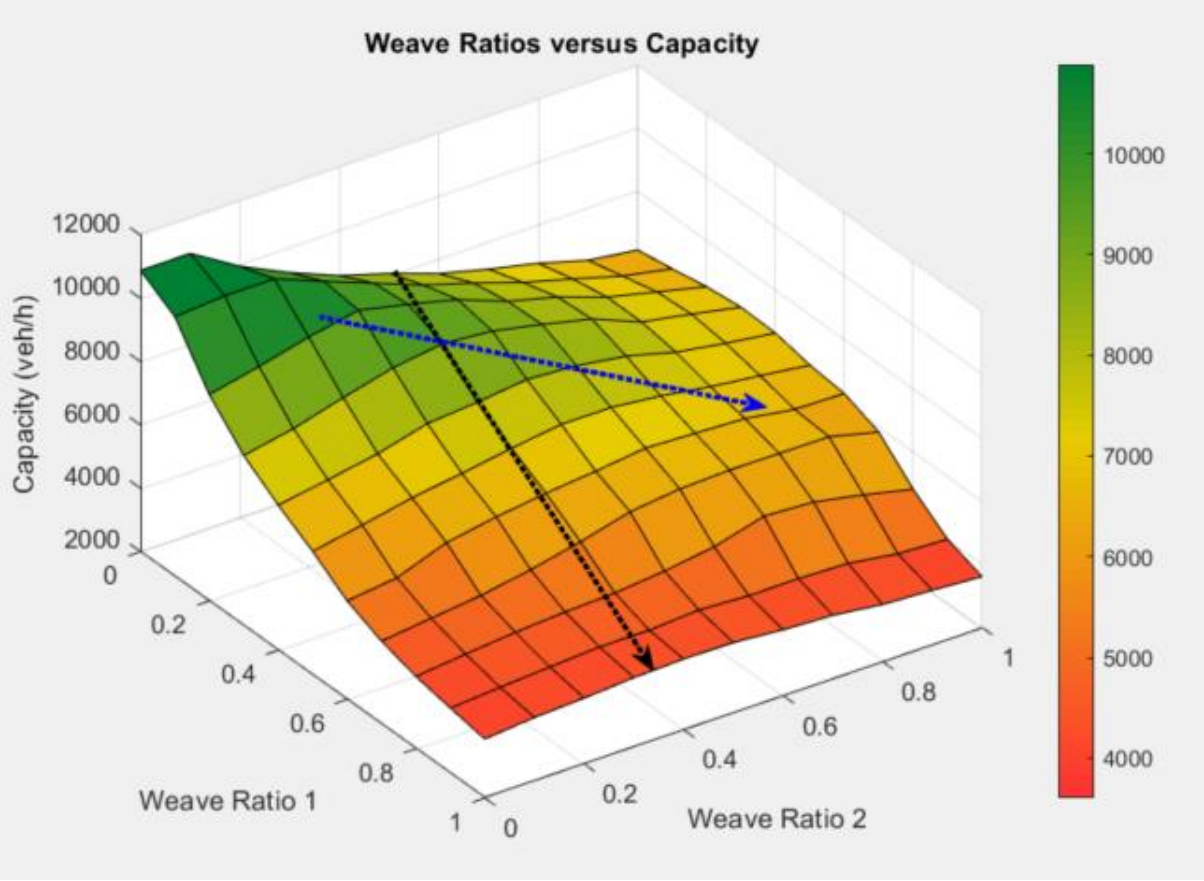



**Figure 4.9: 3D plot of the capacity values, on the z axis, for the symmetrical 3+2 configuration with on the weaving ratios for origin 1 and 2 respectively on the x and y axis.**

### 4.1.5 Allowed Speed

Since quantitative research is already present in *CIA*, it is known that the capacity reaches an optimum at speed limits of 100 kilometres per hour. This paragraph investigates the impact of different configurations and weaving ratios on this optimum. In Figure 4.10 the impact of the speed limit on capacity is shown for six different configurations. Since capacity values for *2222* configurations and all other configurations differ substantially, the capacity values are divided on two axes. This figure shows that there is actually an optimal speed which realises the highest capacity for a certain configuration.

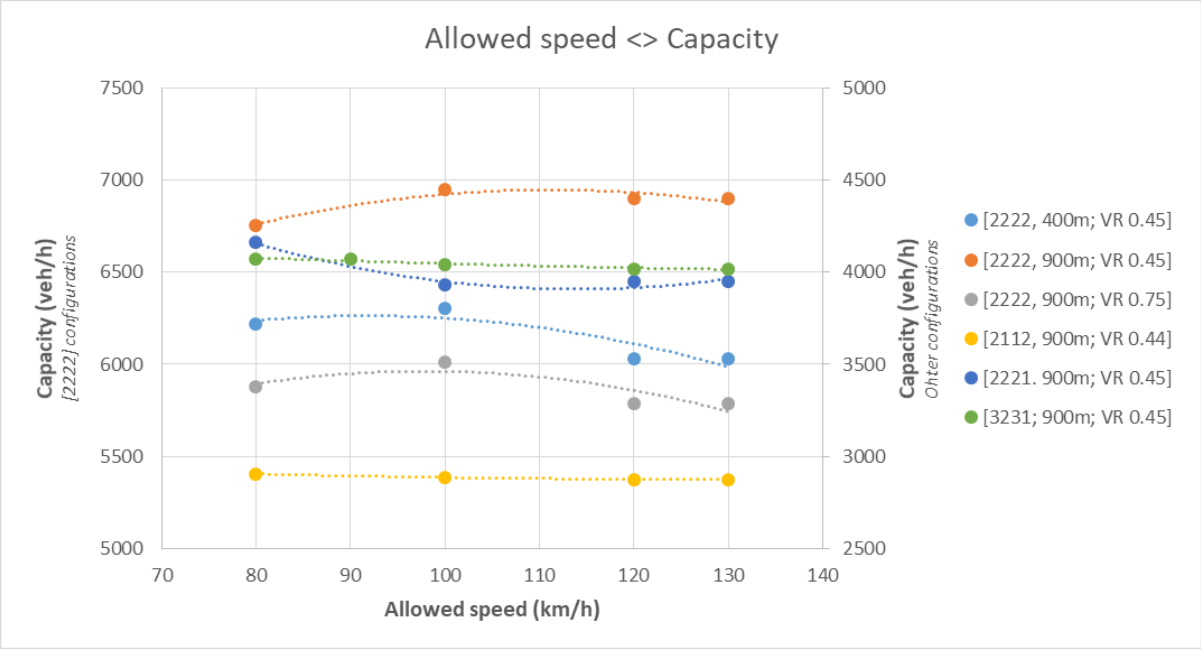



**Figure 4.10: Relation between the allowed speed and capacity according FOSIM simulations. On the left vertical axis the capacity values for [2222] configurations are shown (blue, orange and grey) and on the right vertical axis, capacity values for all other configurations are shown.**

Moreover, the impact of the speed limit on capacity is related with other capacity influencing variables. For the configurations with a taper, which are the dark blue and green datapoints, the capacity is the highest at a maximum speed of 80 kilometres per hour instead of 100 kilometres per hour. This is explained by the fact that vehicles do have less time for cooperation and anticipation due to the taper merge. Here, a lower speed will homogenize traffic flow resulting in improved opportunities for lane-changing.

Furthermore, the speed limit has a limited effect on long configuration compared to short configurations. This could also be caused by the fact that on longer weaving sections, vehicles have more time for cooperation and anticipation, while on short weaving sections a reduced allowed speed will homogenize traffic.

### 4.1.6 Speed Differences

In the theoretical framework it was found that speed differences between weaving and non-weaving vehicles do negatively influence capacity (Hidas, 2005; Marczak et al., 2013). Furthermore, a quantitative analysis of the effect of the weaving configuration on speed differences and thus capacity is elaborated in Paragraph 4.1.1.

Speed differences between weaving and non-weaving traffic can also arise due to sharp turns on (single) weaving lanes. For example a 2+ 1 weaving section with a sharp turn on the single lane or cloverleaf weaving sections.
As shown in Figure 4.3, the mean speed of the continuous traffic on the most right lane is approximately equal to 98 km per hour. This causes the speed differences between the weaving and continuous traffic to be the lowest when the speed reduction factor (to simulate slower traffic on specific lanes (i.e. sharp turns) in FOSIM) on the single weaving lane is between 0.8 and 0.9. A too low or too high speed reduction factor will cause too slow or too fast weaving traffic. In other words: the merging process is most efficient when the speed differences between vehicles are low, which can be realized by an optimal speed reduction factor.

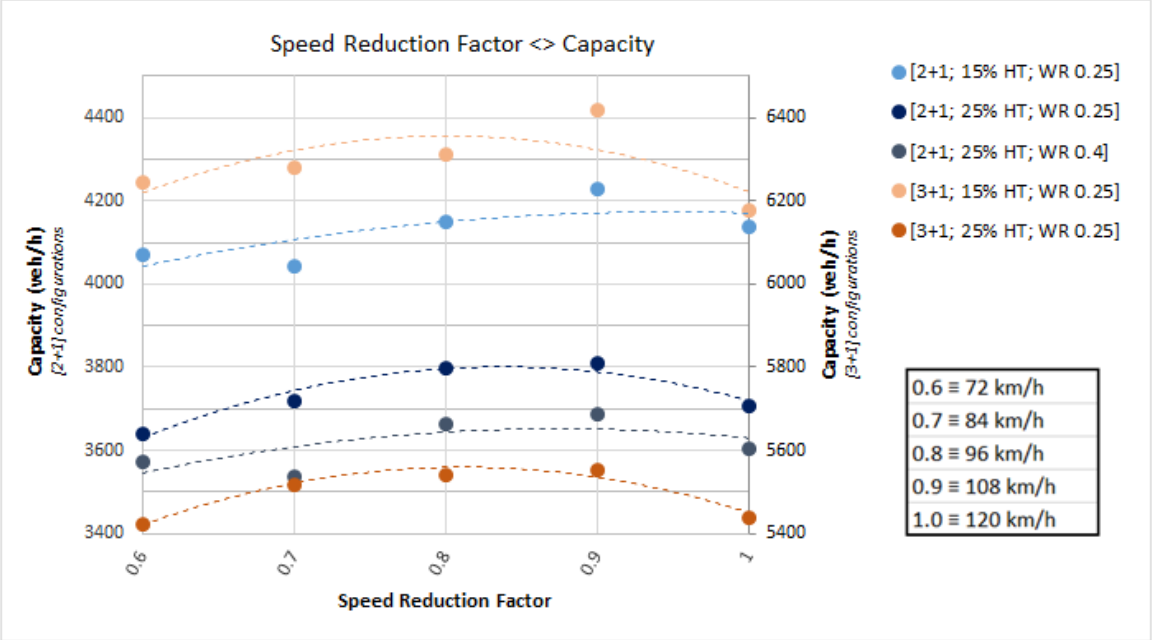



**Figure 4.11: Graph with the speed reduction factors (and actual speeds in the bottom right) and capacity values. On the right axis the capacity values for the 2+1 configurations are shown and on the left axis the capacity values for the 3+1 configurations. The abbreviations in the legend represent respectively the ratio heavy traffic (%HT) and the ratio of merging/weaving traffic (WR).**

In this figure, Figure 4.11, the capacity values for two different types of merge configurations are shown. On the left axis the capacity values for 2+1 configurations (blue data points) are shown where the right axis shows for capacity values for 3+1 configurations (orange data points).

Polynomic trend lines provide the best fit to the data points. Since the mean speed of the continuous traffic on the most right lane is approximately equal to 98 kilometers per hour, the capacity will theoretically be the highest at a speed reduction factor between 0.8 and 0.9 (≡ 96 – 108 km/h).

According to Figure 4.11, the highest capacity per configuration lies indeed around a speed reduction factor equal to 0.9 on the weaving lane. Furthermore, at lower ratios of heavy traffic, the capacity with a speed reduction factor of 0.9 is higher compared to configurations with a higher ratio of heavy traffic. This is because a higher share of trucks on the most right lane decreases the mean speed of the traffic on this lane.

### 4.1.7 Division of Traffic Flows

This variable is not inventoried in the theoretical framework in Chapter 3, however during the research it is found that this variable does influence road capacity. CIA assumes that the traffic flows over the origins are equally divided. In the case of the weaving section shown in Figure 4.12, according to CIA, the proportion of the traffic flows will be equal to 2:1. Since this is the ratio of the number of lanes from origin 1 to the number of lanes from origin 2.

This is logical since the number of lanes from origin 1 is equal to two and the number of lanes from origin 2 is equal to one. Hence CIA assumes: $Prop_{Int} = Lanes_{O1}/Lanes_{O2}$, where $Prop_{Int}$ represent the proportion of traffic flows, $Lanes_{O1}$ the number of lanes from origin 1 and $Lanes_{O2}$ the number of lanes from origin 2. However, in practice this statement of proportionality does not hold, which will be shown in Section 4.3.2. The reason is that asymmetric traffic intensities affect the amount of weaving traffic and thus the volume ratio.

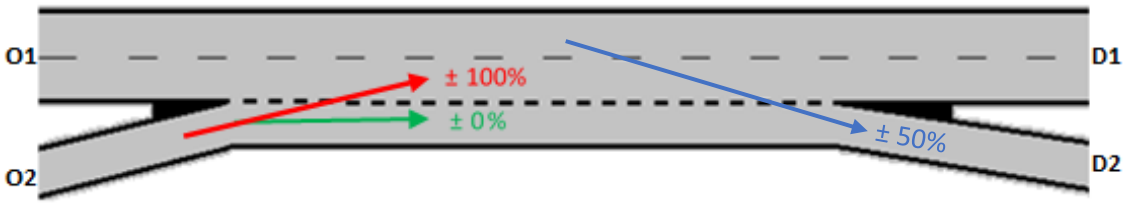



**Figure 4.12: Graphical representation of a symmetric 2+1 weaving section. Where it is illustrated that approximately 100% of the traffic from origin 2 change direction to destination 1.**

According to the assumptions of *CIA*, the volume ratio of this configuration will be calculated by: $VR = \frac{WR1*Traffic_{Intensity},O1 + WR2*Traffic_{Intensity},O2}{Total\ Traffic} = \frac{WR1*L_{O1}+WR2*L_{O2}}{L_{O1}+L_{O2}} = \frac{0,5*2+1*1}{2+1} = 0,66.$

However, when the proportion of traffic flows is not equal to the proportion of lanes (2:1), the volume ratio will change while the two weaving ratios (WR1 and WR2) are still constant. Assume that less traffic is present from origin two, for this example the traffic flow from origin 2 will be six times lower. Then the proportion of traffic intensities will become 2:$\frac{1}{5}$ ≡ 2:0.2 ≡ 10:1.

This will change the volume ratio to: $VR = \frac{WR1*Traffic_{Intensity},O1 + WR2*Traffic_{Intensity},O2}{Total\ Traffic} = \frac{0,5*10+1*1}{10+1} = 0,55.$ Consequently, when the proportion of traffic flow decrease, the volume ratio will increase for this configuration. In Table 4.2 an overview of different configurations with deviating proportions of traffic intensities are shown. Here, it can indeed be seen that the volume ratio changes at deviating proportions of traffic flows. However, when the weaving ratios 1 and 2 are approximately equal, the volume ratio will not change despite deviating proportion of traffic flows.

| Lanes | WR1 | WR2 | Proportion Traffic Intensity | VR |
|---|---|---|---|---|
| **2121** | **50** | **100** | **2:1** | **0.66** |
| 2121 | 50 | 100 | 10:1 | 0.55 |
| 2121 | 50 | 100 | 1.5:1 ≡ 3:2 | 0.7 |
| **2121** | **50** | **10** | **2:1** | **0.37** |
| 2121 | 50 | 10 | 10:1 | 0.46 |
| 2121 | 50 | 10 | 1.5:1 ≡ 3:2 | 0.34 |
| **2222** | **50** | **50** | **2:2** | **0.5** |
| 2222 | 50 | 50 | 6:2 | 0.5 |

**Table 4.2: Overview of the effect of a deviating proportion of traffic intensities on the VR.**

Since FOSIM restricts the use of this variable in a certain manner, this phenomenon will be discussed more extensive in the paragraph on data gathering (Section 4.4) for reasons of readability.

## 4.2 Final set of Independent Variables

In this chapter a quantitative research was elaborated on the capacity influencing variables which were not excluded in Paragraph 3.2.4. All the above described independent variables do significantly influence capacity on weaving sections. Moreover, all variables can be implemented in FOSIM. In other words, supplementary data could eventually by gathered in case CIA did not include these variables (i.e. maximum speed, speed differences and the division of traffic flows).

However, the variables of maximum speed and speed differences are not present or structurally implemented in the macroscopic traffic model of OmniTRANS. First of all, the maximum speed is always set on 120 km/h on weaving sections (see Section 4.3.2). Although, speed limits in real-case are not always equal to 120 km/h (Rijkswaterstaat , 2018). Therefore, it is recommended to implement corresponding speed limits, because capacity values can vary approximately 5% as shown in Figure 4.10.

Next to the maximum speed, speed differences are also not implemented in the macroscopic traffic model. Speed differences mainly arise due to sharp turns located before the gore of the weaving section. Therefore, it is recommended to incorporate these speed reductions in the macroscopic

transport model. This can be done, for example, by coupling a sharp turn to a weaving section, whereby a speed difference between the two origins can be assumed. For now, these two variables are not considered in the neural network. However, it is recommended to implement these variables in macroscopic traffic models in the future, whereby they can be considered within the neural network.

A schematic overview of the final used variables was presented in Figure 4.13. With the set of significant variables defined combined with the currently available parameters in the macroscopic traffic models, an unambiguous answer for the first research question is given. It was namely found that the configuration length, ratio of heavy traffic, weaving configuration (which do hold the number of lanes for every origin and destination) and ratio of weaving traffic (which is composed by the two weaving ratios and the asymmetric traffic intensity) do significantly influence the value of road capacity for weaving sections and can be implemented in the meta model.

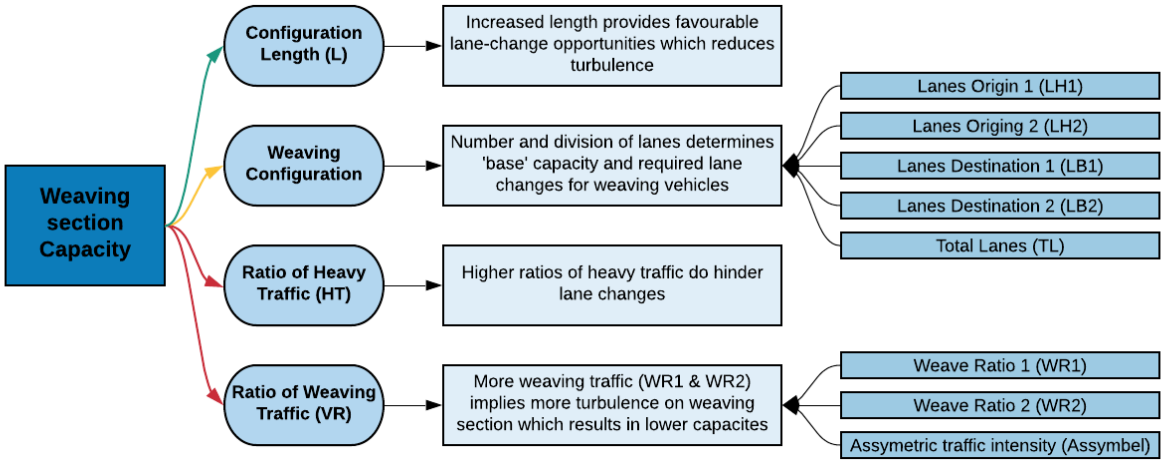



**Figure 4.13: Overview of the final set of variables and the manner in which they influence capacity values (green arrow: positive influence; orange arrow: optimum can be found; red arrow: negative influence).**

## 4.3 Available Datasets and Similarity

In this paragraph the currently available datasets will be presented. First of all, a short summary on the CIA dataset will be given. This dataset is used to train the neural network during the training phase. Furthermore, another dataset is used to validate the neural network on real-case configurations. In this research, it is assumed that the weaving configurations present within the NRM-West network, which is introduced in Chapter 2, are representative for all real-case configurations in The Netherlands.

Finally, this paragraph analyses the similarity between the two above described datasets. Based on the results of this analyse it was decided to gather supplementary data, since meta models are stated to be non-capable of extrapolating (Hagan et al., 2014).

### 4.3.1 CIA Dataset

In Paragraph 3.3 the available dataset of CIA is discussed. It was found that the dependent variables of this dataset are capacity values. These dependent variables are determined by four independent variables, namely: the weaving configuration, the ratio of heavy traffic, the weaving section length and the two weaving ratios (WR1 and WR2).

It should be noted that two of the significant variables, determined in Paragraph 4.1, are not included in this dataset. These two variables are the allowed speed and the division of traffic flows. The datapoints available in this dataset do all hold for configurations with a speed limit of 120 km/h. Since the allowed speed is not taken into account in the meta model (see Paragraph 4.2), no supplementary data with other values for the allowed speed are necessary to simulate. However,

In short, the currently available CIA dataset consist of capacity values of 27 different weaving configurations with varying values for the ratio of heavy traffic (5, 15 and 25%), length and weaving ratios.

### 4.3.2 Set of Real-case Configurations

In the post-training phase, the neural network, which is trained with the (supplemented) CIA dataset, is validated on real-case configurations. These real-case configurations are derived from the NRM-West network. It should be noted that the NRM dataset only contain values for the independent variables. Because of the validation purposes of these configurations, capacity values for these configurations are simulated with FOSIM, which will be discussed in Paragraph 4.4.

The NRM dataset consist of 363 weaving sections present in the study area the NRM-West model. In Table 4.3 the first seven weaving sections and their relevant variables from the dataset are shown for illustrative purposes. In this table it can be seen that variables as volume ratio, configuration codes and the presence of a taper are not present. However, these variables can be calculated and/or derived by means of other variables available in the transport model.

First of all, it can be noticed that the proportion of traffic flows over the origins is not equal to the proportion of lanes, which is assumed by CIA (Paragraph 4.1.7). The ratio between the traffic flows can simply be calculated by dividing the sum of all traffic from origin two by all traffic from origin one: $Int_{Prop} = \frac{\sum Int\ O2}{\sum Int\ 01} = \frac{Int\ O2D1 + Int\ O2D2}{Int\ O1D1 + IntO1D2}$, where *Int* $O_xD_y$ represents the traffic from origin *x* to destination *y*.

Since FOSIM restricts the maximum traffic flows per origin/lane, a factor should be calculated which expresses the traffic flow of the less dominant traffic flow compared to the actual maximum flow per lane in FOSIM. Exceeding this maximum traffic flow per lane will result in a simulation crash which makes batch simulation runs ineffective. This factor is called *Assymbel* and is calculated by dividing

| weefvakid | Int O1 D1 | Int O1 D2 | Int O2 D1 | Int O2 D2 | Lanes O1 | Lanes O2 | Lanes D1 | Lanes D2 | Total Lanes | % HT | Length (m) | % O2-D1 | % O1-D2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 3084 | 638 | 270 | 172 | 3 | 1 | 3 | 1 | 4 | 0.119859 | 800.506681 | 0.61085973 | 0.17141322 |
| 2 | 1430 | 1282 | 1107 | 805 | 3 | 2 | 2 | 2 | 4 | 0.213607 | 1154.4392 | 0.5789749 | 0.47271386 |
| 3 | 3199 | 156 | 2478 | 110 | 2 | 2 | 3 | 1 | 4 | 0.142124 | 788.42473 | 0.95749614 | 0.04649776 |
| 4 | 1521 | 1170 | 1933 | 1247 | 3 | 2 | 3 | 2 | 5 | 0.153581 | 720.12049 | 0.60786164 | 0.43478261 |
| 5 | 2534 | 1411 | 149 | 61 | 3 | 1 | 3 | 1 | 4 | 0.116984 | 752.564162 | 0.70952381 | 0.35766793 |
| 6 | 2564 | 381 | 271 | 74 | 2 | 1 | 2 | 1 | 3 | 0.090738 | 816.808224 | 0.78550725 | 0.12937182 |
| 7 | 2899 | 2154 | 90 | 50 | 4 | 1 | 3 | 2 | 5 | 0.149402 | 729.154468 | 0.64285714 | 0.42628142 |

**Table 4.3: Overview of the first seven weaving configurations in the NRM dataset. Where 'int' stands for the traffic flows (int O1 D1 = traffic flow from origin 1 to destination 1). Furthermore, the number of lanes for all origins, destinations and the total lanes are shown. Finally, the percentage heavy traffic (%HT), length and both weaving ratios can be seen (% O2-D1 ≡ WR2; % O1-D2 ≡ WR1).**

the proportion of traffic intensities by the proportion of lanes. $Assymbel = \frac{Int_{Prop}}{Lanes_{Prop}} = \frac{Int_{Prop}}{(\frac{Lanes_{O2}}{Lanes_{O1}})}$.

Where $Int_{Prop}$ is the proportion of traffic flows, $Lanes_{O2}$ the number of lanes from origin 2 and $Lanes_{O1}$ the number of lanes from origin 1.

With this *Assymbel* factor the less dominant traffic flow can be expressed compared to the more dominant traffic flow. For example, regarding the configuration discussed in Figure 4.12 with 200 veh./h from origin 2 and 1000 veh./h from origin 1, the *Assymbel* factor will be: $\frac{200/1000}{\left(\frac{1}{2}\right)} = 0.4$.

Using this factor, the traffic intensity in FOSIM for the less dominant origin is always set without exceeding the maximum traffic intensity FOSIM can handle. In this case, the traffic flow from origin 2 will be set 0.4 times the (maximum) traffic flow from origin 1.

The volume ratio can, for NRM configuration where traffic intensities are known, be calculated by dividing the sum of the weaving traffic by the total traffic.

$$VR = \frac{\sum Weaving\ Traffic}{\sum Total\ Traffic} = \frac{Int_{O1,D2} + Int_{O2,D1}}{Int_{O1,D1} + Int_{O1,D2} + Int_{O2,D1} + Int_{O2,D2}}$$

For configurations where the traffic intensities are not known, all non-NRM configurations, the volume ratio can be calculated with the formula given in Paragraph 3.2.3.4. However, with the extra variable of *Assymbel,* the volume ratio cannot be calculated by means of the ratio of lanes.

$$VR = \frac{Lanes_{O1} * (WR_1 + \frac{WR_2 * Assymbel}{Lanes_{Prop}})}{Lanes_{O1} * (1 + Assymbel * Lanes_{Prop})}$$

By means of the available information on the number of lanes per origin and destination and the number of total lanes, the presence of a taper and the minimum required lane changes can be calculated. To illustrate: for the second configuration in Table 4.3 the number of incoming lanes is equal to 5 (3+2) while the number of total lanes in the weaving section are equal to 4.

This implies that a taper is present upstream of the weaving section. Based on these required lane changes also the configuration types as defined in the *HCM* can be derived and applied in case the model performance will be increased by adding this parameter.

Finally, the configuration code is an assembly of all relevant independent variables to create a unique identification parameter for every configuration.

| weefvalid | Lane Code | Lanes O1 | Lanes O2 | Lanes D1 | Lanes D2 | TotalLanes | %HT | Length (m) | Speed | WR 2 | WR 1 | VR | Speed Reduction | Assymbol | Config Code |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 3131 | 3 | 1 | 3 | 1 | 4 | 0.12 | 801 | 120 | 0.61 | 0.17 | 0.22 | 1 | 36 | [3131,801,61,17,12,120,36] |
| 2 | 3222 | 3 | 2 | 2 | 2 | 4 | 0.21 | 1154 | 120 | 0.58 | 0.47 | 0.52 | 1 | 106 | [3222,1154,58,47,21,120,106] |
| 3 | 2231 | 2 | 2 | 3 | 1 | 4 | 0.14 | 788 | 120 | 0.96 | 0.05 | 0.44 | 1 | 77 | [2231,788,96,5,14,120,77] |
| 4 | 3232 | 3 | 2 | 3 | 2 | 5 | 0.15 | 720 | 120 | 0.61 | 0.43 | 0.53 | 1 | 176 | [3232,720,61,43,15,120,176] |
| 5 | 3131 | 3 | 1 | 3 | 1 | 4 | 0.12 | 753 | 120 | 0.71 | 0.36 | 0.37 | 1 | 15 | [3131,753,71,36,12,120,15] |
| 6 | 2121 | 2 | 1 | 2 | 1 | 3 | 0.09 | 817 | 120 | 0.79 | 0.13 | 0.2 | 1 | 24 | [2121,817,79,13,9,120,24] |
| 7 | 4132 | 4 | 1 | 3 | 2 | 5 | 0.15 | 729 | 120 | 0.64 | 0.43 | 0.43 | 1 | 12 | [4132,729,64,43,15,120,12] |

**Table 4.4: Overview of the first seven configurations in the NRM dataset. Here only the relevant variables per configurations are maintained. Where WR stands for the weaving ratios and VR for the volume ratio.**

### 4.3.3 Similarity CIA and Real-cases

In this paragraph the representativity between the configurations in the CIA and NRM will be elaborated. The purpose of this analysis is to discover underrepresented ranges of values in CIA which are needed to be simulated.

First of all, Figure 4.14 shows differences and similarities between the CIA and NRM dataset for the number of lanes and configuration length. The histogram reveals that the dataset of CIA is more extensive for configurations with relatively a high number of lanes. For the configurations in CIA it holds that more than 50% of the data points do represent configurations with 5 or more lanes. For the NRM configurations only 15% of the total number of configurations consist of 5 or more lanes. In the NRM dataset, namely more than 70% of the total number of configurations do consist of three or four lanes (i.e. 2+1, 2+2, 3+1 configurations). Furthermore, the dataset of CIA does not match with the NRM configurations regarding configuration length. Firstly, CIA do not provide data points with a higher length than 1100 meters. However, roughly fifteen percent of the NRM configurations have a length larger than 1100 meters. One can expect that a good performing meta model should correctly predict capacity values for these configurations since at large lengths capacity does not increase that much anymore, which is already shown in Figure 4.6. Nevertheless, the only configurations with short lengths in CIA are 2-lanes configurations (1+1 configurations). For configurations with more lanes, the NRM dataset contains shorter configurations than the configurations in CIA, indicated by the much higher share of lower lengths (yellow-pigmented) compared to the higher lengths (green-pigmented). This holds for roughly 25% of all NRM configurations. Especially for these short lengths, it will be hard for a meta model to predict capacity values, because in these situations, capacity is highly influenced by the configuration length (Figure 4.6). For this reason, supplementary data for shorter (and eventually longer) configurations lengths should be added to the dataset.

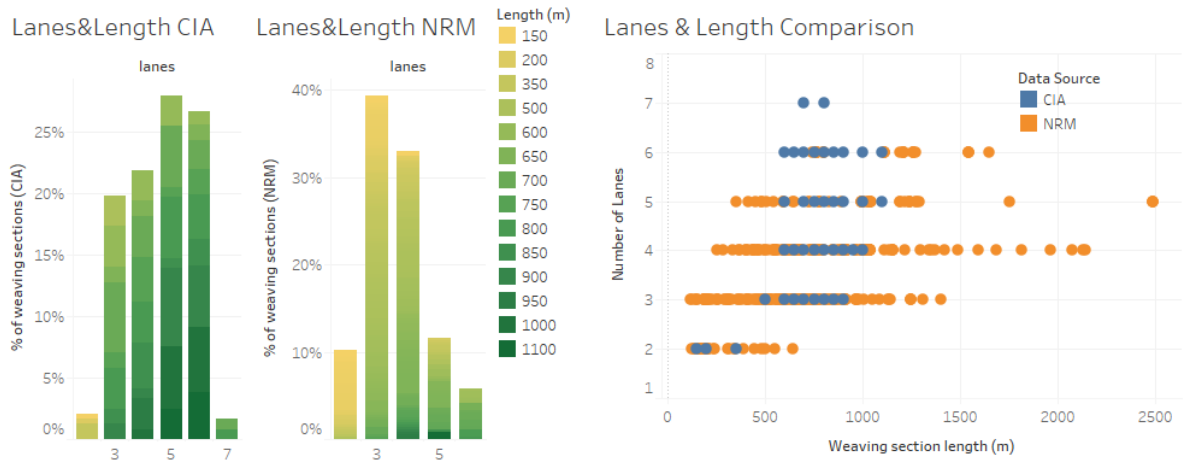



**Figure 4.14: Overview of the CIA and NRM dataset related to the number of lanes and configuration length. In the two histograms the relative share of configurations in the CIA and NRM-dataset and their length is shown. In the scatterplot the individual data points for CIA (blue) and NRM (orange) are shown for the number of lanes of the configuration.**

As already described in the section on the CIA dataset, CIA provides capacity values for weaving sections of five, fifteen and twenty-five percent heavy traffic. In Figure 4.15 all ratios heavy traffic of the NRM configurations are shown in combination with the three different ratios heavy traffic of CIA. This figure shows that the ratios heavy traffic as in NRM mostly lie within the range of the ratios present in CIA. As concluded from Figure 4.7 capacity values between five and twenty-five percent are reasonably to be interpolated. Only for relatively low and high ratios of heavy traffic the relationship between capacity and ratio heavy traffic is not linear. Therefore, only data representing higher ratios of heavy traffic should be added to the dataset, as lower ratios of heavy traffic are not present in the NRM configurations, as shown in Figure 4.15.

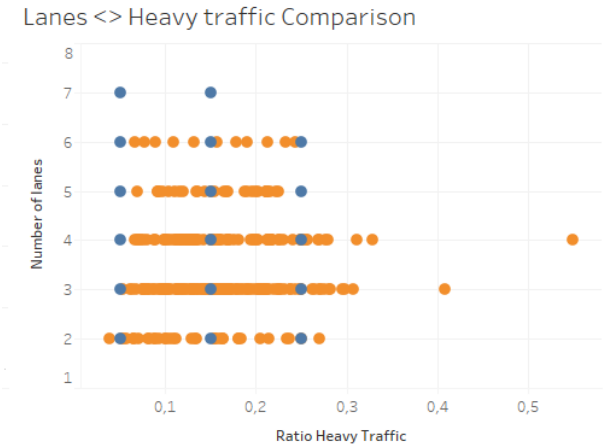



**Figure 4.15: Scatterplot of all ratios heavy traffic in both CIA (blue) and NRM (orange).**

In Figure 4.16 a comparison for the weaving ratios and volume ratios is given. In the upper left scatterplot, the combination of weaving ratios for all configurations is jointly shown. This scatterplot shows that the NRM configuration are represented well by the CIA weaving ratios. However, a relative dense (many NRM-data points) area with a high weave ratio for O2D1/WR2 (0.75>1) combined with a low weave ratio O1D2/WR1 (0>0.2) is not covered well by the CIA data. It is, however, not known if the desired meta model can sufficiently predict capacity values for weaving ratios which transcend configurations. In other words: will the meta model be able to predict capacity values for configuration *x* with certain weaving ratios *y*, which combination is only present for another configuration *z*. This will be identified during the post-training analysis.

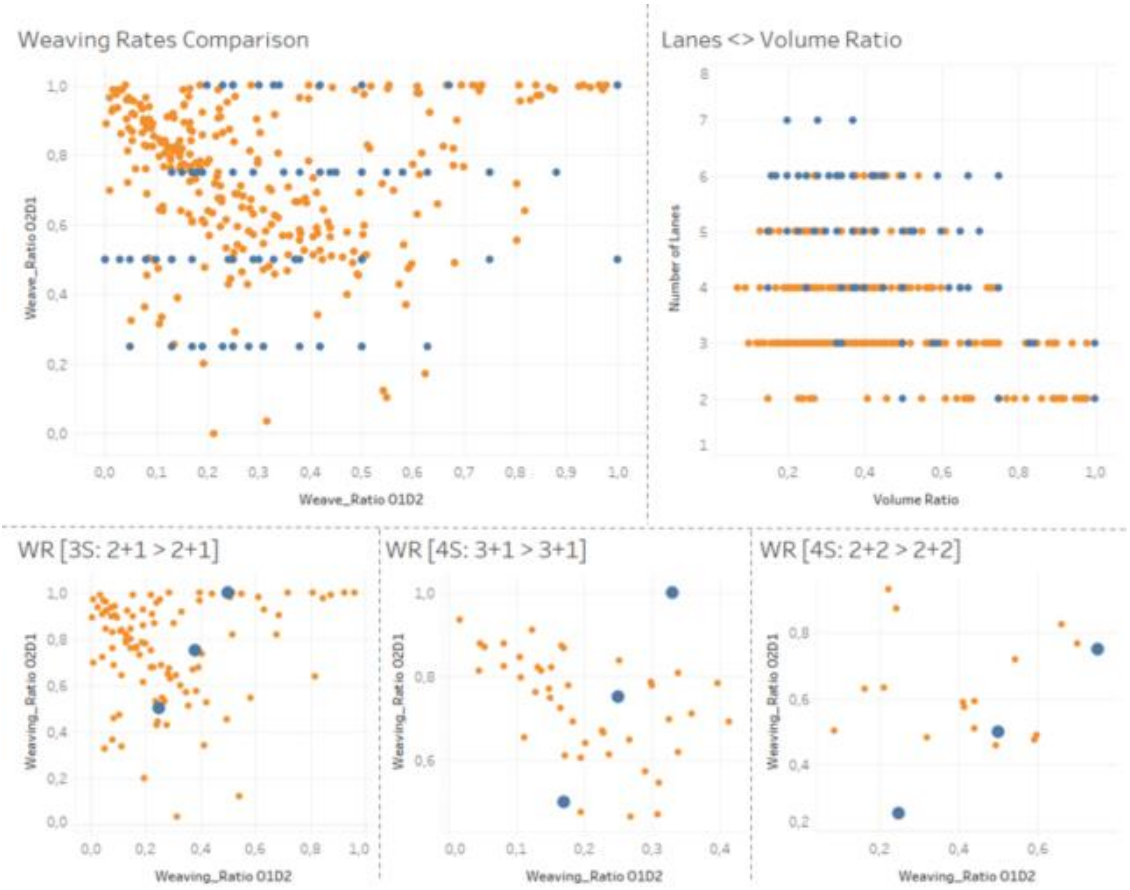



**Figure 4.16: Scatterplots for the weaving ratios and volume ratios in the CIA and NRM dataset.**

For that reason, the combination of weaving ratios for three of the four configurations with the most appearances in the NRM dataset, as shown in Figure 4.16, are shown in the lower three scatter plots. These scatter plots show that the combinations of weaving ratios match poorly for the configurations separately. Moreover, the distance between some NRM configurations and the closest CIA data point is quite large.

Finally, in the scatter plot upper right, the volume ratios for the number of lanes in a configuration are both shown for CIA and NRM data points. Here, the range of the volume ratios in CIA largely cover the range of volume ratios for NRM configurations. However, for configurations with two or three lanes, a large share of data points do have a lower volume ratio than the volume ratios in CIA. Furthermore, as shown Figure 4.18**,** a volume ratio can correspond with several combinations of weaving ratios. For these reasons it is recommended to generate supplementary data where the weaving ratios and volume ratios better match with the configurations in NRM.

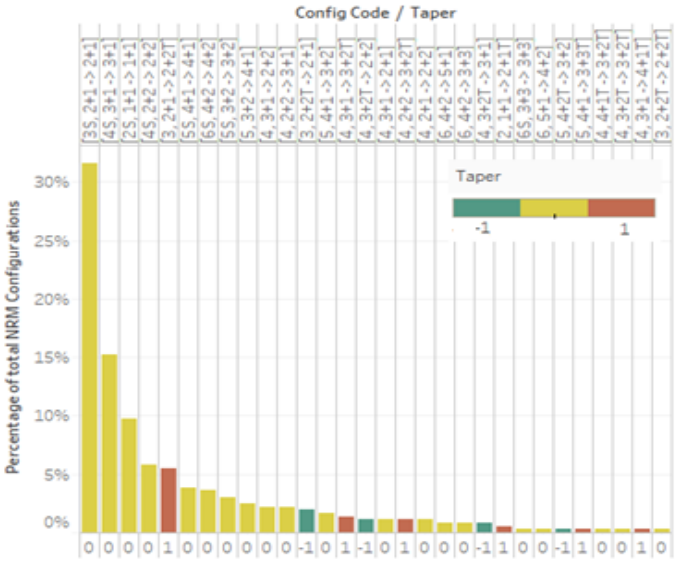



**Figure 4.17: Overview of the share configurations present in the NRM dataset. In the top of the figure the configuration/lane codes for the configurations are shown. Furthermore, the Taper value of the configuration made visible, where: -1 represents a taper upstream, 0 no taper and 1 a taper downstream.**


**Figure 4.18: Scatterplots providing an overview of the weaving ratios, volume ratios and Assymbel-factors.**

As stated above a value for a volume ratio can correspond with several combinations of weaving ratios and the *Assymbel*-factor. In Figure 4.18 scatterplots related to the volume ratios compared to the weave ratios and *Assymbel* factors are shown. This figure shows that the majority of NRM data points have an *Assymbel* factor lower than one. This implies that the traffic flows originating from origin two are over-estimated by CIA, where all *Assymbel* factors are equal to one, since the ratio of flows is equal to the proportion of lanes. Therefore, a large part of the volume ratios is also overestimated by CIA, as weaving ratios for origin two are mostly higher for the majority of data points (which can be seen in the upper scatterplots).

These scatterplots show that the assumption in CIA that the ratio of traffic flows is equal to the proportion of lanes, does not hold for the NRM case. Therefore, NRM configurations with similar weaving ratios in CIA do have a distinctive volume ratio and thus capacity values will not be corresponding. Furthermore, it is illustrated that the volume ratio depends on the combination of values for weaving ratio one, weaving ratio two and the *Assymbel* factor. Since research found that the volume ratio is one of the most capacity-influencing variable, supplementary data points are recommended to simulate with different volume ratios including deviating asymmetrical flow distributions.

In summary, supplementary datapoints are simulated to provide a better match between the CIA and real-case configurations, especially for the following cases:

- More three- and four lane configurations (Figure 4.14);
- Relatively short and long configuration lengths (Figure 4.14);
- Higher (> 25%) ratios of heavy traffic and a denser distribution of ratios of heavy traffic (Figure 4.15);
- Configurations with other ratios of weaving traffic resulting in a denser grid of weave ratios (Figure 4.16Figure 4.15 );
- Configurations with *Assymbel* factors not equal to one hundred (Figure 4.18).

By simulating these configurations more similarities between the training dataset (Section 4.3.1) and the real-cases will be realised (Section 4.3.2). In the next paragraph, the simulation process will be described.

## 4.4 Supplementary Data Simulation

In the theoretical framework (Chapter 3) FOSIM was already introduced, along with the arguments to use FOSIM to generate capacity values. In this chapter the settings (and legitimacy) of gathering data with FOSIM is described using the manual of FOSIM and several validation reports. Initially, the available dataset of CIA would have been supplemented by means of the microscopic traffic simulation software of FOSIM. However, no apparent, unambiguous, settings for the FOSIM simulations are known whereby simulation outputs do not consequently coincide with the capacity values in CIA. For this reason, the complete dataset of CIA is re-simulated, resulting in an identical data grid (configurations and values for independent variables) compared to CIA. However, for this re-simulated set the capacity values are changed. Therefore, the most representative settings for the FOSIM simulations are indicated and researched in this section. With these settings the dataset of CIA is re-simulated. Moreover, these settings have been used to extend the data grid of CIA based on analysis on the configurations in NRM. Hereby, the NRM configurations have not been added simply but gaps between the CIA data grid and NRM set have been identified and filled with supplementary data (see the enumeration above). Furthermore, capacity values for the NRM dataset have been simulated identically to be used for post-validation of the neural network.

### 4.4.1 FOSIM Validation and Implemented Settings

Since CIA does not provide unambiguous setting of the FOSIM simulations for the FOSIM configurations in their manual, several validation reports are used to find optimal settings for reproduction of the *CIA* dataset. The three validation reports which are analyzed were the general validation report (executed by *Sweco*), a report on symmetric weaving sections (CIA 1, *Vermijs, 1997)* and a report on asymmetric weaving sections (CIA-2, *Dijker & Minderhoud, 2001*).

In the validation report it was stated that FOSIM is calibrated for the Dutch road situations in the past, especially focused on weaving sections. However, FOSIM has changed on several points in the last years. For the sake of the importance of FOSIM outcomes and several recent changes, it was desired to validate FOSIM once again (Henkens et al., 2017). The validation report concluded that there is no apparent reason to doubt the simulated capacity values for weaving sections within FOSIM, given the input configurations. However, it is stated that it is hard to represent the reality in FOSIM, mainly due to a lack of information (i.e. lane width and slope) (Henkens et al., 2017). This lack of information in combination with the restrictions of FOSIM, are the reason that some variables have not been taken into account, as stated in Section 4.2.

As stated in the introduction and in the reports on FOSIM, several degrees of freedom in the settings are present which do influence the final outcomes of the simulations runs. These parameters which can be adapted in FOSIM can be summed up as:

- Lane change areas;
- Origin-Destination patterns;
- Ascent of traffic intensities;
- Local features on single lanes.

In contradiction with the statements in *CIA-1* and *CIA-2*, changes in the settings of these parameters do influence the simulation output. In this case, the simulation output are the capacity values. This statement of deviating capacity values is endorsed by the tables in this paragraph describing the results of the sample tests with different settings for the simulations. In the next two paragraphs the settings for the lane change areas and local features on single lanes will be discussed. For the other two parameters the settings are elaborated in Appendix A1, because no (or only small) discrepancies between the several validation reports exist for these parameters.

#### *4.4.1.1 Lane Change Areas*

In *CIA-1*, the lane change areas for the most inner lanes are roughly composed as required lane change areas for 95% of the total length of the weaving length. The remaining part of the weaving length is reserved as a desirable lane change area. For the outer lanes the required lane change area continuous 300 meter downstream of the gore. Furthermore, 600 meter of desirable lane change area is added, which corresponds with the location of signage (Vermijs, 1997). In Figure 4.19, an example of the lane change areas according to *CIA-1* is shown.

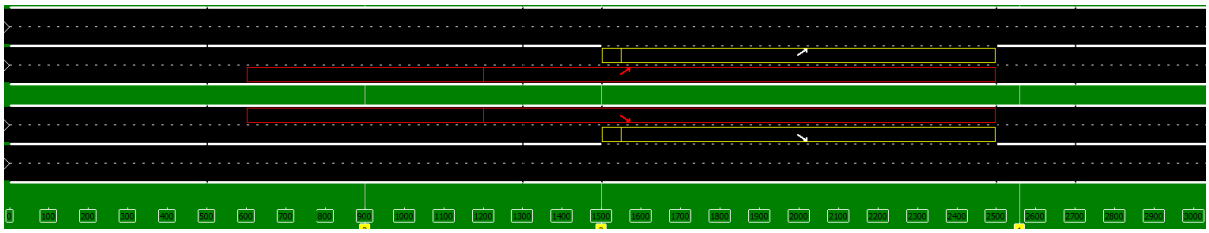


**Figure 4.19: Lane change areas for both destinations for a symmetric 2+2 configuration according to *CIA-1*.**

In *CIA-2* the lane change areas are also, roughly, underpinned on the location of signage. However, in this case the lane change areas are quite different compared to the lane change areas according to the validation report. Moreover, the settings for the lane change areas do differ per configuration and configuration length. This will result in time-consuming batch simulation runs of capacity values for asymmetric weaving sections. In other words, depending on the configuration and configuration length, the lane change areas should be changed, which would result in inconsistent lane change areas and time-consuming simulation set-ups.

The re-simulated capacity values, where the settings of the validation report are used, for the symmetric weaving sections sufficiently corresponded with the capacity values from CIA. Moreover, the contradictions between the several (validation) reports related to FOSIM were minor for the symmetric weaving configurations. Unfortunately, for the asymmetric weaving sections the contradictions between the several reports were

| Configuration Code | Capacity CIA | FOSIM: Val. Report | FOSIM:CIA-2 | FOSIM: CIA-1 |
|---|---|---|---|---|
| [2122T, 600, 50, 50, 15, 120] | 4010 | 3756 | 3864 | 3852 |
| [2231, 850, 25, 75, 15, 120] | 7450 | 7356 | 7422 | 7500 |
| [4132, 1000, 44, 75, 25, 120] | 6730 | 6834 | 7164 | 7050 |
| [4132, 1100, 44, 75, 25, 120] | 6920 | 6870 | 7218 | 7086 |
| [4132, 1200, 44, 75, 25, 120] | 7400 | 7444 | 7626 | 7470 |
| [22T21, 650, 42, 77, 15, 120] | 4730 | 4752 | 4500 | 4854 |
| [22T21, 850, 42, 77, 15, 120] | 4880 | 4896 | 4938 | 4992 |
| [22T21, 850, 17, 50, 15, 120] | 5410 | 5052 | 5070 | 5136 |
| [2112, 700, 63, 25, 15, 120] | 5320 | 5430 | 5454 | 5376 |
| [2112, 900, 63, 25, 05, 120] | 6070 | 6282 | 6288 | 6336 |
| [2112, 900, 88, 75, 25, 120] | 4080 | 4476 | 4500 | 4500 |
| Mean Squared Error: | | 39215 | 69462 | 46873 |
| MSE excl. Taper-config: | | 33993 | 81557 | 55378 |
| MSE Taper-config: | | 48355 | 48295 | 31990 |

**%HT and speed and the capacity according CIA are shown. In the other three columns the capacity values from FOSIM simulations with different settings for lane change areas are shown.**

larger. *CIA-2* roughly uses the location of signage around weaving sections according to the signage schemes of *Rijkswaterstaat* (Rijkswaterstaat, 2017). On contrary, as stated before, the validation report has uniform lane change area-settings for all weaving configurations. Since the discrepancies between the lane change areas for asymmetric weaving sections cannot be neglected, a sampling test for several configurations with varying variables are simulated with different settings of lane change areas.

In Table 4.5, the results of the sample test are shown. In the grey columns the configuration codes with the corresponding capacity values from *CIA* are shown. The configuration code consists of the lanes in the weaving sections, the length of the weaving section, the weave ratio for origin 1 and 2, the ratio of heavy traffic and the speed limit. In the green-coloured columns the capacity values from the simulations are shown. In the first column, the capacity values simulated with the lane-change area-setting of the validation report are shown. The second column contains capacity values with the setting as in *CIA-2* and the last column contains capacity values with a simplified form of settings of *CIA*-1. The settings from *CIA-1* are modified, for reasons of consistency, towards a required lane change area for the complete weaving section length with an additive desirable lane-change area of 300 meters.

In this table it is shown that the settings of the validation report result in the most corresponding capacity values with CIA. Moreover, the mean squared error for these capacity values compared with the capacity values from *CIA* are the lowest.

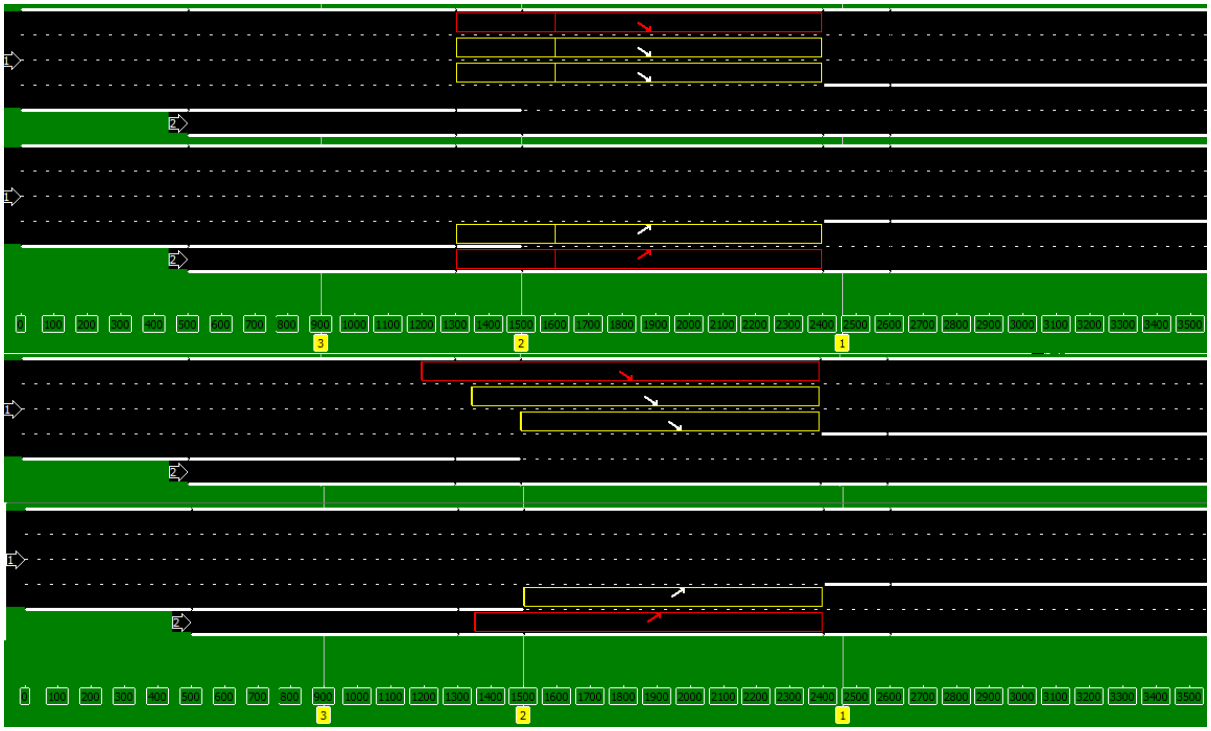


**Figure 4.20: Overview of the lane change areas for the asymmetric configuration [4+1 -> 3+2]. Above lane change areas according to CIA-2 with a require lane-change area for the complete weave length except the first 100 meters and 400-meter additive desired lane changes. Under lane change areas, which are also applied, according to the validation report with required lane change areas for the complete weave length and every additive lane 150 meters supplementary length.**

Next to the fact that the capacity values do correspond relatively good with the CIA values, the outcomes of the simulations make sense. In other words, an increasing weaving length results in increasing capacity values and increasing weaving ratios lead to a decrease of capacity. For this reason, next to the importance of consistency, in this research, the lane change area-settings of the

validation report are used in this case for data generation. As shown in Figure 4.20 these settings comprise required lane change areas for the complete configuration length plus supplementary 150-meter required lane change areas for every additive lane.

#### *4.4.1.2 Local Features on Single Lanes*

Next to the lane change areas, local features on individual lanes do influence simulation outcomes. These local features comprise a speed reduction on the single weaving lanes (i.e. 2+1, 3+1 and 4+1 configurations) and speed reductions on complete 1+1 weaving sections.

The *CIA-1* and *CIA-2* reports also implement a speed reduction for single-lane origins and/or destinations. In the validation report no speed reduction factors are applied. Due to this discrepancy, a sample test is executed for the symmetric 3+1 configuration. Table 4.6 reveals that the settings without a speed reduction on the single lanes correspond better with the CIA capacity values than the configurations with a speed reduction on the singe lane. Therefore, no speed reduction factors will be used during the simulations for data gathering and extension.

The other uncertainty regarding local features on single lanes holds especially for symmetric 1+1 configurations. These configurations do represent the clover-leave interchanges. Hereby, the configurations are relatively short, whereby traffic speeds are lower than for other (continuous) configurations. For configurations with a high-speed reduction factor (80), capacity values are (too) high and many vehicles cannot reach their destination. Since this is not realistic and to validate the settings of the validation report, a sample test is executed with several level of speed reduction factors. The simulation outcomes for all 1+1 configurations in *CIA* are shown in Table 4.7. This table shows that a speed reduction of 60% correspond the best with the capacity values in CIA. For this reason, speed reduction factors of 60% will be used for symmetric 1+1 configurations.

As stated in the introduction of this chapter, an overview of all the used settings within the FOSIM simulations is shown in Appendix A1. Here, also an overview of the other two output-changing parameters is elaborated.

| Configuration Code | Capacity CIA | Capacity Fosim | %-dev | Capacity Fosim SPRD | %-dev |
|---|---|---|---|---|---|
| [3131,600,17,50,5,120] | 8830 | 8742 | -1 | 8454 | -4.3 |
| [3131,700,17,50,5,120] | 9170 | 8892 | -3 | 8496 | -7.4 |
| [3131,800,17,50,5,120] | 8960 | 8946 | -0.2 | 8748 | -2.4 |
| [3131,600,17,50,15,120] | 7520 | 7374 | -1.9 | 7254 | -3.5 |
| [3131,700,17,50,15,120] | 7570 | 7422 | -2 | 7356 | -2.8 |
| [3131,800,17,50,15,120] | 7600 | 7512 | -1.2 | 7458 | -1.9 |
| [3131,600,17,50,25,120] | 6430 | 6516 | 1.3 | 6366 | -1 |
| [3131,700,17,50,25,120] | 6480 | 6636 | 2.4 | 6432 | -0.7 |
| [3131,800,17,50,25,120] | 6530 | 6636 | 1.6 | 6492 | -0.6 |
| [3131,600,25,75,5,120] | 8400 | 8052 | -4.1 | 7872 | -6.3 |
| [3131,700,25,75,5,120] | 8530 | 8298 | -2.7 | 8016 | -6 |
| [3131,800,25,75,5,120] | 8660 | 8454 | -2.4 | 8298 | -4.2 |
| [3131,600,25,75,15,120] | 6970 | 6726 | -3.5 | 6654 | -4.5 |
| [3131,700,25,75,15,120] | 6840 | 6900 | 0.9 | 6792 | -0.7 |
| [3131,800,25,75,15,120] | 7120 | 7116 | -0.1 | 6876 | -3.4 |
| [3131,600,25,75,25,120] | 6020 | 5976 | -0.7 | 5928 | -1.5 |
| [3131,700,25,75,25,120] | 6120 | 6048 | -1.2 | 5886 | -3.8 |
| [3131,800,25,75,25,120] | 6120 | 6192 | 1.2 | 6012 | -1.8 |
| [3131,600,33,100,5,120] | 7700 | 7536 | -2.1 | 4914 | -36.2 |
| [3131,700,33,100,5,120] | 7990 | 7746 | -3.1 | 5004 | -37.4 |
| [3131,800,33,100,5,120] | 8140 | 8028 | -1.4 | 4968 | -39 |
| [3131,600,33,100,15,120] | 6530 | 6396 | -2.1 | 4224 | -35.3 |
| [3131,700,33,100,15,120] | 6440 | 6594 | 2.4 | 4278 | -33.6 |
| [3131,800,33,100,15,120] | 6770 | 6684 | -1.3 | 4368 | -35.5 |
| [3131,600,33,100,25,120] | 5630 | 5556 | -1.3 | 3870 | -31.3 |
| [3131,700,33,100,25,120] | 5780 | 5730 | -0.9 | 3912 | -32.3 |
| [3131,800,33,100,25,120] | 5880 | 5874 | -0.1 | 3972 | -32.4 |

**Table 4.6: Table with results of sample test on 3+1 configurations for speed reduction factors. The column 'Capacity Fosim' represents capacity values without speed reduction on the single lanes and the column 'Capacity Fosim SPRD' represent capacity values with speed reduction factor 0.8 on the single lane.**

| Configuration Code | Capacity CIA | Capacity Fosim 80 | Capacity Fosim 70 | Capacity Fosim 65 | Capacity Fosim 60 |
|---|---|---|---|---|---|
| [1111,350,50,50,5,120] | 2680 | 3474 | 3210 | 3060 | 3000 |
| [1111,350,50,50,15,120] | 2320 | 3054 | 2796 | 2664 | 2568 |
| [1111,350,50,50,25,120] | 2150 | 2760 | 2568 | 2484 | 2382 |
| [1111,350,75,75,5,120] | 2360 | 3234 | 3234 | 2802 | 2634 |
| [1111,350,75,75,15,120] | 2110 | 2808 | 2538 | 2448 | 2238 |
| [1111,350,75,75,25,120] | 1940 | 2598 | 2370 | 2292 | 2196 |
| [1111,150,100,100,5,120] | 2100 | 2994 | 2250 | 2160 | 2136 |
| [1111,150,100,100,15,120 | 1850 | 2568 | 1974 | 1920 | 1896 |
| [1111,150,100,100,25,120 | 1680 | 2418 | 2058 | 1788 | 1740 |
| [1111,200,100,100,5,120] | 2140 | 2970 | 2346 | 2208 | 2172 |
| [1111,200,100,100,15,120 | 1860 | 2508 | 2040 | 1956 | 1908 |
| [1111,200,100,100,25,120 | 1720 | 2400 | 2130 | 1860 | 1794 |
| [1111,350,100,100,5,120] | 2170 | 3156 | 2754 | 2376 | 2232 |
| [1111,350,100,100,15,120 | 1930 | 2688 | 2358 | 2094 | 2004 |
| [1111,350,100,100,25,120 | 1780 | 2484 | 2220 | 2148 | 1926 |
| Max. Speed after speed reduction >> | | 96 | 84 | 78 | 72 |
| Vehicles ~ Destination >> | | 1.64 | 0.43 | 0.24 | 0.16 |

**Table 4.7: Sample test for speed reduction factors on 1+1 weaving configuration. In the coloured columns, the capacity for configurations with several speed reduction factors.**

### 4.4.2 Conclusions on FOSIM Simulations

After an introduction on FOSIM in Chapter 3.4, the validation and used settings are described in this chapter. First of all, the complete datagrid of CIA is re-simulated because no consistent corresponding capacity values were able to be simulated. Hereby, a bias between, on the one hand, the CIA and supplementary datapoints and, on the other hand, the (CIA) training dataset and the simulated validation set (NRM) is prevented. To find consistent settings for (re-)simulation, three available reports on FOSIM are discussed. Two of it were more focussed on the used settings for generating capacity values for respectively symmetric and asymmetric weaving sections, the other was more focussed on the validation of the simulation software for the Dutch road and traffic conditions/situations. Since discrepancies between the used settings did exist, several sample tests were executed to match sufficiently with the available dataset in *CIA*. Regarding the discrepancies between the settings in the validation reports, the degrees of freedom for the FOSIM settings are: the lane change areas, origin-destination patterns, ascent of traffic intensities and local features on single lanes. The final settings for (re-)simulating capacity values are summarized in Appendix A1. In Figure 4.21 a summary of the findings in the sections on FOSIM is shown.

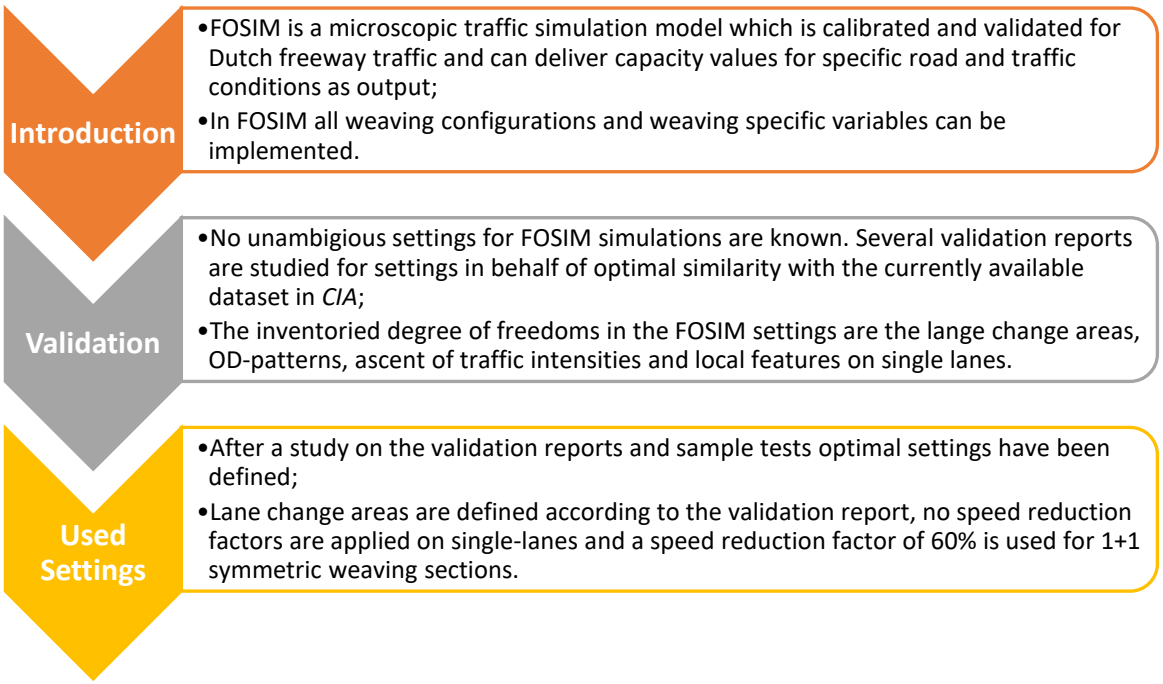



**Figure 4.21: Summary of the chapters (Chapter 3.4 and Chapter 4.4) on FOSIM.**

## 4.5 Final used Training dataset

For the ranges that were underrepresented within the CIA dataset (Sections 4.3.2/4.3.3) supplementary data points have been simulated to enhance the match between the simulated (CIA) dataset and the real-case configuration (NRM) dataset. On page 58 the regions of the supplementary data simulations have been listed. These supplementary data have been simulated using FOSIM (Section 4.4).

Firstly, data points with a relatively lower and higher length have been simulated, shown in Figure 4.22. Hereby, only 2% of the NRM data points lie outside the range of simulated lengths and regarding the number of lanes per configuration separately only 11% of the NRM configurations lie outside the range of simulated lengths.

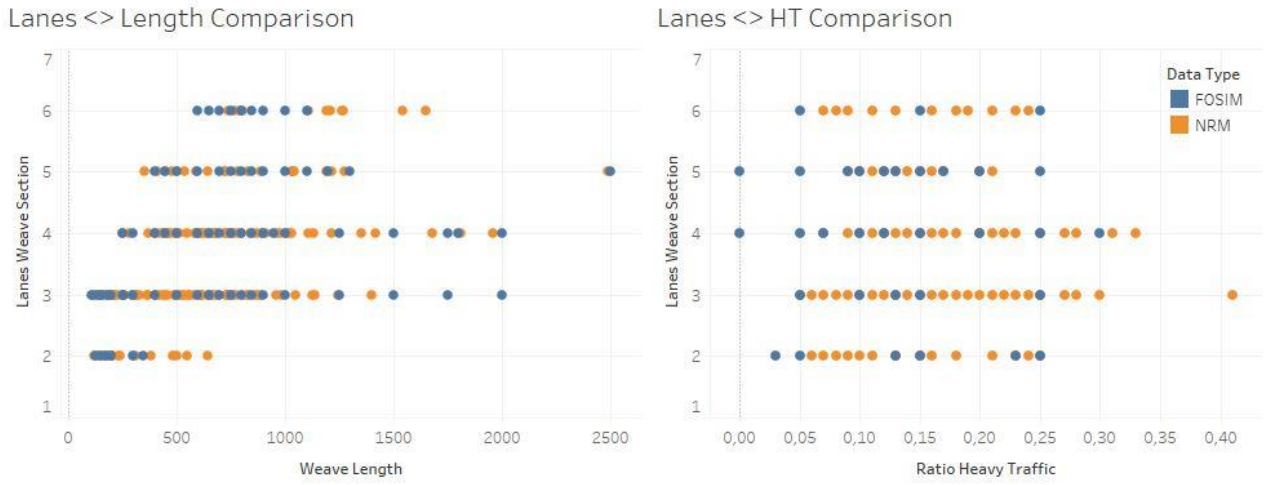



**Figure 4.22: Scatterplots for comparing the simulated dataset and the NRM dataset. In the left scatterplot the length of the weaving section are plotted against the number of lanes of the configuration. In the right scatterplot the ratio heavy traffic is plotted against the lanes of the weaving section.**

Secondly, more configurations with a deviating ratio of heavy traffic are simulated. The CIA dataset only provided capacity values with ratios of heavy traffic of five, fifteen and twenty-five percent. The supplemented dataset contains additional capacity values with ratios of heavy traffic of zero to thirty-five, which can be seen in the left scatterplot of Figure 4.22. Hereby, the meta model should theoretically predict capacity values with relatively low and high ratios of heavy traffic better.

Thirdly, more data points are simulated with deviating weaving ratios. Since a complete grid of weaving ratios is simulated for configuration 3+2, as shown in Figure 4.9, the grid of weaving ratios for all configurations combined as shown in Figure 4.23 is much more dense then the previous grid of weaving ratios in Figure 4.16. In the upper-right scatterplot this is also shown by the increased range of volume ratios. Currently, only 4% of the NRM data points lie outside the range of volume ratios in the simulated dataset.

Despite all the supplementary simulated data points, it is too time-consuming to simulate a complete data grid regarding weave ratios for every configuration. For example, differences still exist for the configurations as shown in the lower three scatterplots. Nonetheless, at the time of writing the question remains if the meta model will be able to sufficiently predict capacity values for configurations with deviating weaving ratios where the combination is only known for other

configurations. To illustrate: will the meta model be able to predict capacity values for configuration 2+1, with weaving ratios [0.9, 1](a combination only present for configuration 3+2)?

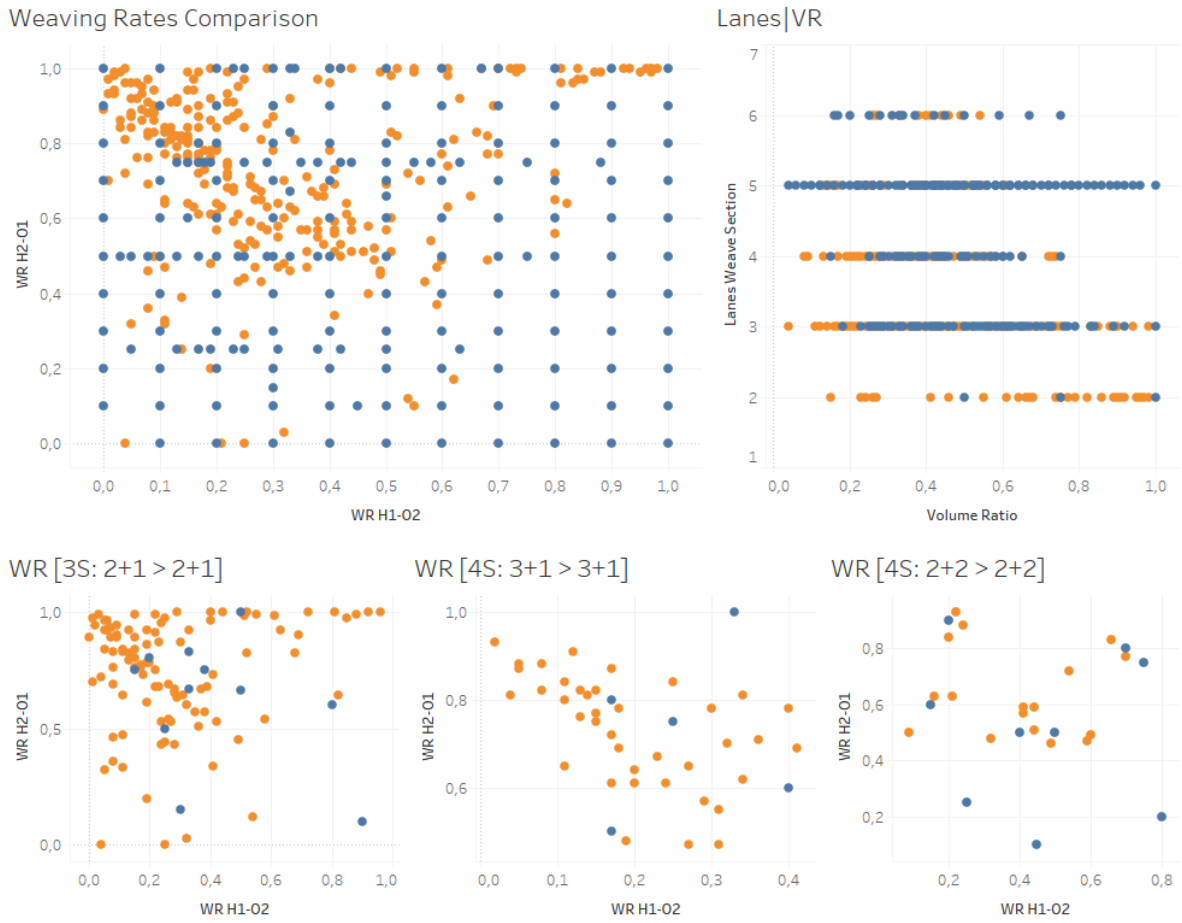



**Figure 4.23: Scatterplots for the weaving ratios and volume in the simulated (blue) and NRM (orange) dataset for all configuration combined (upper scatterplots) and for the three most frequent configurations in the NRM dataset (lower scatterplots).**

Fourthly and finally, Figure 4.24 compares volume ratios, weave ratios and *Assymbel* factors present in the simulated CIA and NRM datasets. The histograms show a better match between the simulated and NRM dataset regarding the number of lanes per configuration. Furthermore, the upper two histograms show that the share of (relatively) low volume ratios is high for the NRM configurations. Only for the 2-lanes configuration (1+1), the majority of configurations have a high volume ratio. Furthermore, the scatterplots from Figure 4.24 (simulated dataset, left; NRM configurations, right), show many similarities. The simulated dataset has several combinations of values for WR1 and *Assymbel* for identical values for the volume ratio. This is shown using different sizes and colors of the data points in the scatterplot of Figure 4.24. Due to inclusion of these combinations, the meta model should theoretically perform better when predicting capacity values for different ratios of weaving traffic, volume ratios and the *Assymbel* factor.

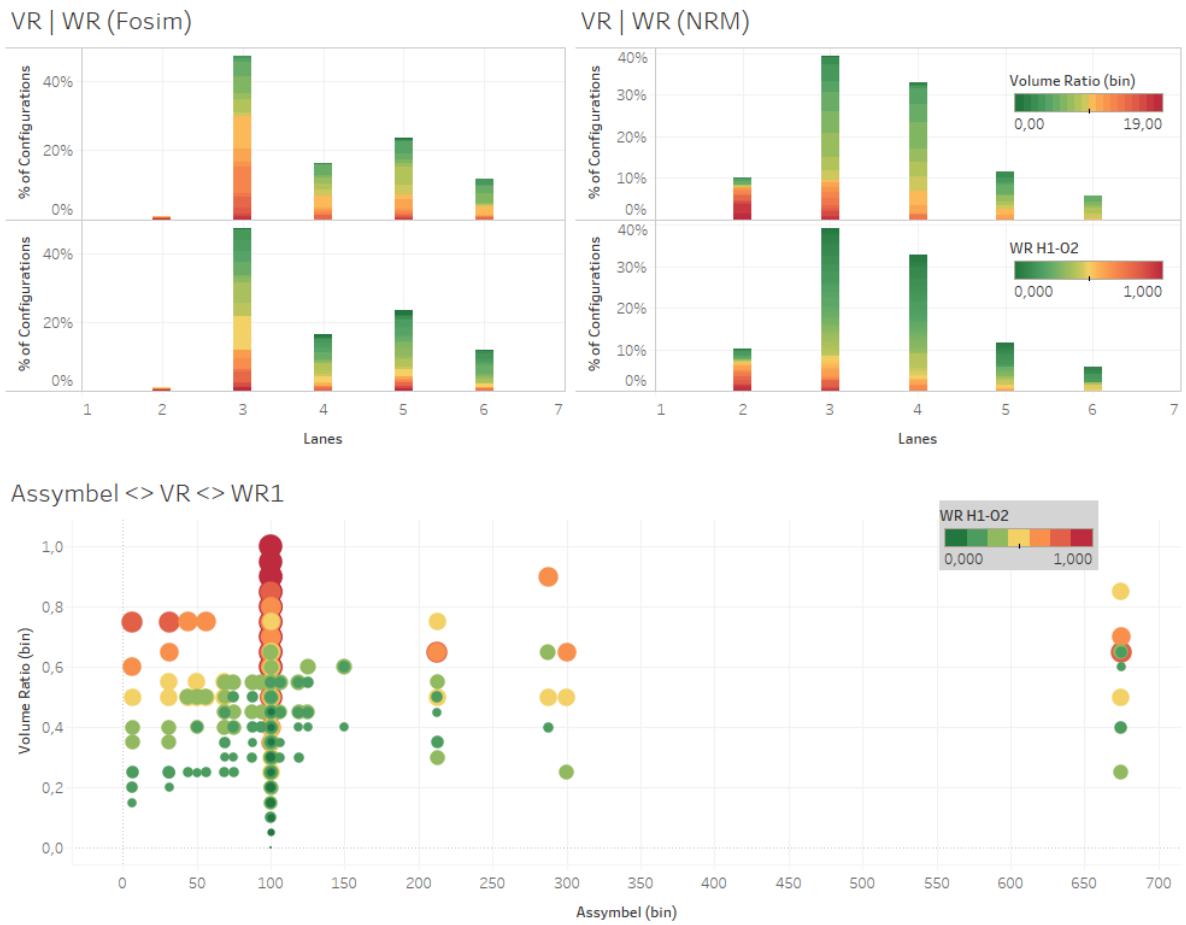



**Figure 4.24: Overview of the relations between the Volume Ratio, Weave Ratio and Assymbel factor in both the simulated FOSIM and NRM dataset.**

## 4.6 Conclusions Pre-Training Phase

In the pre-training phase the significant variables influencing the capacity value were discussed. Using the set of significant variables and the 'two-sided' restrictions on using them, a final set of independent variables is determined. This set was shown in Figure 4.13 and contains the weaving configuration, weaving length, weaving traffic, traffic composition and division of traffic flows. The independent variables are both included in the training dataset (CIA; re-simulated with FOSIM) and the validation dataset (NRM). Since, the capacity values (dependent variables) in the CIA dataset were not able to reproduce, the complete dataset is re-simulated. Moreover, to provide better similarities between the datagrid of CIA and real-case configurations, supplementary datapoints are simulated with FOSIM.

Despite the fact that for some combinations of values for independent variables are not or underrepresented, supplementary data simulation resulted in improved similarities between the (final) CIA dataset and real-case configurations, which are represented by the NRM-West dataset. In the next phase of the model development process, the training phase, the neural network will be trained by means of the supplemented CIA dataset and validated with the real-case configurations present in the NRM-West dataset.

# 5 Training Phase

Now the training and external validation dataset are known, the network can be trained. The network estimates capacity values (dependent variables) by means of a set of capacity influencing variables (independent variables). In Figure 5.1 the components of the training phase, as part of the iterative model development process, are shown. In this chapter the network selection, training algorithm and network training will be discussed.

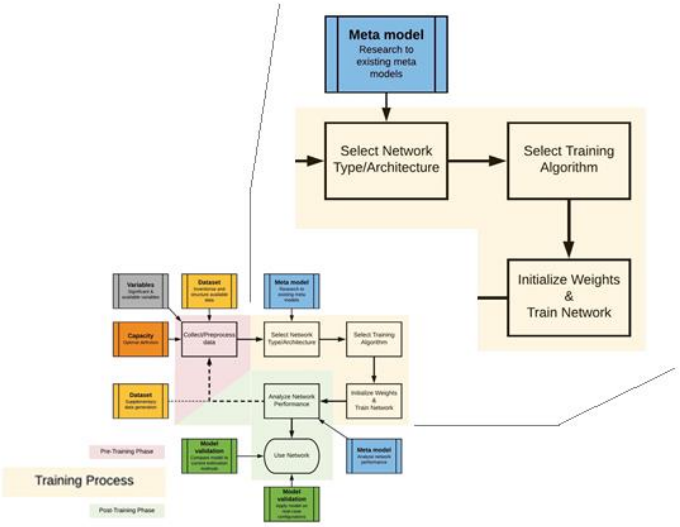



**Figure 5.1: (Iterative) Model development process as already shown in Figure 2.3, with the components included in the training phase enlarged.**

Firstly, the input and output layer will be described, which components are discussed in the pre-training phase. Then, the network architecture and the degrees of freedom within the network architecture will be elaborated upon. Since several networks have been trained, multiple network architectures have been developed. Finally, the training process and its stochasticity, due to the weights and bias initialization, will be discussed.

The neural network output is analysed afterwards in the post-training analysis. Moreover, (selected) trained network is used to be validated using real-case configurations represented by the NRM-West road network. Finally, the neural network will be compared with the currently used capacity estimation methods.

## 5.1 Input and Output layer

In Section 3.6 an introduction to Neural Networks is given. Here, it was elaborated that a neural network consists of three components, namely: the input layer, the hidden layer and the output layer. Moreover, it was stated in section 3.6 that structure of the input and output layer mainly depends on the external problem specification (Hagan et al., 2014). In this case, the external problem specification provides a set of independent and dependent variables that can be implemented in several manners.

The output layer of the network consits of only one node, namely the capacity value as the dependent variable. The network estimates the capacity value by means of a set of independent variables (see Figure 5.2), namly: [$Lanes_{O1}$, $Lanes_{O2}$, $Lanes_{D1}$, $Lanes_{D1}$, *TotalLanes, %HeavyTraffic (HT), Configuration Length, Weave Ratio 1, Weave Ratio 2, Volume Ratio, Speed Reduction Factor*]. It should be noted that the *Speed Reduction Factor* in the input layer is not the speed reduction factor as discussed in Section 4.1.6. Here, the speed reduction factor is a variable implemented in FOSIM for 1+1 configurations and short weaving configurations with a lower speed. Firstly this factor was not included in the input layer. However, implementing this factor resulted in a improved training result.

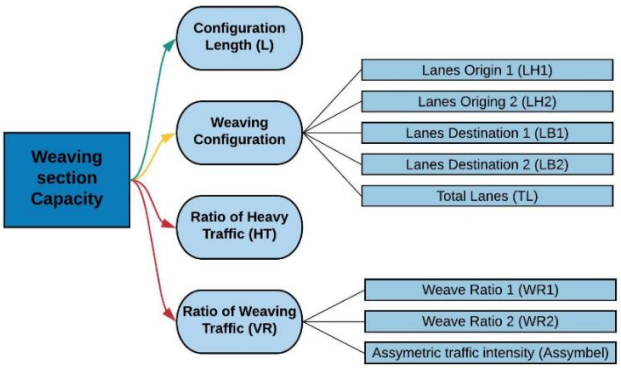



**Figure 5.2: Overview of the independent variables which influences capacity.**

Furthermore, a correlation matrix, shown in Table 5.1, is constructed to identify mutual correlated variables. Table 5.1 shows that the variables concerning the weaving configuration (*$Lanes_{O1}$, $Lanes_{O2}$, $Lanes_{D1}$, $Lanes_{D1}$, TotalLanes*) are strongly mutual correlated. Therefore, other input layers have been implemented where the weaving configuration is expressed in different manners.

First of all, a network was trained with an input layer exluding the variables *Lanes Destination 1* and *Lanes Destination 2*. Secondly, an input layers with the variables *Lanes Origin 1, Lanes Origin 2, Total Lanes, Weaving Type (HCM)* and *Taper Presence (-1,0,1)* replacing the original five variables of the weaving configuration is implemented. However, these alternative input layers did not result in an improved training result of the neural network. This can be explained by the fact that a neural network can estimate capacity values better with the complete weaving configuration known.

| | LanesO1 | LanesO2 | LanesD1 | LanesD2 | TotalL | %HT | Lenght | WR1 | WR2 | VR | SpRd | Asym | Cap |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lanes Origin 1 | x | - | - | - | - | - | - | - | - | - | - | - | - |
| Lanes Origin 2 | 0.03 | x | - | - | - | - | - | - | - | - | - | - | - |
| Lanes Destination 1 | 0.92 | 0.18 | x | - | - | - | - | - | - | - | - | - | - |
| Lanes Destination 2 | 0.3 | 0.39 | 0.17 | x | - | - | - | - | - | - | - | - | - |
| Total Lanes | 0.88 | 0.39 | 0.88 | 0.56 | x | - | - | - | - | - | - | - | - |
| % Heavy Traffic | -0.04 | 0 | -0.04 | 0 | -0.02 | x | - | - | - | - | - | - | - |
| Configuration Length | 0.42 | 0.37 | 0.42 | 0.3 | 0.51 | 0.05 | x | - | - | - | - | - | - |
| Weave Ratio 1 | -0.15 | -0.13 | -0.08 | -0.36 | -0.22 | -0.02 | -0.3 | x | - | - | - | - | - |
| Weave Ratio 2 | -0.32 | -0.03 | -0.39 | -0.09 | -0.28 | 0.03 | -0.03 | 0.08 | x | - | - | - | - |
| Volume Ratio | -0.44 | -0.01 | -0.44 | -0.12 | -0.4 | 0.01 | -0.2 | 0.56 | 0.81 | x | - | - | - |
| Speed Reduction Factor | 0.41 | -0.35 | 0.4 | 0.28 | 0.45 | 0.07 | 0.66 | -0.26 | -0.06 | -0.21 | x | - | - |
| Assymetric Flow Division | 0.05 | 0.06 | 0.06 | 0.08 | 0.08 | 0 | 0.03 | -0.06 | -0.06 | 0.05 | -0.02 | x | - |
| **Capacity** | **0.75** | **0.26** | **0.77** | **0.38** | **0.8** | **-0.25** | **0.51** | **-0.32** | **-0.47** | **-0.59** | **0.5** | **0.08** | **x** |

**Table 5.1: Correlation matrix including the independent and dependent variables.**

Therefore, the (final) input layer consists of twelve variables, namely: [*$Lanes_{O1}$, $Lanes_{O2}$, $Lanes_{D1}$, $Lanes_{D1}$, TotalLanes, %HeavyTraffic (HT), Configuration Length, Weave Ratio 1, Weave Ratio 2, Volume Ratio, Speed Reduction Factor*]. This set of indipendent variables is used by the neural network to estimate capacity values as the dependent variable.

## 5.2 Network Architecture

Now the input layer and output layer are known, the hidden layer can be discussed. Moreover, the final settings within the neural network can be elaborated upon.

The final neural network consists of one hidden layer with 10 neurons. The hidden layer is connected to the input layer with twelve neurons and an output layer which provides the estimated capacity value. The data is normalized using the *mapminmax* function, which normalizes the data so that all inputs fall in the range [-1, 1] (Beale et al., 2018). The transfer functions in the hidden layer are *sigmoid* transfer functions and in the output layer linear transfer functions are applied, which are stated to be the most suitable transfer functions for fitting problems (Hagan et al., 2014). The final network architecture is shown in Figure 5.3.

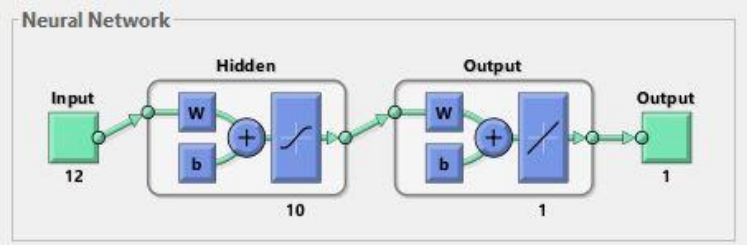



**Figure 5.3: Layout of the final neural network, including the input, hidden and output layer.**

Since the neural network development is an iterative process, several networks with alternative data normalization methods and transfer functions (as inventoried in Section 3.6) have been trained. Nevertheless, post-training analysis (see Chapter 6) found that the above described settings resulted in the best-performing settings. Moreover, number of neurons in the hidden layer have been varied to find the most optimal number of neurons. Too little neurons in the hidden layer resulted in a malperformance of the network, where too many neurons in the hidden layer resulted in an over-trained network (which is illustrated in next section).

## 5.3 Network Training

After constructing the neural network architecture, the network is trained. The training of the neural network is done using the supplemented CIA dataset as described in Section 4.5. This dataset is randomly split up in a training-, testing- and validation dataset. The training dataset is used to fit the network to the data. The validation dataset is used to estimate the prediction error for the model selection. The test dataset is used for assessment of the generalization error of the final chosen model (Hagan et al., 2014). Seventy percent of the total supplemented CIA dataset is used as trainingset and fiveteen percent is used as testing- and validation set.

In Section 3.6 the general training process and mechanism is already described. The final network is trained using the *Levenberg-Marquardt* training rule. Several alternative training rules have been tried, however, the *LM*-training rule provided the best network performance.

The network training process begins with the normalization and the division of the dataset. Then the initial weights and biases will be set on small random values. Since every training run provides other output results due to stochasticity in initial weights and biases and the division of training, testing and validation set, the network architecture is trained 100 times. This *'multistart'* training enables statistical comparison between different settings of the neural network and finds the best settings for the specific neural network related to initial weights and training, testing and validations sets. In the chapter on the post-training analysis, Chapter 6, the selection criteria for the optimal settings of the neural network is elaborated.

The training of the network is stopped if the mean-squared error decreased to zero, the gradient decreased to an (infinite) small number or when *early stopping* is applied. Since the error never decrease to zero during the training process, the magnitude of the gradient and the number of validation checks are used to terminate the training. When the training reaches the minimum of the performance, the magnitude of the gradient will become very small, which terminates training (Beale et al., 2018). However, in most cases the training is terminated by *early stopping*. Early stopping can be set by de parameter *Validation Checks* (or max_fail), which represents the number of successive iterations that the network performance on the validation set fails to decrease. The number of *validation checks* is set on six, increasing the number of validation checks pave the way for network overtraining.

The early stopping mechanism is illustrated in Figure 5.4. The left plot shows that the network performance on the training set improves every iteration. However, after iteration 88, the performance on the validation set fails to decrease. The network training process is then stopped after iteration 94 (≡ 88 + 6 *Validation Checks*). Furthermore, this figure shows that the network is over trained when 50 neurons are used in the neural network, as discussed in previous section. Here, the performance on the training set is much better than for the validation- and test set. Here, the training process is not terminated after the 16$^{th}$ iteration, when training performance still improves where the test- and validation performance only improves very little.

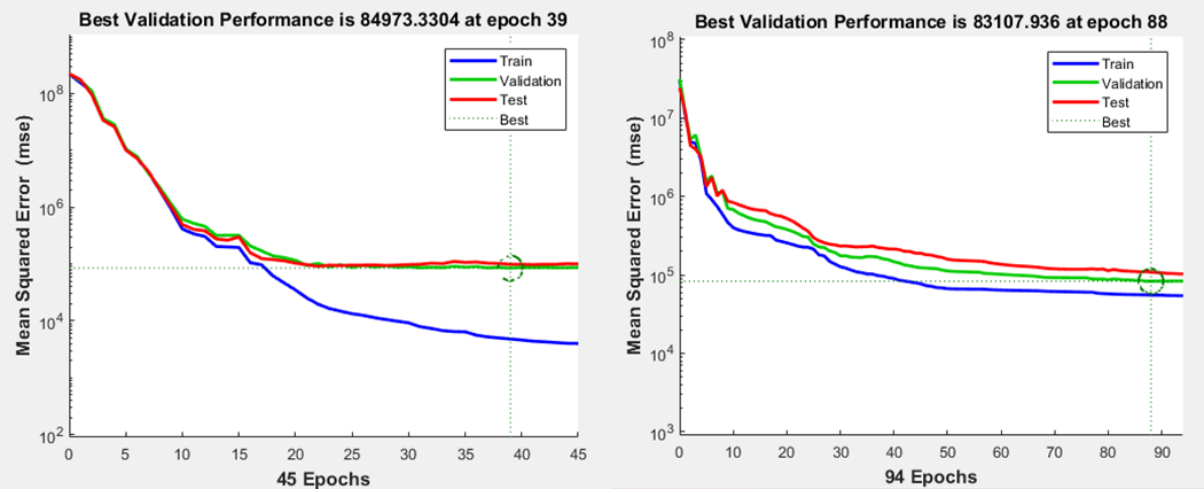



**Figure 5.4: Performance plot for the train-, validation- and test set for a neural network with 50 (left) and 10 (right) neurons in the hidden layer.**

## 5.4 Conclusions Training Phase

In this section the training process of the neural network is described. Several different network architectures and settings have been implemented. The architecture and settings presented in the summary of the training phase (see Figure 5.5) are found to provide the best training results, which will be discussed in the next chapter on the post-training phase.

**Network Architecture**

- Input layer consist of twelve independent variables which are used by the network to estimate one dependent variable, namely: the capacity value.
- (One) hidden layer consisting of 10 neurons and *Sigmoid* transfer functions is used. Output layer uses linear transfer functions.

**Network Initialisation**

- Network trained by using supplemented CIA dataset, which is split into training, testing and validation set (70,15,15%).
- Network is trained by means of a 'multistart' because of stochasticity in network output due to initialisation of weights, biases and division of dataset.

**Network Training**

- Network is trained using the *Levenberg-Marquardt* training rule
- Training process is terminated when MSE decreases to zero, the gradient becomes (infinte) small or by using *early stopping*.

**Figure 5.5: Summary of the network architecture and settings used in the selected network.**

# 6 Post-Training Phase

In this section the results of the trained network will be analysed. The post-training phase of the model development process includes the analysis of the model performance and the network applicability, see Figure 6.1.

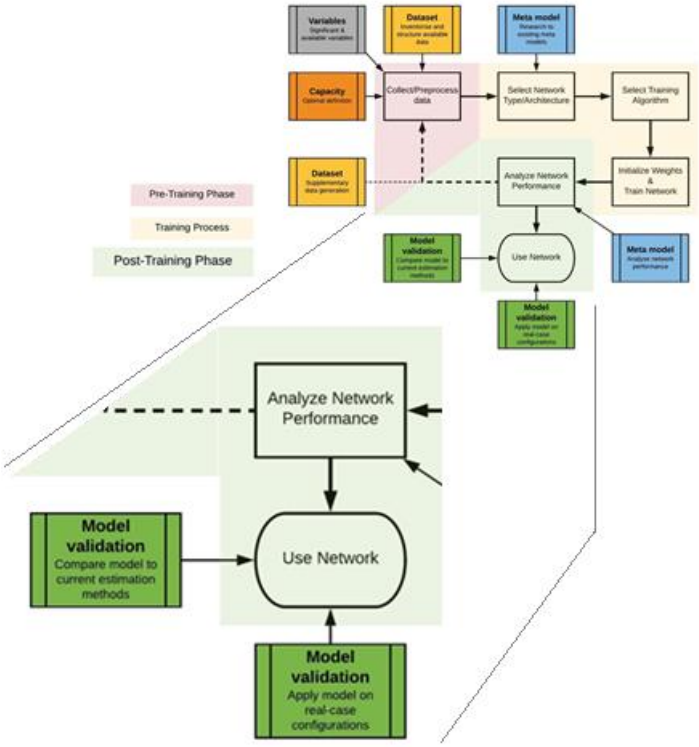



**Figure 6.1: (Iterative) Model development process as already shown in Figure 2.3, with the components included in the post-training phase enlarged.**

For this research the post-training phase includes three different elements. The first analysis discusses the model performance on the training-, test- and validation set. The second component analyses the model performance on real-case configurations. These real case configurations are derived from the NRM-West network, as discussed in Section 4.3.2. Furthermore, the performance of the neural network for the NRM-configurations is compared with the performance of the currently used capacity estimation models: QBLOK and a nearest neighbour-based method. Based on this analysis, a verdict can be given if the neural network is an improvement of the currently used methods. Finally, the network applicability will be elaborated. In this paragraph the errors of the neural network will be compared with the FOSIM simulations and the network performance on configurations which were not consider is elaborated.

## 6.1 Network Training Results

As stated in previous paragraph, 100 *multistarts* have been executed for training a single network. In other words, a certain network architecture has been trained with 100 different initial weights and biases, resulting in different network outputs. Therefore, the most well-performing neural network was needed to be selected. This has been done by means of the following analyses, which are elaborated in the next sections:

- Highest regression coefficient and narrow regression plots (target vs. output capacity values);
- Lowest errors in error histograms;
- Lowest Mean Squared Errors (MSE) and/or lowest Mean Absolute Error (MAE).

Despite applying the early stopping within the training process, the neural network could have been overtrained. Therefore, a model will only be selected if and only if (constraints):

- The relative importance of the independent variables is legitimate;
- The Network Identification Diagrams (NIDs) show sound connections;
- The Neural network is well-performing on a sample tests for pattern recognition (i.e. patterns corresponding with the patterns found in Section 4.1).

First of all, the network performance on the train-, test- and validation set is analysed. These three sets were randomly divided from the supplemented CIA dataset. The network performance on this set is analysed by means of three different visualisations, namely: regression plots, error histograms and 3D error-plots. Moreover, a histogram of the relative importance of variables in the neural network and a *Neural Interpretation Diagram* can indicate model overtraining or generalization.

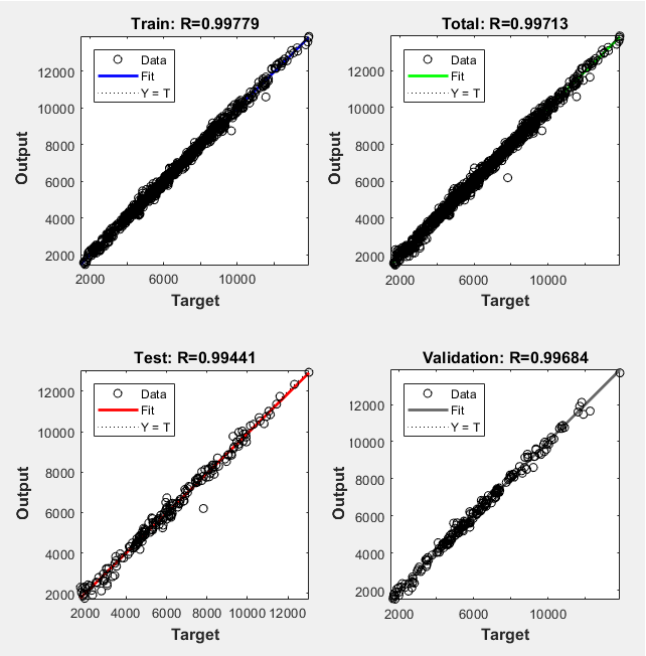



**Figure 6.2: Regression plots, as output of the neural network training, for all the three sets (training, test and validation) separately and a regression plot with all data points jointly.**

The first selection criterium of the neural network is the regression coefficient and regression plots. Here, the output data is plotted against the target data. In this case: the predicted capacity values and the simulated capacity values of FOSIM. The regression plots are made for the training, testing and validation set separately and all data points within the supplemented CIA dataset jointly. These regression plots are shown in Figure 6.2. No large differences between the output results for the different sets are present. This indicates, as discussed in Chapter 3.6, that the network is, most probably, not overfitted or this means that the training, testing and validation dataset are very similar (Hagan et al., 2014).

The R-squared values for all three sets are almost equal to one, which suggests a perfect fit. Only one small outlier can be identified in the regression plot (at target capacity of 8000 veh/h). This configuration has a relatively low value for the *Assymbel* factor, which are not present much in the dataset, due to the FOSIM restrictions on maximum traffic flow per lane (see Chapter 4.4). Next to the regression plots an error histogram, with bins of 0.5 percent, has been created. From this histogram, shown in Figure 6.3, it can be derived that 23% of the errors lie in the range of -1 to 1% and 80% of the errors lie in the range of -5 to 5%, but it also shows that some errors in the range of -25 to 30% are present.

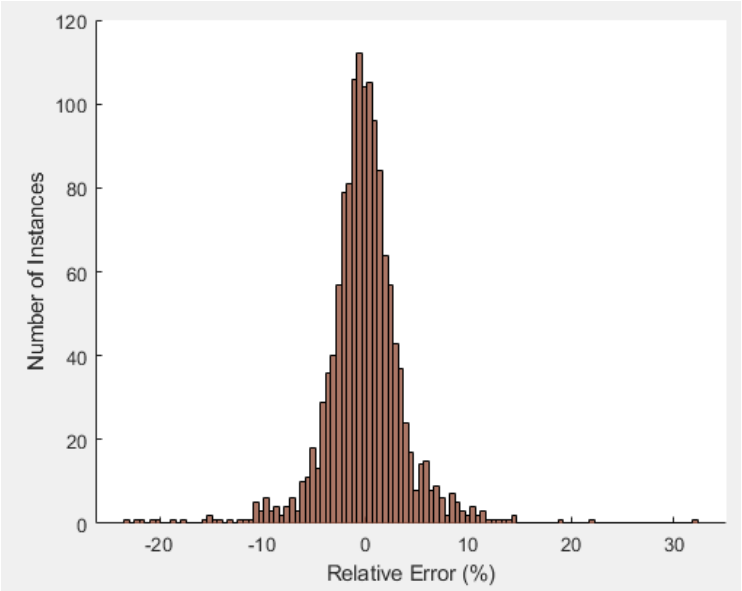



**Figure 6.3: Error histogram, with bin sizes of 0.5%, for all data points.**

Furthermore, potential structural (relatively large) errors or errors for certain values of variables or a combination of values for variables can be identified by means of several 3D-plots. Two examples of these plots are shown in Figure 6.4.

In the left scatterplot of Figure 6.4 the combination of weaving ratio 1 and weaving ratio 2 have been plotted on the x and y axis, where the volume ratio is plotted on the z axis. The green and red data points do represent the configurations with the 10% highest errors. From this scatterplot it can be concluded that the errors related on the weave ratios and volume ratios are (randomly) divided. In other words: the model does not structurally estimate incorrect capacity values for certain combinations of weave ratios and volume ratios. In the right scatterplot the number of lanes per origin are plotted on the x- and y axis and the configuration length on the z axis. This scatterplot shows that several combinations of lanes, the model predicts structurally worse. For example, capacity values for the 1+1 configuration are structurally over- and underestimated.

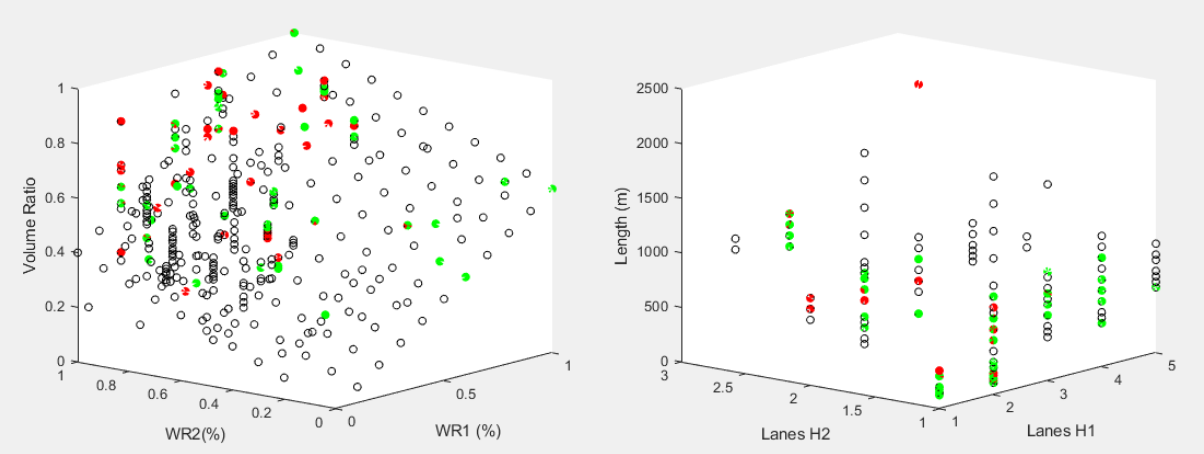



**Figure 6.4: 3D plots with variables on the x, y and z axis. The green data points are the 5% most underestimated configurations and the red data points are the 5% most overestimated configuration.**

In brief, the performance of the neural network is sufficient for the supplemented CIA dataset. However, on first sight, capacity values for some configurations (i.e. 1+1 configurations) are under- or overestimated. For these instances, network validation will be decisive for gathering supplementary data.

Next to the performance criteria shown in regression plots, histograms and 3D error plots, unexpected overtraining of the neural network is researched. Several methods exist to 'illuminate the black box' of the Neural Network. These methods are discussed by *Olden & Jackson, 2002* and *Ibrahim, 2013*. In one of these methods a histogram is created with the relative importance of input variables on the predicted output. This relative importance is calculated using the magnitude of the weights per independent variable (Ibrahim, 2013; Olden & Jackson, 2002).

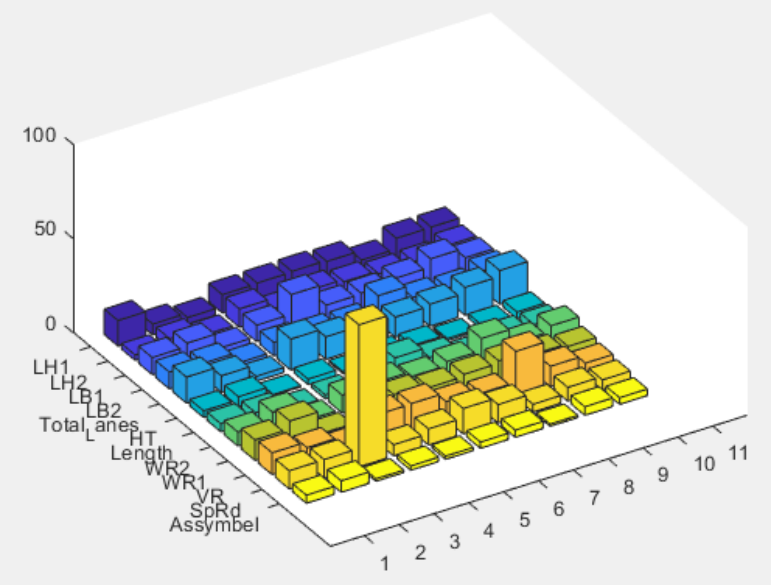



**Figure 6.5: Histogram with the relative importance (z-axis) of the variables (y-axis) for 10 network training runs (x-axis).**

The histograms of ten different training runs have been displayed in Figure 6.5. Since a broad study on the influence of the variables on the capacity value is executed, it can be analysed using these histograms if the results of the neural network are plausible. Despite good results for the third training run, it can be concluded that the neural network did not reach a global optimum, since it is not realist that the speed reduction factor (SpRd) determines the output for almost eighty percent (see third row of histograms).

The histogram with the relative importance of the independent variables for the selected network is shown in Figure 6.6. The relative importance of the variables is seen as legitimate. The number of lanes and the ratio of weaving traffic (VR, WR1 and WR2) are the most determinant variables for estimating capacity, which was also found in literature. Moreover, no peculiarities as in Figure 6.5 are found in this histogram. In other words, this histogram indicates that no overtraining did happen.

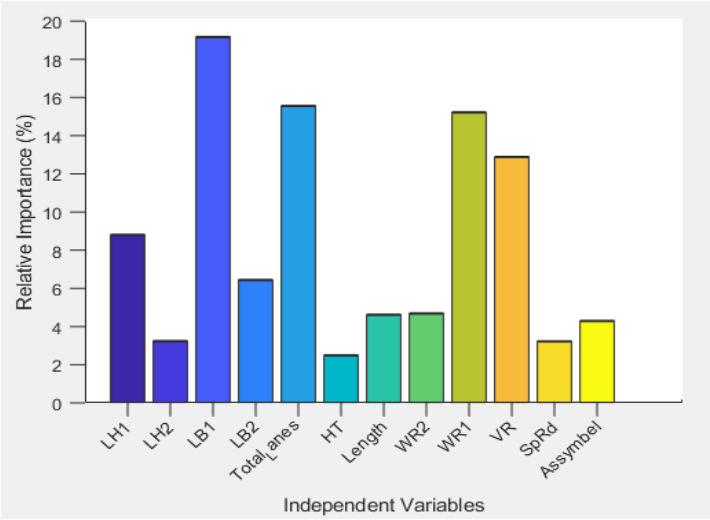



**Figure 6.6: Histogram of the relative importance per independent variable.**

Next to the histograms of relative importance, *Olden & Jackson, 2002* introduces a 'Neural Interpretation Diagram' (NID). Neural Interpretation Diagrams provide a visual interpretation of the connection weights among the neurons and layers in a neural network (Olden & Jackson, 2002). In Figure 6.7 the Neural Interpretation Diagram of the selected neural network is show. Here, the relative magnitude of the connection weights is represented by the line thickness (i.e. thicker lines represent greater weights where thinner lines represent smaller weights). The line shading represents the direction of the weight (i.e. green lines represent positive, excitatory signals and red lines represent negative, inhibitor signals). For visibility reasons, only weights of a certain magnitude (i.e. connections at least 5% of the maximum weight) are shown in the NIDs.

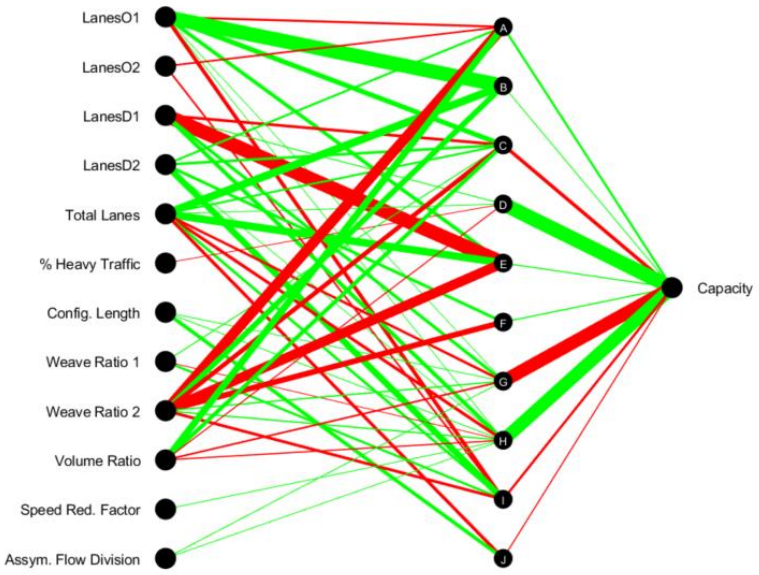



**Figure 6.7: Neural Interpretation Diagram (NID) of the selected neural network.**

For creating an improved view of the effect that each input variables has on the output variable and interactions among predicter variables, a neural interpretation diagram with a multiplication of the two connection weights (input to hidden layer and hidden to output layer) have been created, which is shown in Figure 6.8. However, *Olden & Jacksen, 2002* stated that the interpretation of connections weights, and more specifically NIDs, is not an easy task because of the complexity of connections among the neurons.

Focusing on *Neuron E*, it is apparent that road capacity decreases as *weaving ratio 2* (WR2) increases. However, interactions between several configuration components (*LanesO1*, *LanesD1* and *Total Lanes*) enter the same hidden neuron. This indicates that an interaction between weaving ratio and configuration do influence road capacity, which lies in line with the conclusions in Chapter 4.

Furthermore, in previous chapters (see Section 4.1) it was found that the *Asymmetric Flow Division* could have a positive or negative influence on road capacity depending on the weaving ratios 1 and 2. The NID shows that both an excitatory and inhibitor signal connects the *Asymmetric Flow Division* to *Neuron H and G* respectively. In *Neurons H* and *G* the *Asymmetric Flow Division* interacts with the *Volume Ratio, Weaving Ratio 1* and *Weaving Ratio* 2, supporting the statements in Section 4.1.

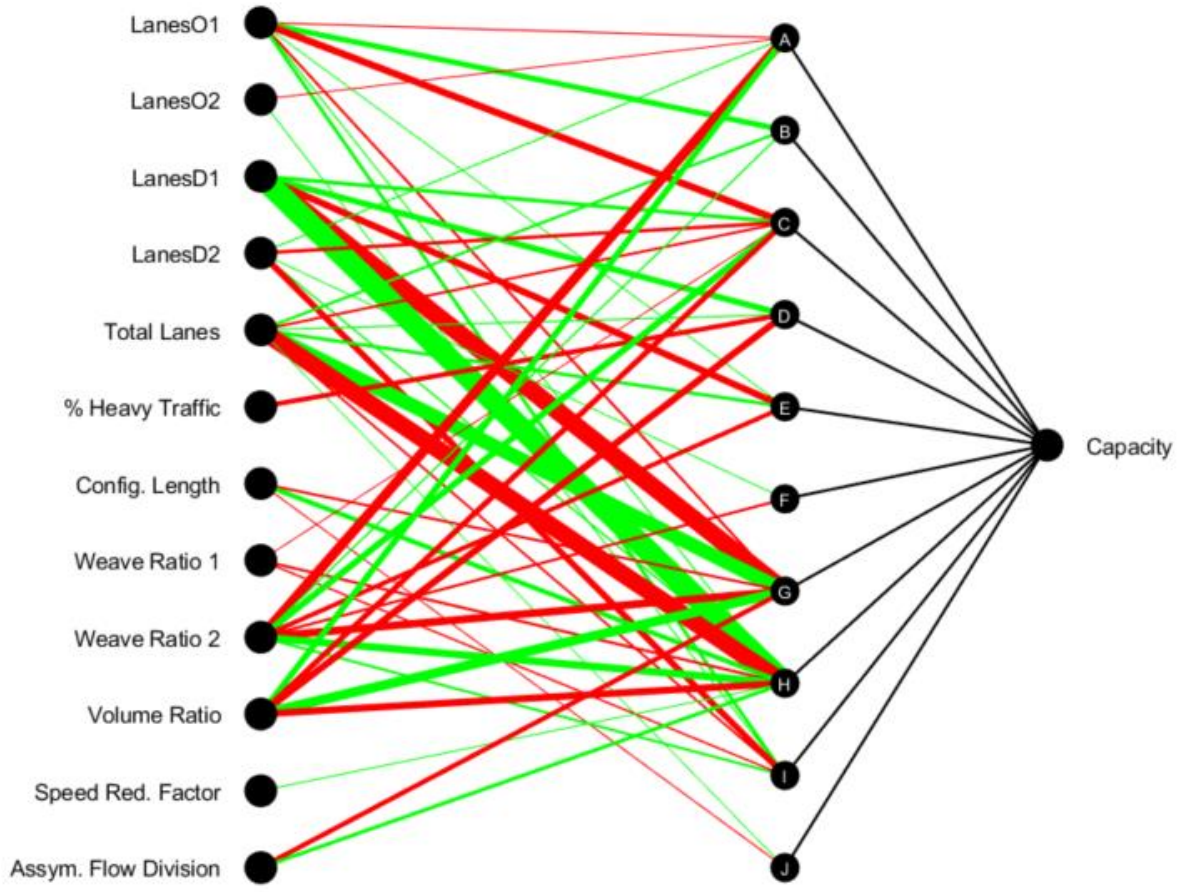



**Figure 6.8: Neural Interpretation Diagram with the weights between the input and hidden layer and output and hidden layer multiplied.**

The final selection criterium was the ability of pattern recognition by the neural network. For this criterium, several configurations have been defined with varying values for configuration length, ratio of heavy traffic and volume ratio. First, the relation between configuration length and capacity is analysed. In the upper graph of Figure 6.9 it is shown that the capacity structurally increases as the configuration length increase, when all other variables remain constant. For very short configurations, the parameter of the speed reduction factor (SpRd) is modified, in accordance with the training dataset. This lies in line with the conclusions in Chapter 4.1 (see Figure 4.6).

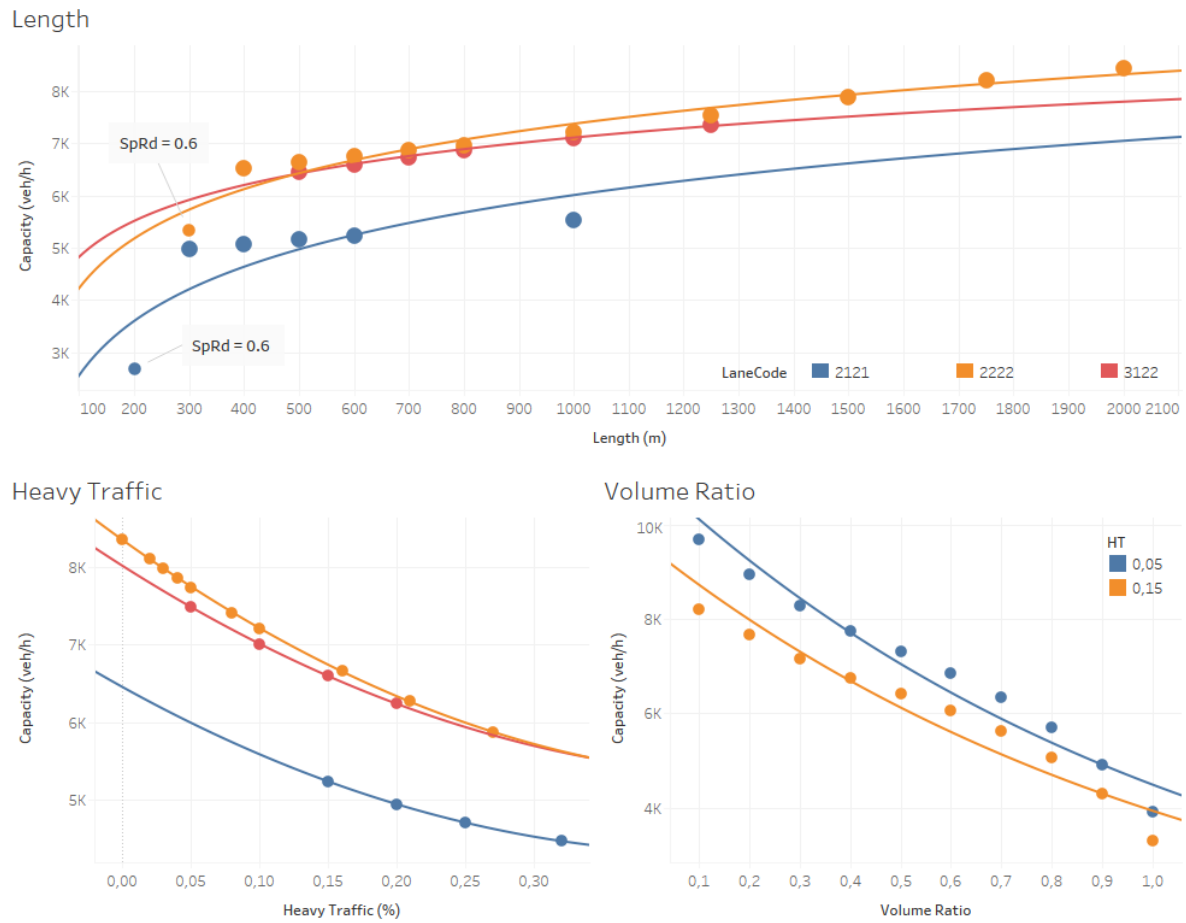



**Figure 6.9: Overview of the ability of pattern recognition of the neural network. The upper graph shows the relation between the estimated capacity and configuration length. The bottom left graph shows the relation between the estimated capacity and ratio of heavy traffic, where the bottom right graph shows the relation between the volume ratio and estimated capacity.**

Secondly, the bottom left graph in Figure 6.9 shows that capacity structurally decreases as the ratio of heavy traffic increases. This is also in accordance with the conclusions found in Chapter 4.1 (see Figure 4.7). Moreover, this pattern is also found for ratios of heavy traffic between 2 and 5% (orange line), these ratios of heavy traffic were not present in the training dataset (see Figure 4.22). This supports that the model is well able to interpolate for this variable.

Finally, the bottom right graph in Figure 6.9 shows the relation between the volume ratio and estimated capacity values. Here, the capacity also decreases as the volume ratio decreases, which lies in line with the conclusions in Chapter 4.1. Based on this figure, it can be stated that the network recognise patterns and only gets the wisdom out of the data which is in it. This concept is called *generalization* and cause that the network will perform as well in new situations as it does on the data on which it was trained (Hagan et al., 2014).

## 6.2 Network Validation

Now, the network performance is analysed for the training dataset and overtraining, with sufficient results, the performance of the neural network on real-case configurations is analysed. The real-case configurations are represented by the weaving sections present in the NRM-West network which are simulated within FOSIM. Hereby, the trained neural network can be used for estimating capacity values for the NRM-configurations.

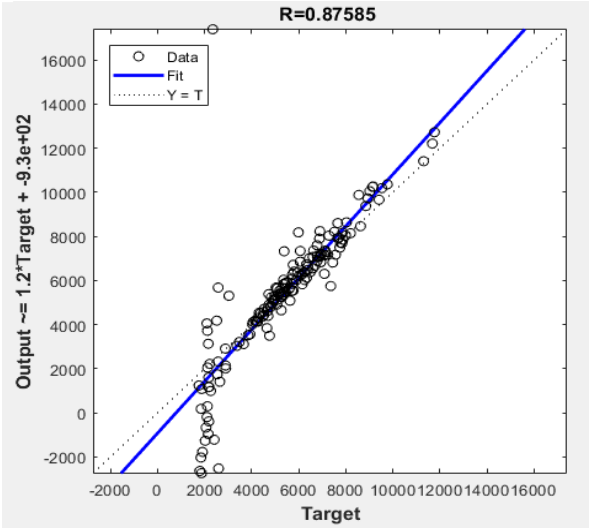



**Figure 6.10: Regression plot with the target values for NRM-configurations and the predicted (output) values of the neural network.**

In first instance, the capacity estimations for the NRM-West network appeared to be sufficient. However, as shown in Figure 6.10, large errors did arise in the target range of [1500, 2500]. These capacity values do match with the capacity values for the 1+1 configurations. Moreover, negative capacity values are predicted for these target values. This indicates extrapolations, since no negative capacity values could theoretically exist. As stated by *Hagan et al., 2014*, supplementary data should be gathered for these region of target values. For the 1+1 configurations, CIA only provided weaving ratios equal to [50,50; 75,75; 100,100] for short weaving sections, while in real-case configurations more variation in weaving ratios and length is present. After gathering supplementary data for this region of target data, the model performance on the real-case configurations improved.

Figure 6.11 shows a regression plot and error histogram for the trained network on the real-case configurations. It is shown that the regression between the target and output is improved compared to the regression plot in in Figure 6.10. This is explained by the fact that supplementary data is simulated for 1+1 configurations (see Section 4.4), which improved network performance. The network performance, expressed in the regression coefficient, is slightly worse compared to the network performance on the train-, test- and validation set (TTV) of the supplemented CIA dataset.

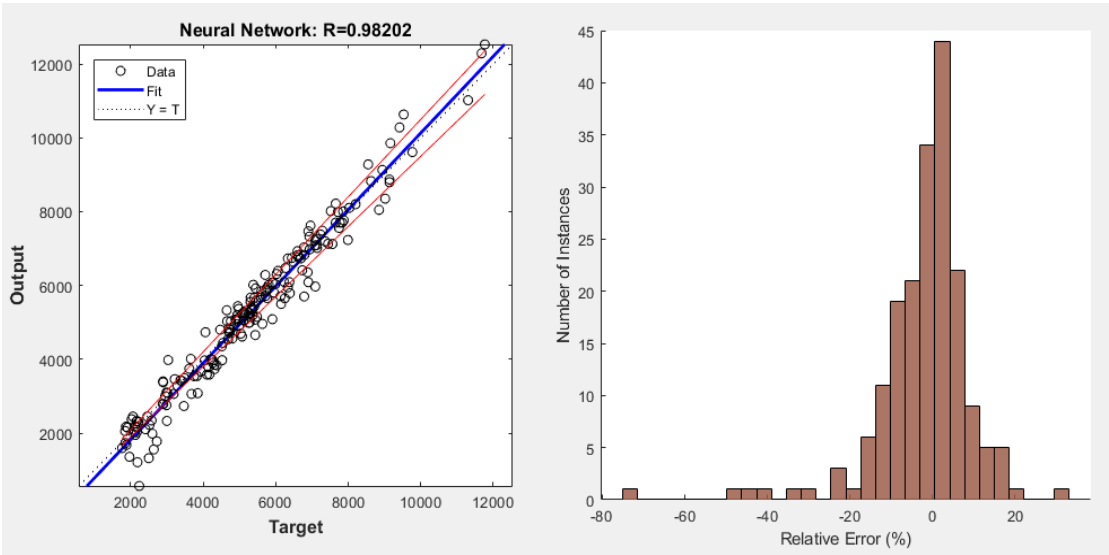



**Figure 6.11: Regression plot and error histogram of the results of the trained network on the real-case NRM configurations. The blue line in these figures are the trendlines for the capacity estimations, where the black dotted lines represent the perfect fit (R=1) with supplementary red lines supply ±5% error compared with the perfect fit.**

The histogram of Figure 6.11 shows the relative errors on real-case configurations are relatively larger than the errors on configurations in the TTV sets, which are shown in Figure 6.3. It has been derived that 28% of the configurations lie between an error range of [-2.5%, 2.5%] and 51% between the range [-5%, 5%]. However, an error of 80% is also seen for the real-case configurations. The larger errors lie, primarily, in the region of 1+1 configurations, the large errors are represented by 1+1 configurations with deviating ratios of weaving traffic and long lengths, which were not directly present in the training dataset.

Solely the regression plot and error histogram do not bring in a verdict if the neural network is an improvement compared to the currently used capacity estimation methods. Therefore, the performance of the neural network compared to the currently used capacity estimation methods is discussed. These two methods are *QBLOK* and a *Nearest Neighbour* method, which are also described in Section 2.3. Capacity values predicted within *QBLOK*, which is the standard network loading model of the Dutch NRM models, are predicted by means of a reduction compared to the base capacity, which is a black-box procedure. The nearest neighbour-based method compares the NRM configuration with the present configurations in the CIA manual and selects the capacity value of the most similar configuration.

These methods also estimated capacity values for the configurations present in the NRM-West network. Since, FOSIM simulation have been executed for these configurations, regression plots and error histograms can be made to show the performance of these capacity estimation methods.

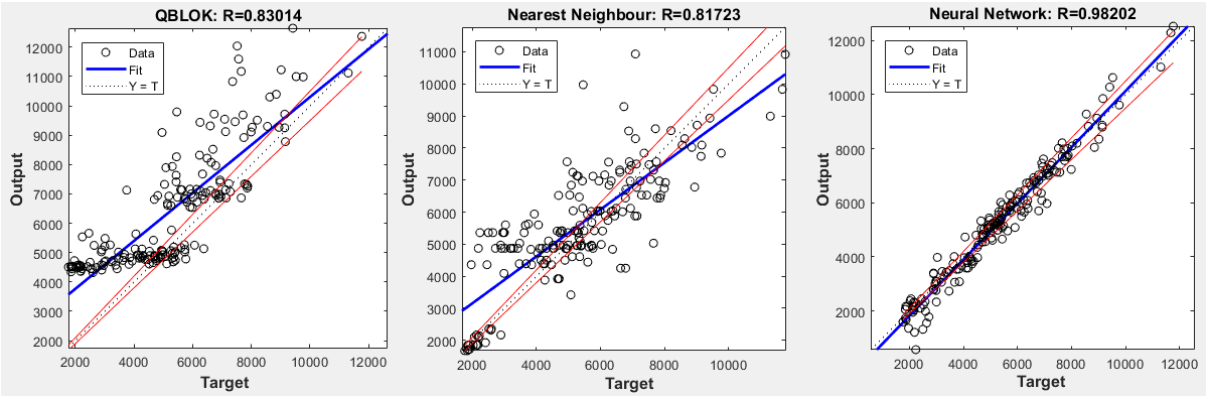



**Figure 6.12: Regression plots for the alternative capacity estimation methods and the neural network (left: QBLOK, middle: Nearest-Neighbour, right: Neural Network).**

Figure 6.12 shows the regression plots for the QBLOCK method (left) and *Nearest Neighbour* method (right). The blue line in these figures are the trendlines for the capacity estimations, where the black dotted lines represent the perfect fit (R=1) with supplementary red lines supply ±5% deviation of the perfect fit. The left scatterplot shows that the capacity values which are estimated by QBLOCK are structurally overestimated. The capacity estimations made by the *Nearest Neighbour* method are more dispersed around the perfect fit. However, the regression coefficient is lower for the *Nearest Neighbour* method than for the QBLOK method.

Next to the regression plots, error histograms have been plotted. Figure 6.13 shows error histograms for both the QBLOK (left) and *Nearest Neighbour* method (right). The left scatterplot confirms that QBLOK overestimates capacity values. Roughly 85% of the capacity values estimated by QBLOK are overestimated compared to the FOSIM capacity values, which is stated to be the 'ground-truth'. Moreover, for roughly 15% of the configurations, the relative error is more than 100%. The maximum error of the neural network is approximately 80%, which is seen only one time in the set of real-case configurations.

Structurally overestimated capacity values, which is the case for QBLOK estimated values, results in less congestion on the road network. Hereby, bottlenecks which are present in real-case are not indicated by the macroscopic traffic model. This effect has consequences for the complete road network, because (amongst others) weaving sections function as node in the road network. The malperformance of these methods were the reason for this research of developing a data-analytic model for estimating capacity values.

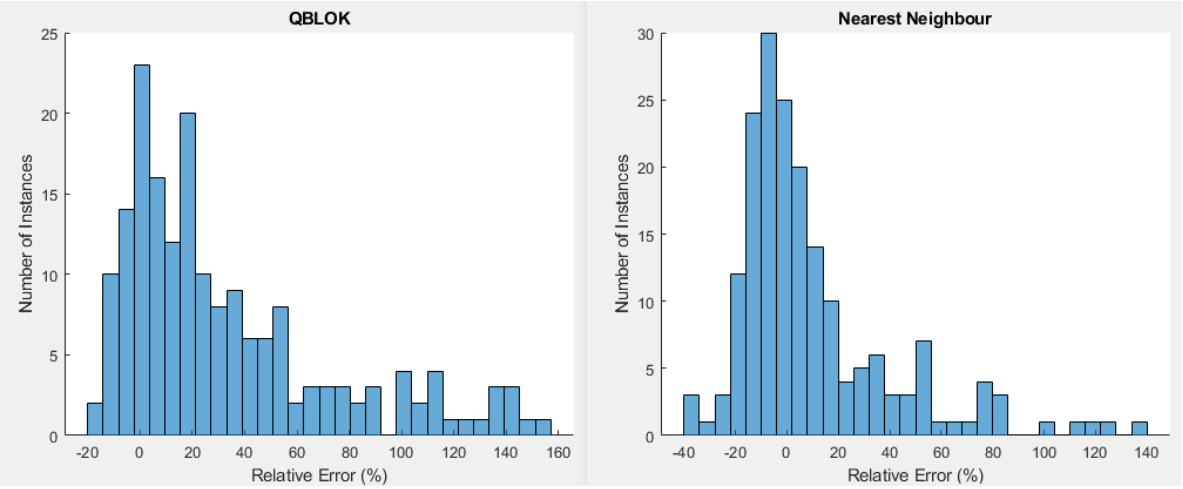



**Figure 6.13: Error histograms for the alternative capacity estimation methods (left: QBLOK, right: Nearest-Neighbour)**

The right scatterplot shows that the errors of the *Nearest Neighbour* method are more dispersed around zero. However, the share of overestimated configurations is also larger than the share of underestimated capacity values. For the *Nearest Neighbour* method 21% of the estimated capacity values lie in the error range of [-5%, 5%] and 43% in the range of [-10%, 10%]. For the QBLOK method this is worse, namely: only 16% of the estimated capacity values lie in the range of [-5%, 5%] and 31% in the range of [-10%, 10%]. Whereas, 51% of the estimated capacity values by the *Neural Network* lie in the error range of [-5, 5%].

In brief, when comparing the performance of the different capacity estimation methods, it can be concluded that the capacity estimation method of the neural network outscores the currently used capacity estimation methods.

## 6.3 Network Applicability

Next to the model validation on real-case configurations, this paragraph goes in on the model performance compared with FOSIM and elaborates on the configurations which are not considered within the neural network.

First of all, the magnitude of the errors of the neural network estimations are compared with the standard deviations of the FOSIM simulation per configuration. In other words, assuming that the median value of 100 FOSIM simulations is indeed the capacity value, the error per configuration of the neural network is compared with the standard deviation of the 100 FOSIM simulations.

In Figure 6.14 the error of the neural network is visually compared to the standard deviation of FOSIM per configuration. In the upper left and figure below, the model error is subtracted of the standard deviation of the FOSIM simulations. Negative (green) values do represent configurations where the neural network error is lower than the standard deviation of FOSIM. Positive (red) values do represent configurations where the neural network error is larger than the standard deviation of FOSIM.

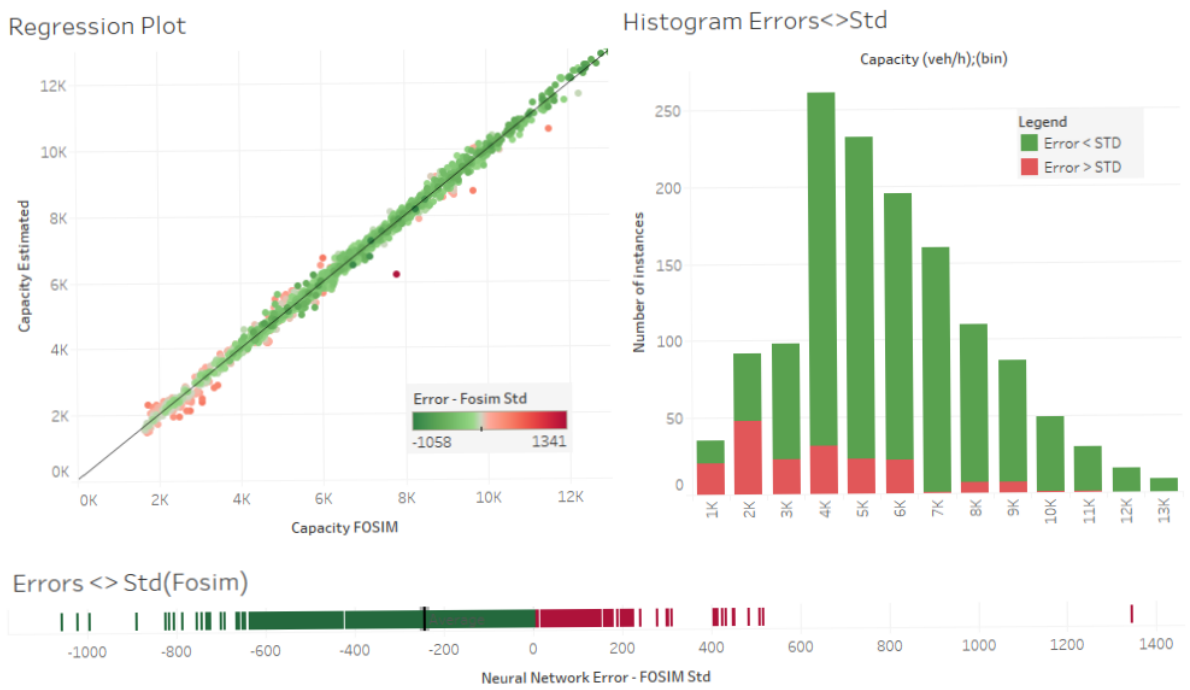



**Figure 6.14: Analysis of the neural network errors with the standard deviation of FOSIM. Upper left: regression plot of the FOSIM capacity (x-axis) and estimated capacity (y-axis). Upper right: histogram with the share of configurations where the neural network error is larger or smaller than the standard deviation of FOSIM. Below: spread**

The upper two plots of Figure 6.14 show that primarily for configurations with low capacity values, the error of the neural network exceeds the standard deviation of the FOSIM simulation. This can be explained by the fact that for these configurations, the standard deviation is also lower than for configurations with high capacity values. To illustrate: the mean standard deviation for 1+1 configurations is equal to 114, for 2+2 configurations 348, for 3+2 configurations 446 and for 5+1 configurations 588 veh/h. This can be explained by the fact that more traffic is simulated on more extensive configurations, which results in more stochasticity in driving behaviour and thus capacity values. Therefore, only relatively small errors are permitted for the neural network to lie within the range of the standard deviation of FOSIM.

The figure below shows the spread of the errors compared with the FOSIM standard deviation. It is shown that the mean of the errors compared to the FOSIM standard deviation (Neural Network Error minus FOSIM standard deviation) lies at -217 veh/h for all configurations within the train-, test- and validation set. Hereby, it can be concluded that for the majority of configurations within the train-, test- and validation set, the neural network error lie within the range of the FOSIM standard deviations.

Finally, some weaving configurations are not taken into account during training and validating the neural network. These configurations are represented by the configurations where FOSIM was not able to simulate capacity values for. This was due to the fact that FOSIM is restricted in the maximum traffic flow per lane (see Paragraph 4.4).

The neural network should ideally estimate high capacity values for these configurations. To illustrate: since FOSIM was not able to simulate capacity values for these configurations, the estimated capacity value by the neural network should lie around the 'base capacity' value. This occurs primarily in the case of configurations with low ratios of weaving traffic, a high or low division of traffic flow, long lengths and/or low ratios of heavy traffic.

Figure 6.15 shows a regression plot of the base capacities (target values; x-axis) and the output of the neural network (output; y-axis) for the configurations which could not be considered by the neural network. This figure shows that only a small share of the capacity values is estimated above the black dotted line.

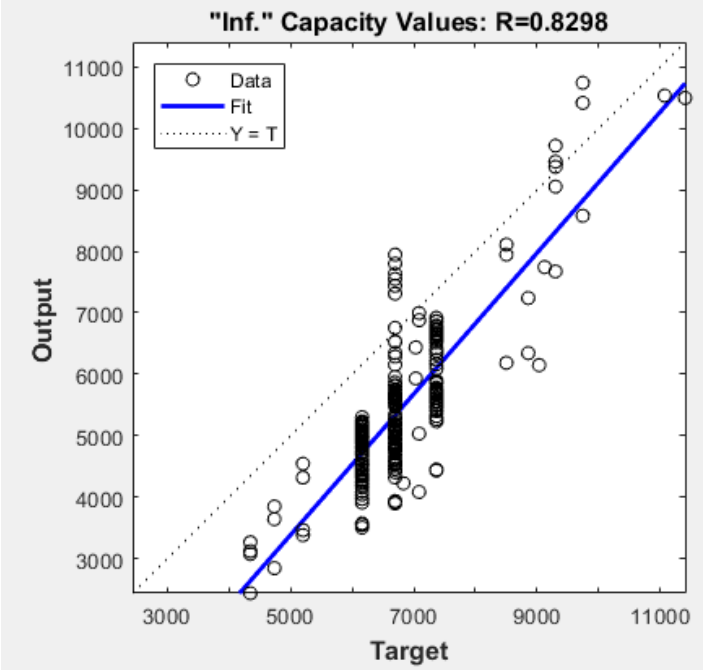



**Figure 6.15: Regression plots of the estimated capacity values for the "∞-configurations", where the target values are represented by the base capacity for that certain configuration (i.e. standard capacity for number of lanes and ratio of heavy traffic (reduction factors)).**

When assuming capacity values for these configurations (should) lie around the 'base capacity' value, the model performance on these configurations is substantial lower. Therefore, eventual follow-up studies can be used for improving model applicability by solving the capacity values for these configurations.

## 6.4 Conclusions Post-Training Phase

In this section the output of the neural network is analysed. Moreover, the network is validated on real-case configurations and currently used capacity estimation methods. Regarding these analyses, it can be concluded that the neural network is a convincing improvement to the currently available capacity estimation methods.

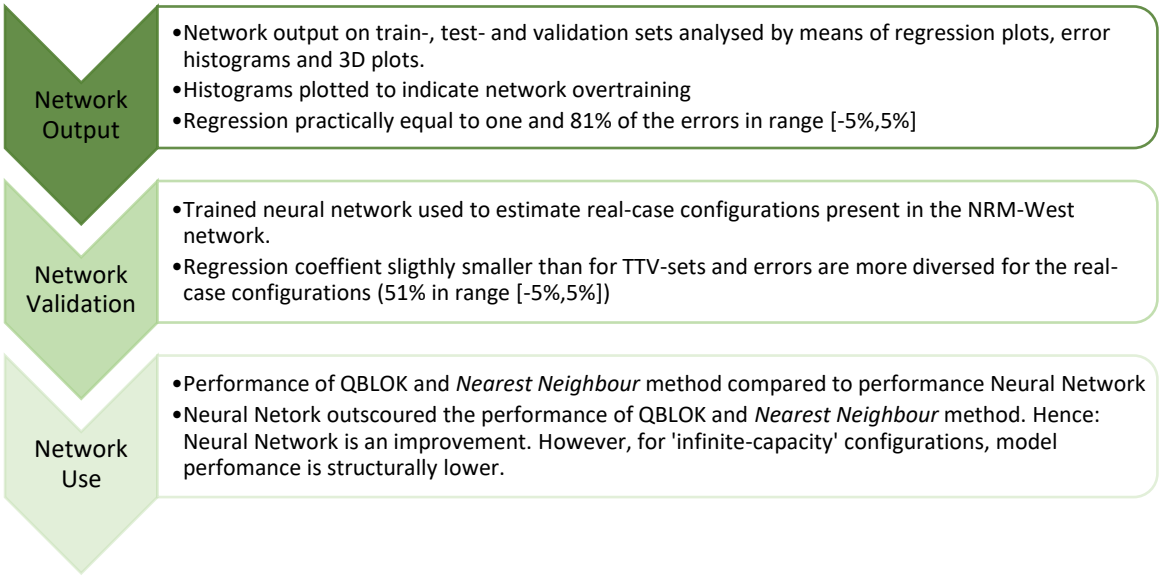



**Figure 6.16: Summary of the Post-Training phase which included the network output analysis, validation and use.**

# 7 Conclusions

The goal of this research was to develop a model which estimates capacity values, suitable for implementing within macroscopic capacity constrained traffic assignment models, by means of a set of explanatory variables. This research developed a neural network which estimates capacity values using the values of twelve independent parameters. Furthermore, the goal of the research was made more specific by adding that the model should estimate capacity values relatively fast, deliver truthful capacity values and is an improvement compared to the currently used capacity estimation methods. The developed neural network succeeded in delivering estimated capacity values fast which are an improvement to the currently used capacity estimation methods. As stated, since the output of this model is sufficient - high regression coefficients and errors in a respectable range – it can be stated that a neural network succeeded in achieving the research goal. Moreover, the neural network outscored the performance of the currently used capacity estimation methods. In the next paragraphs the answers on the research questions are elaborated separately.

First of all, a suitable capacity definition for implementing in the meta model was necessary to find. It was found that the *free capacity* definition in *CIA* was corresponding with the used capacity definition within the macroscopic traffic assignment models. Since capacity is determined by a set of independent variables, a large set of variables were inventoried. However, not all variables did significantly influence capacity, were not relevant for implementing within the model or were not available in the macroscopic traffic model or microscopic simulation model. Finally, the weaving configuration, configuration length, weaving ratios, ratio of weaving traffic and division of traffic flows were stated to be significant and available variables. Despite the fact that the speed limit and speed differences were significantly influencing capacity, these variables are not implemented because of unavailability in macroscopic traffic models. This problem was found to theoretically overcome by implementing speed limits or speed reductions for sharp turns within the macroscopic traffic model.

Secondly, it was found that a simply closed-form (mathematical) solution would not have been suitable for the capacity estimation problem. Therefore, several data-analytic models were assessed, using predictive performance, pattern recognition and pattern recognition. It was found that neural networks were best capable of pattern recognition and predictive performance. To train the neural network, data has been gathered. Data to train the network was gathered in *CIA*, where weaving configuration to validate the model were found in the NRM-West network. Supplementary data, and capacity values of the weaving sections in the NRM-West network, have been derived from FOSIM simulations. However, during simulations it was found that the capacity values in *CIA* were not reproducible. Therefore, the complete datagrid of *CIA* is re-simulated to prevent a bias within and between datasets.

Afterwards, a trial and error procedure (iterative) have determined the best manner of implementing and structuring the variables within the neural network. It was found that the best performance was realised when a set of twelve independent variables (*$Lanes_{O1}$, $Lanes_{O2}$, $Lanes_{D1}$, $Lanes_{D1}$, TotalLanes, %HeavyTraffic (HT), Configuration Length, Weave Ratio 1, Weave Ratio 2, Volume Ratio, Speed Reduction Factor*) was used to estimate the capacity values as dependent variable.

Finally, several network architectures and settings have been implemented and trained. Trial and error procedures have found that a neural network with 10 neurons in one hidden layer resulted in the best performance. Furthermore, *tansig* and linear transfer functions are found to deliver the best performance. The network output was analysed, afterwards, using mean squared errors, regression plots, error histograms and 3D plots as performance indices. It was found that the regression coefficient of the neural network is equal to 0.997 for the train-, test- and validation set. Moreover, the trained neural network was validated on real-case configuration, which are present in the NRM-West network. The performance of the neural network on these configurations, expressed as the regression coefficient, is equal to 0.98. At last, a comparative study is carried out for the use of a neural network as capacity estimation method versus the two currently used capacity estimation methods. Here, it was found that performance of the neural network outscored the performance of the currently used capacity estimation methods, the *QBLOK* and a *Nearest-Neighbour* method. The regression coefficients of these methods on the real-case configurations were, respectively, equal to 0.83 and 0.82. These values were indeed outscored by the neural network with a regression coefficient of 0.98.

Concluding, all research questions of this research have been answered and the research succeeded in developing a model for estimating capacity values for macroscopic traffic models using a set of explanatory variables. The developed neural network delivers, in accordance with the goal of the research, relatively fast capacity values, truthful capacity values and is an improvement of the currently used capacity estimation methods.

# 8 Discussion

As stated in the conclusion, the neural network succeeded in the goal of this research, namely developing a model that can estimate capacity values relatively fast, delivers truthful capacity values and is an improvement of the currently used capacity estimation methods.

The predictive performance of the neural network lies in the line of expectations, since previous research to the application of neural network in the field of transport engineering found similar (well-performing) predictive power of the neural network (Awad, 2004; Kadari et al., 2015; Semeida, 2012; Semeida, 2013; Yap et al., 2015). The research of *Awad*, *2004* resembles the most with this research, because *Awad* developed a neural network to estimate capacity values using the capacity values from the exhibits of *HCM.* This research is, however, not a simple repetition of previous research since this neural network is especially developed for Dutch road networks and capacity values, which do differ with the conditions in *HCM*. Moreover, the neural network developed in this research is external validated by means of a set of real-case configurations. This suppplementary validation - next to the validation set split from the supplemented CIA dataset – is executed to prevent that the network is trained on a structural datagrid composed by CIA and/or HCM. This can possibly also explains the fact that the performance on the internal validation set (split from the CIA set) is better than the external validation set, which was not part of the datagrid composed by CIA. Finally, estimated capacity values of the neural network are compared with the capacity values estimated by the currently used methods. From here it followed that the estimation performance of the neural network outscored the performance of the currently used capacity methods.

However, previous research have compared the performance of the neural networks to the performance of other data-analytic models (i.e. multiple (non)linear regression). This study can not demonstrate this, since no other data-analytic models have been trained due to the convincing foundings in literature on the performance of neural networks.

It should however be noted that all capacity values in this research are simulated with FOSIM. This holds for both the supplemented CIA dataset as the NRM-West configurations. Since these capacity values depends on the used settings of FOSIM, an in-depth validation of the settings and FOSIM for weaving sections would be recommended. Moreover, FOSIM provides restrictions to the maximum traffic flow, resulting in configurations where no capacity value can be simulated for. For this reason, configurations, where no capacity value can be simulated for, are excluded from the neural network. Idially, measured capacity values are preffered to use to train and validate the neural network. However, it was found that these capacity values are hard and costly to derive. Furthermore, it was found that capacity is influenced by the maximum speed and speed differences (i.e. due to sharp turns). However, these values were not available in the macroscopic traffic model, whereby it was not possible to implement these in the neural network. It is therefore reccommendend to implement these variables in the macroscopic traffic simulation models.

Furthermore, a neural network is stated to be a black-box. In other words, the traceability of the capacity estimations is very low. A small analysis has been executed by means of an histogram plot of the relative importance of the explanatory variables and *neural interpretation diagrams*. However, this do not give a foolproof insight in the hidden layer of the neural network. Moreover, it was found that neural networks are not capable of dealing with extrapolations. Therefore, it is recommended that the neural network is also validated on other real-case configuration from other Dutch road networks. In this research it is (simply) assumed that the NRM-West configurations is representable for the (complete) Dutch road network. This should however be validated in future research. Extrapolations could result in exceptional high or low capacity values (i.e. negative capacity

values). This is also a disadvantage of a neural network compared to other methods, since other methods are more related to a 'realistic fundament'. In other words, a nearest neighbour method will never predict negative capacity values.

For implementing the neural network within macroscopic traffic simulation models, which is desirable to improve model outputs, the developed neural network should be transferred, first, to a code or model which can directly be implemented in macroscopic traffic models as OmniTRANS. Furthermore, it should be researched in which manner the neural network should be implemented in the model: the main two directions of implementing the neural network is by implementing it within the dynamic network loading sub-model or by implementing it after every iteration. Besides, future research should eventually find an improvement of the neural network for the configurations that were not able to be considered. For these configurations, FOSIM was not able to deliver a capacity value, while these configurations are present in real-case.

All in all, it is recommended for future research that capacity values can be measured or simulated for all weaving configurations without any restrictions. Moreover, it should be found if the NRM-West model is a sufficient representation of the Dutch road network, to prevent extrapolations of the neural network. Furthermore, to improve capacity estimations, compared to the real situation, speed limits and speed reductions are recommended to implement within traffic assignment models. Finally, future research should be executed to bring a verdict for the manner of implementation of this neural network, as capacity estimation method, within traffic assignment models.

# Bibliography


Adamowski, J., Chan, H. F., Prasher, S. O., Ozga-Zielinski, B., & Sliusarieva, A. (2012). Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neral network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. *Water Resources Research*.

Al-Jameel, H. (2011). Towards the Simulation of Ramp Weaving Sections. University of Salford, Manchester.

Al-Kaisy, A., J. Stewart, & Van Aerde, M. (1999). A simulation approach for examining capacity and operational performance at freeway diverge areas. *Canadian Journal of Civil Engineering*, 760-770.

Awad, W. H. (2004). Estimating traffic capacity for weaving segments using neural networks technique. *Applied Soft Computing*, 395-404.

Barton, R. R., & Meckesheimer, M. (2006). Chapter 18 Metamodel-Based Simulation Optimization. In S. G. Henderson, & B. L. Nelson, *Simulation* (pp. 535-574). Amsterdam: North-Holland; Elsevier.

Beale, M. H., Hagan, M. T., & Demuth, H. B. (2018). *Neural Network Toolbox™; User's guide.* Natick, MA: The MathWorks, Inc.

Bliemer, M. C., Raadsen, M. P., Brederode, L. J., Bell, M. G., Wismans, L. J., & Smith, M. J. (2017). Genetics of traffic assignment models for strategic transport planning. *Transport Reviews*, 56-78.

Bliemer, M., Brederode, L., Wismans, L. J., & Smits, E. (2012). Quasi-dynamic traffic assignment: static traffic assignment with queueing and spillback. *The Transportation Research Board (TRB) 91st Annual Meeting, Washington DC, January 22-26, 2012* (pp. 1-24). Washington, DC, USA: Transportation Research Board (TRB).

Burghout, W., Koutsopoulos, H., & Andreasson, I. (2006). A discrete-event mesoscopic traffic simulation model for hybrid traffic simulation. *Intelligent Transportation Systems Conference*.

Cremer, M. (1979). *Der Verkehrsfluß auf Schnellstraßen: Modelle, Überwachung, Regelung.* Springer-Verlag.

Dharia, A., & Adeli, H. (2003). Neural network model for rapid forecasting of freeway link travel time. *Engineering Applications of Artificial Intelligence*, 607-613.

Dijker, T., & Knoppers, P. (2006). *Fosim 5.1 Gebruikshandleiding.* Delft: Delft University of Technology; Transport & Planning Department.

Dijker, T., & Minderhoud, M. (2001). *CIA-2: Deelonderzoek CAPWEEF - Fase 2; Het bepalen van de capaciteit van Asymmetrische Weefvakken met Simulatie.* Ministerie van Verkeer en Waterstaat.

Duliba, K. A. (1991). Contrasting Neural Nets with Regression in Prediction Performance in the Transportation Industry. *System Sciences*.

FOSIM. (2018). *Beschrijving model*. Retrieved from FOSIM: https://fosim.nl/beschrijving-model.shtml

Geistefeldt, J. (2011). Capacity effects of variable speed limits on German freeways. *Procedia Social and Behavioral Sciences*, 48-56.

Gora, P., & Bardoński, M. (2017). Training neural networks to approximate traffic simulation outcomes. *Models and Technologies for Intelligent Transportation Systems (MT-ITS)*.

Grontmij. (2009). *Capaciteitsmetingen autosnelwegen - Samenvatting van de resultaten voor 75 locaties.*

Hagan, M. T., Demuth, H. B., Beale, M. H., & De Jesús, O. (2014). *Neural Network Design; 2nd edition.* Martin Hagan.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The Elements of Statistical Learning; Data Mining, Inference, and Prediction.* Springer.

Henkens, N., Mieras, W., & Bonnema, D. (2017). *Validatie FOSIM.* De Bilt: Sweco.

Hidas, P. (2005). Modelling vehicle interactions in microscopic simulation of merging and weaving. *Transportation Research Part C: Emerging Technologies*, 37-62.

Ibrahim, O. (2013). A comparison of methods for assessing the relative importance of input variables in artificial neural networks. *Journal of Applied Sciences Research*, 5692-5700.

Intini, P., Colonna, P., Berloco, N., & Ranieri, V. (2016). The Impact of Route Familiarity On Divers' Speeds, Trajectories and Risk Perception.

Jiang, X., & Adeli, H. (2004). Clustering-Neural Network Models for Freeway Work Zone Capacity Estimations. *International Journal of Neural Systems*, 147-163.

Kadari, B. R., Vedagiri, P., & Rathi, N. (2015). Models for pedestrian gap acceptance behaviour analysis at unprotected mid-block crosswalks under mixed traffic conditions. *Transportation Research Part F*.

Kleijnen, J. P. (2007). Kriging metamodeling in simulation: a review. *CentER Discussion Paper Series*.

Leferink, B. G. (2013). *The Continious Line Continued...; A research to the effect of a continious line at a highway access.* Enschede.

Ligterink, N. (2016). *On-road determination of average Dutch driving behaviour for vehicle emissions.* Utrecht: TNO.

Livingstone, D., Manallack, D., & Tetko, I. (1997). Data modelling with neural networks: Advantages and limitations. *Journal of Computer-Aided Molecular Design*, 135-142.

Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A*, 291-305.

Marczak, F., Daamen, W., & Buisson, C. (2013). Key variables of merging behaviour: empirical comparison between two sites and assessment of gap acceptance theory. *Social and Behavioral Sciences*, 678-697.

Miaou, S.-P. (1994). The relationship between truck accidents and geometric design of road sections: poisson versus negative binomial regressions. *Accident Analysis & Prevention*, 471-482.

Minderhoud, M. M., Botma, H., & Bovy, P. H. (1997). Assessment of roadway capacity estimation methods; Highway capacity issues and analysis. *Transportation research record*, 59-67.

Nordfjærn, T., Jørgensen, S., & Rundmo, T. (2010). An investigation of driver attitudes and behaviour in rural and urban areas in Norway. *Safety Science*, 348-356.

Olden, J. D., & Jackson, D. A. (2002). Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling* , 135-150.

Payne, H. J. (1979). FREFLO: A Macroscopic Simulation Model of Freeway Traffic. *Transportation Research Board*, 68-77.

Polus, A., Craus, J., & Livneh, M. (1991). Flow and Capacity Characteristics on Two-Lane Rural Highways. *Transportation Research Record*.

Possel, B. (2017, september 14). Verbetering Modellering Reistijden: Netwerkkalibratie.

Rakha, H., & Crowther, B. (2002). Comparison of Greenshields, Pipes, and Van Aerde Car-Following and Traffic Streams Models. *Transportation Research Record: Journal of the Transportation Research Board*.

Rakha, H., & Zhang, Y. (2006). Analytical Procedures for Estimating Capacity of Freeway Weaving, Merge, and Diverge Sections. *Journal of Transportation*, 618-628.

Rijkswaterstaat . (2018, Juni 04). *[Wegenkaart] Maximumsnelheden vanaf 1 juli 2018.* Retrieved from Ministerie van Infrastructuur en Waterstaat: http://publicaties.minienm.nl/documenten/wegenkaart-maximumsnelheden-per-1-juli-2018

Rijkswaterstaat. (2015). *Capaciteitswaarden Infrastructuur Autosnelwegen.* Dienst Water, Verkeer en Leefomgeving (Rijkswaterstaat).

Rijkswaterstaat. (2017). *Handreiking Bewegwijzeringsschema's.* Rijkswaterstaat; Ministerie van Infrastructuur en Milieu.

Rijkswaterstaat. (2017, april 2017). *Inhaalverbod voor vrachtverkeer vernieuwd.* Retrieved from Rijkswaterstaat: https://www.rijkswaterstaat.nl/wegen/wetten-regels-en-vergunningen/verkeerswetten/inhaalverbod-vrachtverkeer/

Rijkswaterstaat. (2018, August 30). *Verkeers- en vervoersmodellen LMS en NRM*. Retrieved from Rijkswaterstaat: https://www.rijkswaterstaat.nl/wegen/wegbeheer/aanleg-wegen/nederlands-regionaal-model-nrm-en-landelijk-model-systeem-lms.aspx

Roess, R. P. (1988). Development of Weaving Area Analysis Procedures for the 1985 Highway Capacity Manual. *Transportation Research Record 1112*, 17-22.

Semeida, A. M. (2013). New models to evaluate the level of service and capacity for rural multi-lane highways in Egypt. *Alexandria Engineering Journal*, 455-466.

Smith, B. L., & Demetsky, M. J. (1997). Traffic flow forecasting: comparison of modeling approaches. *Journal of Transport Engineering*, 261-266.

Soriguera, F., Martínez, I., Sala, M., & Menéndez, M. (2017). Effects of low speed limits on freeway traffic flow. *Transportation Research Part C*, 257-274.

Transportation Research Board. (2000). *Highway Capacity Manual (HCM).* Washington, DC: Transportation Research Board, National Research Council.

Treiber, M., & Kesting, A. (2013). *Traffic Flow Dynamics; Data, Models and Simulation.* Springer.

van Beinum, A., Farah, H., Wegman, F., & Hoogendoorn, S. (2016). A critical assessment of methodologies for operations and safety evaluations of freeway turbulence. Washington: TRB.

Van Beinum, A., Farah, H., Wegman, F., & Hoogendoorn, S. (2018). Driving behaviour at motorway ramps and weaving segments based on empirical trajectory data. *Transportation Research Part C*, 426-441.

Vermijs, R. (1997). *Capaciteitswaarden symmetrische weefvakken (CIA-1).* Delft.

Vermijs, R. (1998). New Dutch Capacity Standards for Freeway Weaving Sections Based on Micro Simulation. *Transport Research Board*.

Wardrop, J. G. (1952). Some theoretical aspects of road traffic research. *Proceedings of the Institute of Civil Engineers, Part II*, 325-378.

Yap, Y. H., Gibson, H. M., & Waterson, B. J. (2015). Models of Roundabout Lane Capacity. *Journal of Transport Engineering*.

Zhang, Y. (2005). *Capacity Modeling of Freeway Weaving Sections.* Blacksburg, Virginia: Virginia Polytechnic Institute and State University.

| | Eigen FOSIM instellingen | | CIA | | Sweco validatierapport | | Capweef-1 (symmetrisch) | Capweef-2 (assymetrisch) | HCM |
|---|---|---|---|---|---|---|---|---|---|
| | *Symmetrisch* | *Assymetrisch* | *Symmetrisch* | *Assymetrisch* | *Symmetrisch* | *Assymetrisch* | *Symmetrisch* | *Assymetrisch* | |
| HB Patroon (ruimteherkomsten) | Zowel evenredige verdeling als assymetrisch (als variable) | | Intensiteiten evenredig verdeeld over stroken | | Intensiteiten everendig verdeeld over de stroken (4+1 = I_hrb = 4* I_inv) | | Intensiteiten evenredig verdeeld (4) | Intensiteiten evenredig verdeeld (4) | Analytische benadering |
| HB Patroon (tijdintensiteitverloop) | Maximale intensiteit op t=5400, t=0: 0,5*max intensiteit (1) | | X | | Intensiteit neemt 75%tgl/strookklu toe tot max. intensiteit op t=5400 | | Verkeersaanbod neemt toe met 100à300 mvt/hr/str | X | |
| Strookwisselgebieden | i.o.m. Sweco rapport | *Zie toelichting rechts* | X | | Gehele weefvaklengte verplicht, elke extra strook +150m stroomopwaarts | | Verplicht: WVL-100 + 1100 gewenst; extra 400 verplicht bij meer stroken (3) | Varieren per configuratie, gebaseerd op bebording (5) | |
| Lokale kenmerken op weefvakken x+1 | Geen snelheidsreductie op invoegende stroom (zie beneden) | | X | | X | | Invoegende stroom SPRD 80% (3) | X | |
| Infrastructuur x+1 weefvakken | 1-strooks herkomsten hebben lengte van 1000m ipv 1500m (6) | | X | | X | | 1-strooks herkomsten hebben een kortere lengte | X | |
| Lokale kenmerken op weefvak 1+1 | Snelheidsonderdrukking 60% komt meest overeen met CIA (z.b.) | | Representeren rangeerbanen bij klaverbladen | | Geeft aan dat SPRD boogstralen kunnen nabootsen (geen kwantificatie) | | Snelheidsonderdrukking 60% (2) | X | |
| Invoegers | Verschillende ratio's invoegend verkeer | | Geen verschillende invoegratio's ! | | Verschillende verhoudingen (3) intensiteit HRB+INV doorgerekend i.t.t. CIA | | X | X | |

1) Maximale intensiteit is zelf vastgesteld a.d.h.v. FOSIM experimenten per aantal stroken en percentage zwaar verkeer
2) Snelheidsonderdrukking van 60% op weefvakken van 1+1, echter meer vervolgonderzoek naar validiteit gewenst
3) Strookwisselgebieden: bij een rijstrookwissel verplicht weefvaklengte -100m; 100 m gewenst. Voor meerdere strookwisselingen Weefvaklengte +300 meter verplicht + aanvulling tot 1200m voor het puntstuk gewenst (zie ander tabblad)
4) Meer onderzoek gewenst naar assymetrische wegvakbelasting
5) Per configuratie (bijv. 3+1-2+2) meerdere combinaties van strookwisselgebieden gegeven, nl. één met de aanwezigheid van voorwegwijzer (bebording) en een met de afwezigheid van voorwegwijzer. Bij voorwegwijzer: grofweg/meestal SWGB van 1000m verplicht + 400 gewenst, zonder voorweg.wijzer: grofweg SWGB van weefvaklengte + 300m gewenst (zie ander tabblad)
6) Eén-strooks herkomsten hebben een kortere lengte ivm herkomsten met meerdere stroken om collonevorming achter langzaam rijdende voertuigen te voorkomen (inhalen nl. niet mogelijk)

**Toelichting: Snelheidsreductie op invoegende stromen van 1-strooks herkomsten**

Test op 3+1 weefvak met snelheidsreductie (80% zie Capweef-1) en zonder snelheidsreductie levert resultaten die laten zien dat Capaciteiten beter overeenkomen met CIA zonder snelheidsreductie.

Echter: Bij een snelheidsonderdrukking komt het minder vaak voor dat voertuigen de bestemming niet bereiken. Met snelheidsreductie halen gemiddeld 0.2 voertuigen per run NIET de gewenste bestemming. Zonder snelheidsreductie zijn dat 3.7 voertuigen per run, met een max van 11 per run. (Output levert niet of deze voertuigen de bestemming niet halen voor of ná het optreden van congestie)

Desondanks gekozen voor configuratie zonder snelheids-onderdrukking vanwege betere overeenkomst met capaciteitswaarden in CIA.

| Configuration Code | Capacity CIA | Capacity Fosim | %-afwijking | Capacity Fosim SPRD | %-afwijking |
|---|---|---|---|---|---|
| [3131,600,17,50,5,120] | 8830 | 8742 | -1 | 8454 | -4.3 |
| [3131,700,17,50,5,120] | 9170 | 8892 | -3 | 8496 | -7.4 |
| [3131,800,17,50,5,120] | 8960 | 8946 | -0.2 | 8748 | -2.4 |
| [3131,600,17,50,15,120] | 7520 | 7374 | -1.9 | 7254 | -3.5 |
| [3131,700,17,50,15,120] | 7570 | 7422 | -2 | 7356 | -2.8 |
| [3131,800,17,50,15,120] | 7600 | 7512 | -1.2 | 7458 | -1.9 |
| [3131,600,17,50,25,120] | 6430 | 6516 | 1.3 | 6366 | -1 |
| [3131,700,17,50,25,120] | 6480 | 6636 | 2.4 | 6432 | -0.7 |
| [3131,800,17,50,25,120] | 6530 | 6636 | 1.6 | 6492 | -0.6 |
| [3131,600,25,75,5,120] | 8400 | 8052 | -4.1 | 7872 | -6.3 |
| [3131,700,25,75,5,120] | 8530 | 8298 | -2.7 | 8016 | -6 |
| [3131,800,25,75,5,120] | 8660 | 8454 | -2.4 | 8298 | -4.2 |
| [3131,600,25,75,15,120] | 6970 | 6726 | -3.5 | 6654 | -4.5 |
| [3131,700,25,75,15,120] | 6840 | 6900 | 0.9 | 6792 | -0.7 |
| [3131,800,25,75,15,120] | 7120 | 7116 | -0.1 | 6876 | -3.4 |
| [3131,600,25,75,25,120] | 6020 | 5976 | -0.7 | 5928 | -1.5 |
| [3131,700,25,75,25,120] | 6120 | 6048 | -1.2 | 5886 | -3.8 |
| [3131,800,25,75,25,120] | 6120 | 6192 | 1.2 | 6012 | -1.8 |
| [3131,600,33,100,5,120] | 7700 | 7536 | -2.1 | 4914 | 36.2 |
| [3131,700,33,100,5,120] | 7990 | 7746 | -3.1 | 5004 | 37.4 |
| [3131,800,33,100,5,120] | 8140 | 8028 | -1.4 | 4968 | -39 |
| [3131,600,33,100,15,120] | 6530 | 6396 | -2.1 | 4224 | 35.3 |
| [3131,700,33,100,15,120] | 6440 | 6594 | 2.4 | 4278 | 33.6 |
| [3131,800,33,100,15,120] | 6770 | 6684 | -1.3 | 4368 | 35.5 |
| [3131,600,33,100,25,120] | 5630 | 5556 | -1.3 | 3870 | 31.3 |
| [3131,700,33,100,25,120] | 5780 | 5730 | -0.9 | 3912 | 32.3 |
| [3131,800,33,100,25,120] | 5880 | 5874 | -0.1 | 3972 | 32.4 |

**Toelichting: Snelheidsreductie op 1+1 weefvakken**

Verschillende batch-runs voor 1+1 weefvakken uit het CIA zijn uitgevoerd met verschillende snelheidsonderdrukkingen over het hele weefvak. Hieruit blijkt dat, in overeenstemming met Capweef-1, dat een snelheidsreductie van 60% capaciteitswaarden oplevert die het dichtst bij CIA liggen. Na een snelheidsreductie van 60% ligt de (max.) snelheid op 72 km/u.

Configuration code bestaat uit: [Lanes H1, H2, B1, B2; Lengte; Weefratio 1; Weefratio 2; %HT; Max.Speed]

| Configuration Code | Capacity CIA | Capacity Fosim 80 | Capacity Fosim 70 | Capacity Fosim 65 | Capacity Fosim 60 |
|---|---|---|---|---|---|
| [1111,350,50,50,5,120] | 2680 | 3474 | 3210 | 3060 | 3000 |
| [1111,350,50,50,15,120] | 2320 | 3054 | 2796 | 2664 | 2568 |
| [1111,350,50,50,25,120] | 2150 | 2760 | 2568 | 2484 | 2382 |
| [1111,350,75,75,5,120] | 2360 | 3234 | 3234 | 2802 | 2634 |
| [1111,350,75,75,15,120] | 2110 | 2808 | 2538 | 2448 | 2238 |
| [1111,350,75,75,25,120] | 1940 | 2598 | 2370 | 2292 | 2196 |
| [1111,150,100,100,5,120] | 2100 | 2994 | 2250 | 2160 | 2136 |
| [1111,150,100,100,15,120] | 1850 | 2568 | 1974 | 1920 | 1896 |
| [1111,150,100,100,25,120] | 1680 | 2418 | 2058 | 1788 | 1740 |
| [1111,200,100,100,5,120] | 2140 | 2970 | 2346 | 2208 | 2172 |
| [1111,200,100,100,15,120] | 1860 | 2508 | 2040 | 1956 | 1908 |
| [1111,200,100,100,25,120] | 1720 | 2400 | 2130 | 1860 | 1794 |
| [1111,350,100,100,5,120] | 2170 | 3156 | 2754 | 2376 | 2232 |
| [1111,350,100,100,15,120] | 1930 | 2688 | 2358 | 2094 | 2004 |
| [1111,350,100,100,25,120] | 1780 | 2484 | 2220 | 2148 | 1926 |
| Max. Snelheid na speedreduction >> | | 96 | 84 | 78 | 72 |
| Gemiddeld aantal veh. Dat verkeerde bestemming bereikt >> | | 1.64 | 0.43 | 0.24 | 0.16 |

**Toelichting: Strookwisselgebieden assymetrische weefvakken**

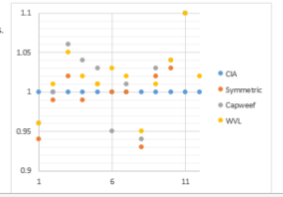
Op basis van drie voorkomende strookwisselgebied-instellingen zijn er op verschillende configuraties simulatieruns uitgevoerd. Uit tabel blijkt dat de Sweco instellingen (zelfde als symmetrisch) toch het beste overeenkomen met de capaciteitswaarden van CIA. Daarnaast geven alle configuraties logische uitkomsten. In andere woorden: bij een toename van lengte ook een toename van capaciteit, bij een reductie van de weefratio's een toename van capaciteit, bij een afname van %HT ook een toename van capaciteit.

Alleen bij taper-configuraties blijkt de instellingen van Sweco het minst goed te preseteren i.v.m. de CIA-capaciteitswaarden. Desondanks (voor de reden van uniformiteit) ook hier de strookwissel-gebieden op dezelfde wijze ingesteld.

In de onderste puntenwolk staan de resultaten van de simulatieruns in een iets andere vorm, namelijk de CIA capaciteiten zijn hier blauw en gelijk aan 1 (index = 1) . De capaciteitswaaren van de fosim runs zijn uitge-drukt t.o.v. de index. Hier valt te zien dat de oranje punten (symmetric strookwisselgebieden) algemeen gezien het dichst bij de CIA-waarden liggen, met wel wat enkele uitschieters.

Echter: Taper in- en uitvoegers in CAPWEEF-2 zijn niet meer in overeenkomst met de huidige taper configuraties. In CAPWEEF-2 zijn tapers namelijk niets meer dan een extra strook die vervelat naar 200 meter van het puntstuk.

| Configuration Code | Capacity CIA | FOSIM: Symmetric | FOSIM: Capweef-2 | FOSIM: WVL+ 300m GW |
|---|---|---|---|---|
| [2122T, 600, 50, 50, 15, 120] | 4010 | 3756 | 3864 | 3852 |
| [2231, 850, 25, 75, 15, 120] | 7450 | 7356 | 7422 | 7500 |
| [4132, 1000, 44, 75, 25, 120] | 6730 | 6834 | 7164 | 7050 |
| [4132, 1100, 44, 75, 25, 120] | 6920 | 6870 | 7218 | 7086 |
| [4132, 1000, 38, 50, 25, 120] | 7400 | 7444 | 7626 | 7470 |
| [2T21, 650, 42, 75, 15, 120] | 4730 | 4752 | 4500 | 4854 |
| [2T21, 850, 42, 75, 15, 120] | 4880 | 4896 | 4918 | 4992 |
| [2T21, 850, 17, 50, 15, 120] | 5410 | 5052 | 5070 | 5136 |
| [2112, 700, 63, 25, 15, 120] | 5320 | 5430 | 5454 | 5376 |
| [2112, 900, 63, 25, 05, 120] | 6070 | 6282 | 6288 | 6336 |
| [2112, 900, 88, 75, 25, 120] | 4080 | 4476 | 4500 | 4500 |
| Mean square error: | 39215 | 69462 | 46873 | |
| MSE excl. Taper-config: | 33993 | 81557 | 55378 | |
| MSE Taper-config: | 48355 | 48295 | 31990 | |




Appendix A1