1. Zonder een goed wederzijds begrip van de deelsystemen spoorwegen, intern trasport, overslag en opslag is het niet mogelijk de optimale spoorterminal voor een omgeving als de Maasvlakte te ontwerpen.

2. Overslag op treinen wordt veelal als bijzonder eenvoudig aangemerkt. In een omgeving als de Maasvlakte is dit echter allerminst waar vanwege (1) het gebruik van MTS-trailers, (2) het grote aantal "tappunten" op de Maasvlakte, (3) de deels strikte beladingseisen voor treinen, en (4) de directe aanvoer van laadepoplen over grote afstanden.

3. Simulatie van een systeem heeft, naast het nut van analyse, een niet te onderschatten belang voor het doorgronden van dat systeem. Dit laatste is vaak zelfs van nog groter belang dan de eigenlijke analyse. Voorwaarde is wel, dat zowel de model-specificatie als de simulaties door dezelfde personen worden uitgevoerd.

4. De uiteindelijk gevonden oplossing voor een probleem wordt in hoge mate beïnvloed door de gekozen oplossingsmethode. Het is daarom aan te raden om via verschillende ingangen een oplossing te zoeken.

5. De oplossingsmethode voor het transportation problem en het transshipment problem zijn, in die gevallen waar het op te lossen probleem daadwerkelijk een transport-probleem betreft, voornamelijk van belang voor bedrijven die hun vervoerscapaciteit bij derden inkopen – en niet zozeer voor bedrijven die over een eigen wagenpark beschikken.

6. Het gebruik van optimaliseringstechnieken is een in de praktijk veronachtzaamde en onderschatte methode om tot een goed ontwerp te komen. Hier valt nog veel winst mee te behalen.

7. De inzet van ingenieurs voor het ontwerpen van een spoorterminal is niet altijd een gelukkige keuze; in veel gevallen zal een ervaren lid van de drie-ploegendienst op een spoorterminal ook tot een goede oplossing komen.


9. Het verschil tussen een westers mens en een vervelend kind is niet zo groot als men wel denkt. Beiden hebben namelijk de eigenschap de hun gegeven mogelijkheden tot het uiterste uit te buiten, en waar mogelijk te proberen hun grenzen te overschrijden. In het geval van een vervelend kind leidt dit vaak tot opstandig gedrag en een blauw zitvlak, doch in het geval van een westers mens vanwege zijn al te grote welvaart tot een, wat eufemistisch genoemd, verhoogd consumptief gedrag. Hierdoor legt de westere mens een meer dan rechtmatig beslag op de schaarse grondstoffen en energiebronnen op de aarde zonder zich hiervan de gevolgen te realiseren. Zolang de westere mens zich niet van dag tot dag bewust is van zijn rol op aarde, zal hij niet goed met zijn welvaart om kunnen gaan en zichzelf niet kunnen beperken. Grappig is, dat iedereen het hiermee eens is. Totdat de volgende vliegvakantie naar China wordt geboekt.

1Waarmee wordt bedoeld: de volwassen westere mens; het was niet de bedoeling om vervelende kinderen als onmenselijk te betitelen.
A structured terminal design method

With a focus on rail container terminals

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof.ir. K.F. Wakker,
in het openbaar te verdedigen ten overstaan van een commissie,
door het College van Dekanen aangewezen,
op dinsdag 28 maart 1995 te 16.00 uur

door

Erik Jan Alexander VAN ZIJDERVELD,

werktuigkundig ingenieur,
geboren te Papendrecht.
Dit proefschrift is goedgekeurd door de promotor:
Prof.ir. M. van Holst

Samenstelling promotiecommissie:
Prof.dr.ir. W.J. Beek
Prof.Dr.-Ing. I.A. Hansen
Prof.dr.ir. R. Dekker
Prof.dr.ir. J.J.M. Evers
Prof.drs. C.J. Ruijgrok
Prof.dr.ir. G.J. Wormmeester
Prof.ir. C.P. Keizer

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Zijderweld, Erik Jan Alexander van

A structured terminal design method : with a focus on rail container terminals / Erik Jan Alexander van Zijderweld. - Delft :
Delft University of Technology, Faculty of Mechanical Engineering and Marine Technology. - III.
Thesis Technische Universiteit Delft. - With index, ref. - With summary in Dutch.
ISBN 90-370-0120-3
Subject headings: terminal complexes ; design / container terminals / rail terminals.

© Copyright 1995 by Erik Jan Alexander van Zijderweld.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form
or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of the author.

PRINTED IN THE NETHERLANDS
BY UNIVERSITEITSDRUKKERIJ, DELFT
## Contents

Acknowledgements............................................................................................................. vii
Summary ................................................................................................................................. ix
Samenvatting .......................................................................................................................... x

**Chapter 1. Introduction** ..................................................................................................... 1

**Chapter 2. Introduction to idiom for transport systems** .................................................. 3

**Chapter 3. Current and future situations at the Maasvlakte** .............................................. 9
3.1. Introduction ..................................................................................................................... 9
3.2. Current situation ............................................................................................................. 9
3.3. Future situation ............................................................................................................. 9

**Chapter 4. Contemporary terminal design methods** ......................................................... 13
4.1. Introduction ................................................................................................................... 13
4.2. Design cases ................................................................................................................. 13
4.3. Design tools .................................................................................................................. 15
4.4. Pros and cons of contemporary design methods.......................................................... 16

**Chapter 5. Proposal for a computer aided terminal design method** ................................. 17
5.1. Introduction ................................................................................................................... 17
5.2. Starting-points.............................................................................................................. 17
5.3. Subdivision into three levels of detail........................................................................ 20
5.3.1. Overview of elements at UCM level of detail........................................................... 23
5.3.2. Overview of elements at SDM level of detail........................................................... 28
5.3.3. Overview of elements at DDM level of detail........................................................... 31
5.4. Aimed results ............................................................................................................... 33
5.4.1. Aimed results at UCM level of detail..................................................................... 33
5.4.2. Aimed results at SDM level of detail..................................................................... 35
5.4.3. Aimed results at DDM level of detail..................................................................... 38

**Chapter 6. Load schedules and contra schedules** ............................................................ 39
6.1. Introduction ................................................................................................................... 39
6.2. Load schedules............................................................................................................. 42
6.3. Contra schedules ......................................................................................................... 46
6.3.1. Regular contra schedules ....................................................................................... 48
6.3.2. Intensity contra schedules ..................................................................................... 51
6.4. Matching load schedules with contra schedules........................................................ 53
6.4.1. Regular contra schedules ...................................................................................... 54
6.4.2. Intensity contra schedules ..................................................................................... 57
6.4.3. Timing of load units .............................................................................................. 60
6.5. Some examples of the use of load schedules and contra schedules............................... 67
6.5.1. Train-to-sea transport........................................................................................... 67
6.5.2. Sea-to-sea transport.............................................................................................. 68
6.5.3. Sea-to-truck transport............................................................................................ 69

**Chapter 7. Equipment** .................................................................................................... 71
7.1. Introduction ................................................................................................................... 71
7.2. Elements for internal transport ................................................................................... 72
7.2.1. Equipment types..................................................................................................... 72
7.2.2. Equipment pools .................................................................................................... 75
# Contents

7.2.3. Equipment networks ................................................................. 76
7.2.4. Equipment routings ................................................................. 77
7.3. Options for internal transport ...................................................... 81
7.3.1. Flow patterns ................................................................. 81
7.3.2. Local internal transport ......................................................... 82
7.3.3. Stacking ................................................................. 84
7.3.4. Transshipment ................................................................. 85
7.3.5. Closed- and broken internal transport ....................................... 86
7.3.6. Equipment relocation ........................................................... 87
7.4. Optimisation of equipment usage at SDM level .................................. 88
7.4.1. Basic definitions ................................................................. 90
7.4.2. Two optimisation models ....................................................... 95
7.4.3. INTRA-window equipment hopping ............................................ 96
  7.4.3.1. Load unit sub-model .......................................................... 96
  7.4.3.2. Equipment sub-model ........................................................ 107
  7.4.3.2.1. Multiple load unit equipment types ...................................... 124
  7.4.3.2.2. Pulling equipment types ................................................... 126
  7.4.3.3. Coupled load unit sub-model and equipment sub-model .............. 143
7.4.4. Inter-window equipment hopping ............................................ 150
  7.4.4.1. Load unit sub-model .......................................................... 152
  7.4.4.2. Equipment sub-model ........................................................ 154
  7.4.4.3. Coupled load unit sub-model and equipment sub-model .............. 155
7.4.6. Conversion of continuous results into discrete results .................. 157

## Chapter 8. Stacks .............................................................................. 163

8.1. Introduction ............................................................................. 163
8.2. Local stacking and normal stacking ............................................ 164
  8.2.1. Local stacking at bounding elements ......................................... 164
  8.2.2. Local stacking at train handle areas ........................................... 165
  8.2.3. Stacking at internal stacks ....................................................... 167
8.3. Initial stack sizes ................................................................. 169

## Chapter 9. Railway aspects .............................................................. 171

9.1. Introduction ............................................................................. 171
9.2. Standard operations at rail service centres .................................... 171
  9.2.1. Handling ........................................................................ 172
  9.2.2. Local movements ................................................................. 172
  9.2.3. Coupling ........................................................................ 174
  9.2.4. Brake tests ....................................................................... 174
  9.2.5. Locomotive storage ............................................................... 175
9.3. Train concepts ......................................................................... 175
  9.3.1. Shuttle types - introduction ..................................................... 175
  9.3.2. Shuttle types - options ............................................................ 178
  9.3.3. System entry- and exit operations ............................................ 181
  9.3.4. E-Locomotive operations ......................................................... 182
  9.3.5. D-Locomotive operations ......................................................... 183
  9.3.6. Handling operations ............................................................... 184
9.4. Train services ........................................................................... 191
  9.4.1. Timetables ...................................................................... 191
  9.4.2. Operations ...................................................................... 192
  9.4.3. Loading and unloading ............................................................ 193
  9.4.4. Bundle selection .................................................................. 197
9.5. Infrastructure ............................................................................ 197
  9.5.1. Rail bundles .................................................................... 197
  9.5.2. Rail points ....................................................................... 198
  9.5.3. Rail tracks ...................................................................... 200
  9.5.4. Rail routes ........................................................................ 201
  9.5.5. Block safety systems ............................................................. 201
Acknowledgements

Born on a windy day, the dwarf was left alone in the moors. A small turf hut was his shelter against the cold winds and stormy nights. Every now and then a kind visitor from outside of the moors fed the dwarf and gave him some attention to make him feel more comfortable in his hut... and the dwarf started to grow.

Some day, the dwarf was wandering over the gloomy moors, the landscape he had become familiar with, when suddenly a gull was next to him. The silver gull, serene as the moon, spoke to him and thus showed him a way out of the moors. Not long after this occurrence a small caravan of wise men, three in number, entered the moors and visited the dwarf in sympathy. They asked the dwarf to change his scenery from the gloomy moors to the impressive oak woods far away. Pleased as could be, the dwarf accepted the offer and set off for the oak trees, accompanied in distance by the wise men.

On his way to the woods of oak, the dwarf came across a broad river, flowing silently from east to west and from west to east, but no one could tell its depth. Its serenity was deceiving as the river was known to be dangerous to anyone who did not know him well. Therefore, to cross the river the dwarf had to develop some new skills which he was willing to do. Helped by the three wise men, he got acquainted with the dangers of the river and with the techniques to overcome these dangers. After a while the dwarf managed to assemble a sufficiently solid raft and crossed the river, accompanied by the silver gull which reappeared every now and then next to him.

Now seeing the woods of oak from afar, the dwarf strolled into a lawn of soft grass from which several small oak trees popped up. These small oak trees were tended by little people living in the fields, and were tested regularly to guarantee a good quality of the breed. When the oak trees were found mature and of good quality they were removed from the lawn by the little people and were planted in the oak woods.

The dwarf was fascinated by the works of these little people and deepened his knowledge about their procedures to grow the oak trees. Doing so he found that these little people developed an entirely new breeding plan for each individual oak tree, which cost much effort and resulted in an unfair situation in which one young oak tree was given substantially more attention and care than others. The dwarf decided that he might help to stop this situation of injustice by developing a standard breeding plan that could be used for almost any oak tree, offering equal care to all young trees. This plan would benefit both the little people and the young oak trees, the dwarf believed. The three wise men supported the idea and allowed the dwarf to stroll around a little more in the lawn of the little people.

After some time the dwarf came up with a general breeding plan and it was accepted by the three wise men; they allowed him to put some more effort in the breeding plan to complete it. The scope of this work appeared to be thus broad that the dwarf was not able to perfect his breeding plan, and when his time was come only part of his ambitions had become true. But this did not distress the dwarf as he believed the breeding plan he had made was good and could serve as a basis for further development.

The dwarf never made it to the oak trees. But he did not care, as he had found his rest in the lawn with the little people.

♦ Concluding remarks
Thanks to mom, dad, Irene, Judith, Maarten, Conny, Jos, Johan, Dodo, Jeanine, Majel, Ben, Gabri, and Dick for their support in their own way. Special thanks to the three wise men Anko Nagel, Mouriets van Schuylenburg and Cees Horvers, the latter being replaced after good services by Maarten Huug Bode. Further thanks to Ruud van der Ham, Hans Schuitj, Benjamin Jansen, Piet Koning, Ben Immer, Hans Bruins, Cor Labuyère, Cor Hoenders, Jaap Otjes and Jaap Vreugdenhil. And, last but certainly not least, to my professor and promotor prof.ir. M. van Holst.

Summary

This thesis describes a model to design and analyse logistic aspects of terminal complexes at which standard load units like containers and swap bodies are handled, while focusing primarily on the role of rail terminals within these terminal complexes. The model aims to facilitate and standardise analysis of terminal complexes, thus encouraging development of new terminal concepts and improving insight in the interrelation of various elements in those terminal complexes. New about this model is the ability to model and analyse this interrelation, with relatively little effort and to a relatively high level of detail, even for very complicated terminal complexes. This new approach is now feasible thanks to late improvements in computer technology. Part of the model has been implemented in a computer programme that runs under Windows™ and is called TORCMOD\(^1\).

The terminal complexes that can be described by the model may consist of any combination of following types of locations:
- rail terminals,
- marine terminals,
- barge terminals,
- truck gates, and
- stacks,

each type of location having its own characteristics. One of the main characteristics is the position of the location relative to the boundary of the modelled system: at or inside the system’s boundary. Between these various locations load units can be transported by means of internal transport. The model offers several general structures to model and analyse almost any configuration for internal transport, and to determine the optimal flow of load units across the entire terminal complex. This optimal flow then answers to aspects like preferences on the presence of load units at the various locations, and preferences on the use of (competing) equipment pools on particular transport links. It also includes the mutual influence of all individual flows of load units across the terminal complex.

Because of its focus on rail terminals the model also contains numerous operational aspects that are found at these rail terminals and are characteristic for railway operations. These aspects cover amongst others brake tests, claiming of infrastructure and light engines.

One of the major problems when modelling large-scale terminal complexes is to specify and balance the flows of load units at the boundaries of the modelled system. Proper balancing is required for analysis of the modelled system for longer periods of (say) over twenty-four hours. The model described in this thesis always guarantees a balanced system by using load schedules and contra schedules that are linked by means of equipment routings and flow patterns. This structure allows to analyse the modelled system for a period of time that is longer than the previously mentioned twenty-four hours.

The model is based on three different levels of detail which are called UCM, SDM and DDM respectively. These levels of detail are discerned with the intention to facilitate modelling, to improve insight in the modelled system, and to allow quickly different types of analysis with the same model.

For as far as completed, the implementation of the model has been applied for the future Maasvlakte terminal complex at Rotterdam. The thus obtained analysis results are validated against those found by another study on the same terminal complex, which was performed by a logistic consultants firm.

\(^1\)Abbreviation of: Transhipment Of Rail Containers Modelling tool.
Een gestructureerde terminal ontwerp methode
Met de nadruk op rail container terminals

Samenvatting

Dit proefschrift beschrijft een model voor het ontwerp en de bestuiving van terminal complexen die zijn ingericht voor de overslag van standaard laaduitheden als containers en wissellaadbakken. Het zwaartepunt van het model ligt bij de *spoorterminals* binnen dergelijke terminal complexen. Het doel van het model is, om het bepalen van het gedrag van dergelijke terminals te vergemakkelijken en te standaardiseren, en om zo verder onderzoek naar nieuwe terminal concepten te stimuleren. Tevens heeft het model tot doel om het inzicht in de complexe processen die zich op dergelijke terminals afspeLEN te verbeteren. Het vernieuwende aspect van het model is het modelleren van de verschillende processen op deze terminals in hun samenhang, tegen een relatief geringe modelleringsinspanning en tot op een relatief hoog niveau van detail, zelfs voor zeer complexe situaties. Het model is gedeeltelijk geïmplementeerd in een computer programma dat onder Windows™ draait en TORCMOD² heet.

De terminal complexen die met het model beschreven kunnen worden, kunnen een willekeurige configuratie van de volgende locatie typen zijn: (1) rail terminals, (2) zee terminals, (3) binnenvaart terminals, (4) vrachtwagen poorten, en (5) opslagplaatsen, waarbij ieder van deze locatie typen zijn eigen karakteristieken heeft. Eén van de belangrijkste karakteristieken hierbij is, of de locatie al dan niet op de *grens* van het gemodelleerde systeem ligt.

Tussen al deze locaties kunnen laadehenden worden getransporteerd door middel van *intern transport*. Het model bevat een aantal algemene structuren en elementen waarmee dit interne transport gemodellerd en geanalyseerd kan worden, en waarmee tegelijkertijd de *optimale* inzet van het interne transport bepaald kan worden. Deze optimale inzet is onder andere afhankelijk van de opgegeven voorkeuren voor de aanwezigheid van laadehenden op de verschillende locaties, en van de voorkeuren voor het gebruik van (onderling vervangbare) transportmiddelen op bepaalde transportverbindingen. Bij deze optimalisatie wordt tevens rekening gehouden met de wederzijdse beïnvloeding van de verschillende stromen en de beschikbare vrijheid in planning die voor de diverse stromen kan verschillen.

Eén van de belangrijkste problemen bij het modelleren van grootschalige terminal complexen, is het garanderen van een *evenwicht* in de stromen die het terminal complex binnenkomen en weer verlaten. Dit evenwicht, wat nodig is voor het uitvoeren van analyses die langere tijdperiodes bestrijken (ruwweg vanaf 24 uur), wordt binnen het model altijd gegarandeerd door gebruik te maken van de zogenaamde *load schedules* en *contra schedules* die met elkaar zijn verbonden door middel van de zogenaamde *equipment routings* en *flow patterns*. Hierdoor wordt het mogelijk om analyses voor langere periodes dan 24 uur met het model uit te voeren.

Aangezien het zwaartepunt van het model ligt bij de spoorterminals, bevat het model diverse operationele aspecten die kenmerkend zijn voor deze spoorterminals. Voorbeelden hiervan zijn remproeven, blok beveiligingssystemen en losse locomotieven.

Binnen het model worden drie verschillende detail niveaus onderscheiden, die UCM, SDM en DDM worden genoemd. Het doel van deze drie detail niveaus is, om het modelleren te vergemakkelijken, om het inzicht in het gemodelleerde systeem te vergroten, en om op eenvoudige wijze verschillende typen analyses met hetzelfde model uit te voeren.

Voor zover geïmplementeerd, is het model ook daadwerkelijk toegepast voor het toekomstige terminal complex op de Maasvlakte. De zo verkregen resultaten zijn vergeleken met de resultaten die voor dezelfde Maasvlakte zijn gevonden door een logistic consultants bureau, en gebruikt voor validatie van het model.

---

² Afkorting van: Transhipment Of Rail Containers Modelling tool.
Chapter 1. Introduction

Future through growth. This appears to be the main slogan for freight transport in the next few years. Whatever report on future freight streams is consulted, they all predict a vast increase of transport volumes in the next few years. No exception to this rule are the predicted freight streams to and from the port of Rotterdam, as can be seen from figure 1.1. This figure shows the number of containers\(^1\) that is to be handled in future, both for Rotterdam as a whole and for the Maasvlakte, which is part of Rotterdam.

All freight streams inevitably induce transport, so streams of containers induce transport, too. Transport by means of marine vessels, coasters, barges, road trucks or trains. In a number of cases these containers are transported by more than one means of transport on their way from origin to destination; this is called multi-modal transport. A typical example hereof is given in figure 1.2.

Figure 1.2 shows how containers are transported by a multi-modal transport chain which consists of three partial transports: One partial transport by means of a road truck, one by means of a train, and a second one by means of a road truck. All means of transport cover a part of the total way from the container’s origin to the container’s destination.

\* Focus on terminals

Two special locations can be seen from figure 1.2 which are marked as "terminals". At these terminals the containers are transshipped between truck and train, and vice versa. It is this type of location at which focus is put in this thesis: Terminals for transshipment of containers\(^2\) from and onto trains. Which is not constrained to terminals as shown in figure 1.2 only in which containers are transshipped between trains and road trucks, but also covers large-scale terminal complexes at which multiple rail terminals have to co-operate with one or more marine terminals or barge terminals as found at the future Maasvlakte. These latter terminal types will be explained in Chapter 2.

\* Need for new terminal capacity

Multi-modal transport is believed to play an increasingly important role in future, especially multi-modal transport in which trains are involved as depicted in figure 1.2. Some examples of reports that state this shift towards rail transport are SVV-II/D [29], NEA [21] and CPB [3].

When combining the growth of future freight streams and the increasing importance of rail transport, it is obvious that rail terminals will play a more important role in future. Therefore a new need for rail terminal capacity will arise in the next few years, which can be satisfied either by expanding the capacity of existing rail terminals or by constructing totally new rail terminals.

In both cases the future rail terminals will have a capacity which is great compared to current rail terminals and by consequence will have a relatively complicated structure and will be costly in terms of required investments. In this way a need to predict the capacity of these new rail terminals in an accurate way before they are actually constructed is created. The costs involved with this preparatory research of new rail terminals, prior to actual

---

\(^1\) Refer to Chapter 2 for a picture of a container.

\(^2\) And other standard load unit types.
construction, are small compared to the expenses that may occur in case of a malfunctioning terminal that needs adaptations after construction.

♦ Need for new terminal design methods
To obtain a reasonable prediction of the actual capacity of a newly constructed rail terminal some kind of terminal design method is needed. Contemporary terminals are mainly developed by (1) calculations by hand and (2) computer simulation models that are written for that specific new terminal. While these methods were adequate for the smaller rail terminals they will not be sufficient for future, larger rail terminals as costs would become too high: Calculations by hand become thus extensive that they either are not accurate enough anymore or take too much time. Computer simulation models, specifically written for one rail terminal, become complicated too and therefore take much time to develop. In combination with the ever increasing demand for new and larger rail terminals, a new demand for structured and integral design methods for larger rail terminals will arise.

At this moment several computer applications exist that can be used for parts of the design of rail terminals. Some examples hereof can be found in Pierick [23], Schwanhäußer [26] and Sitzmann [28], [27], which describe how to develop rail infrastructure of sufficient capacity, either based on queuing theory or on discrete computer simulation. As these applications cover only part of the functionality of a rail terminal they are of limited use only. For this reason some other computer applications offer an integral analysis of rail terminals (for example KLV-Simu by Brunner [21]), but these applications can only be used for inland rail terminals - and not for partially marine bound terminal complexes as found at the Maasvlakte. Refer to Chapter 4 for more details on these methods.

In summary, a terminal design method that can be used for integral analysis of a terminal complex as found at the future Maasvlakte system is not available yet, and therefore needs to be developed. This new terminal design method should guarantee a sufficiently detailed prediction of a large rail terminal that is not constructed yet, while in the meantime it should eliminate the drawbacks of contemporary design methods. A proposal for such a design method is given in this thesis. It is partially implemented in a computer programmed that is called TORCMOD and runs under Windows™.

The new terminal design method described in this thesis covers a number of elements that is found at rail terminals and at other terminals like marine terminals and truck terminals. A quick impression of these elements can be obtained from Chapter 2 while an overview of the way these elements are included in the new terminal design method is given in Chapter 5. All elements mentioned in Chapter 5 are elaborated in more detail in Chapters 6 through 9, of which Chapter 6 describes flows of load units across the system's boundary, and Chapter 7 inside the system's boundary. Finally, an example of the use of this new method for the future Maasvlakte situation is given Chapter 10 along with the results that were obtained by using the method.

3Refer to Chapter 4 for more details.
4Most of these computer applications focus primarily on design of main lines with more or less regular time tables, rather than on design of railway yards. Nevertheless, they may be used for parts of railway yards.
5An abbreviation of Transhipment Of Rail Containers Modelling tool.
6For as far as currently implemented in TORCMOD.
7Refer to Chapter 3.
Chapter 2. Introduction to idiom for transport systems

This thesis focuses on finding a method to facilitate development and analysis of new terminal complexes for the transhipment of standard load units like containers. In order to increase understanding of the elements that can be found at such terminal complexes and of the coherence between them, this chapter gives a short overview of most of these elements, based on an example of the future situation at the Maasvlakte. This chapter, however, does not aim to give a detailed insight in all options and all processes; it just aims to give a rough impression of the elements that are found at most terminal complexes.

This future situation at the Maasvlakte is shown in figures 2.6 and 2.7: Figure 2.6 depicts an artist's impression of the future Maasvlakte, while figure 2.7 marks a number of important elements that can be seen from figure 2.6. The elements that can be seen from figure 2.7 are:

1. load units (containers)
2. marine vessels
3. marine quays
4. marine quay-craines
5. marine terminals
6. marine stacks
7. barge terminal
8. barges
9. rail service centre
10. rail tracks
11. routes for internal transport.

A number of these elements will be given some more attention.

Load units
The model described in this thesis is developed for terminals at which standard load units are transhipped only; it is not developed for terminals at which bulk like coal or wheat is transhipped, nor oil or gas. The most important type of standard load unit is the ISO container of which an example is given in figure 2.1. Some other examples of standard load units are the EuroContainer and swap body (not depicted). The main difference between EuroContainers and ISO containers are the dimensions: While ISO containers are mainly used for sea-transport by means of specially designed vessels, EuroContainers are mainly used for continental transport by means of road trucks. The main differences between swap bodies and ISO containers are (1) the inability to stack swap bodies and (2) the way both types of load units are picked up: While ISO containers are picked up at the “holes” in the top four corners (the corner fittings), swap bodies are only picked up from the bottom. Even more, swap bodies are exclusively used for continental transport by road trucks and are designed such that they interface in an optimal way with these road trucks.

Terminals
The basic idea about terminals is that standard load units like containers are transhipped from one means of transport (like a marine vessel) that delivers a load unit onto another means of transport (like a road truck) that fetches a load unit. By consequence a terminal is a kind of complex interface between both means of transport.

Most of the times it will not be possible to tranship directly between both means of transport for one of the following reasons: (1) Direct transhipment would make the processes at the terminal too complex; in case of transhipment from a marine vessel onto road trucks this would result in complex controlling of all individual road trucks, to make sure that they arrive in the right order to fetch their container. (2) Direct transhipment would yield a complex terminal design for those terminals that include more than two modes of transport: All modes of transport that are handled should come very close to each other, which immediately causes difficulties.
for terminals that include barges, marine vessels, road trucks and trains. (3) Direct transhipment would require the simultaneous presence of both means of transport between which load units are transhipped. Especially direct transhipment of load units between marine vessels and trains is virtually impossible due to the great diversity of destinations of load units, the strict loading sequences of trains and vessels, and the great length of trains.

For these various reasons load units are generally stored between both transhipment movements, which is illustrated in figure 2.2. The middle block of this figure labelled "storage and handling" stands for all processes that occur at the terminal between both transhipment movements, which in reality are more complex than just suggested. Important to note is the dotted box in figure 2.2 which is called "system's boundary". This boundary indicates that load units either enter or leave the considered terminal by one of both transhipment movements, which will turn out to be an important aspect for the model that is described in this thesis. By consequence, the boundary is always formed by terminals, i.e. locations at which load units are transhipped onto or from means of transport like marine vessels, barges, trains or road trucks.

Figure 2.2: Decoupling both transhipment movements through storage.

Figure 2.3: Cross-section of a terminal complex, consisting of (1) a marine terminal at the left- and (2) a barge terminal at the right-hand side [17].

This aspect of decoupling both transhipment movements by terminals is also depicted in figure 2.3, showing a cross-section of the Maasvlakte terminal system from figure 2.6. In this cross-section all processes inside the terminal between unloading from a marine vessel (at the left-hand side) and loading onto a barge (at the right-hand side) are shown; all these processes are part of the box labelled "storage" of figure 2.2, which in fact is more complicated than just storage of load units between two subsequent transhipment movements.

From figure 2.3 a number of operations that are specific for terminals for standard load units like containers can be seen. These operations are labelled ① through ⑤ in figure 2.3 and will be discussed one by one. The operation labelled ① represents the "first transhipment" from figure 2.2 through which a load unit enters the terminal system, and therefore is always located at the system's boundary. As suggested previously, this can either be transhipment from a marine vessel, a barge, a train or a road truck. The operation labelled ⑤ can be considered as the counterpart of operation ①: It represents the "second transhipment" from figure 2.2 through which a load unit leaves the terminal system. By consequence, this operation is always located at the system's boundary, too.

Between these two transhipment operations three other operations are discerned, two of which are labelled "internal transport". This internal transport refers to transport of load units inside the system's boundary.
Normally, this internal transport is done by means of machines (equipment) that is owned by the considered terminal. This in contrast to the marine vessels and trains that are no part of the terminal and therefore do not remain all time at the terminal. Some examples of the various types of machines that exist will be given shortly.

The final operation from figure 2.3 to discuss is operation ⑨ which is labelled "stacking". In contrast to the previously discussed operations of internal transport, at which load units are transported across the terminal complex, during the operation of stacking load units remain in rest. In other terms, during stacking load units remain on the same location at the terminal complex.

Stacking equipment
To place load units into the stack and to retrieve them from the stack again (i.e. to stack and to unstack them) some special machines are needed. The type of machine shown in figure 2.4 is generally called a rail mounted gantry crane, abbreviated to RMG\(^1\). This type of crane drive over special rail tracks which allows fast and accurate movements of the gantry crane.

Next to the rail mounted gantry crane some other related crane types exist: These are the rubber tyred gantry crane (abbreviated to RTG) and the bridge crane. The main difference between the rubber tyred gantry crane and the RMG is that the first one drives on rubber tyres and the latter one drives on rail. The main difference between the bridge crane and the RMG is that the first one is supported on one side by a levelled rail.

The locations at which load units are stacked are called stacks. These locations play an important role as they serve to decouple subsequent movements of internal transport. They also can be used for sorting of load units by grouping for example those load units that have the same destination in one part of the stack. From figure 2.6 a number of stacks can be seen.

Internal transport equipment
Next to the equipment for stacking and unstacking load units as discussed at figure 2.4 some other types of equipment are found at most terminals. These equipment types are used for internal transport, which corresponds to operations ⑧ and ⑨ of figure 2.3. Two different types of equipment for internal transport, commonly used at the Maasvlakte, are discussed in this chapter: Automated guided vehicles (AGVs) and straddle carriers (SCs).

An example of an AGV as used at ECT container terminals, Rotterdam, is given in figure 2.5. Most particular about these AGVs is their lack of a driver. They can drive around the terminal without the direct interference of a human operator which reduces operational costs. Automated guided vehicles are controlled by a complex computer system that communicates with the AGVs by means of, for example, radio signals. It is expected that for large container terminals AGVs will be used more commonly in the near future as they increase flexibility in planning internal transport\(^2\).

As can be seen from figure 2.5 AGVs can be used for horizontal transport of containers only; they can not be used for transhipment of containers as any gear needed for transhipment lacks. An advantage of this selection is that the AGVs remain relatively simple and therefore relatively cheap. A drawback of these AGVs is that they always need the presence of some other equipment, like gantry cranes, to load and unload them.

\(^1\)In fact, this particular RMG is a very special, automated version that is used at ECT container terminals, Rotterdam.

\(^2\)A drawback of automated internal transport is the lack of flexibility in solving unexpected problems, an ability that is one of the strong points of human operators.
Figure 2.6: *Artist's impression of the future Maasvlakte system [30]*.
Figure 2.7:  *Elements in the future Maasvlakte system.*
The second type of equipment for internal transport is the *straddle carrier (SC)* as shown in figure 2.8. In contrast to AGVs these SCs are both manned and capable of transshipping load units\(^3\).

When stacking load units these SCs drive over the pile of load units that are arranged in long rows or "streets". Between both rows a path is available for the straddle carrier to drive over, forcing the straddle carrier to drive all the way over such a "street" of containers. In this way relatively much potential stacking space is wasted for these paths, reducing the number of load units that is stacked per square meter compared to gantry cranes as shown in figure 2.4. But in practice this drawback is commonly compensated by stacking containers up to four high at stacks that are served by straddle carriers, thus increasing the number of containers stacked per square meter\(^4\).

---

\(^3\)By consequence, SCs are a typical example of equipment incorporating a mix of functions: They can be used both for pure tranship and for pure internal transport.

\(^4\)In the meantime introducing the problem of "digging" in a four-high pile of containers for that particular one that is always at the bottom...
Chapter 3. Current and future situations at the Maasvlakte

3.1. Introduction

Things are changing rapidly in container industries. Not only the number of containers that is transported will increase dramatically in the next few years (refer to Chapter 1), but also the terminals at which these containers are transshipped are at a turning point: New equipment types like automated guided vehicles (AGVs, refer to Chapter 2) are developed, and new organisational forms are investigated, needed to cope with the increased complexity of the container terminals.

A good example of these quick changes can be found at the Maasvlakte. The Maasvlakte is part of the port of Rotterdam, is located near to the sea and is largely addicted to container transhipment. In Chapter 2 an artist’s impression of the future Maasvlakte system has been shown, along with an explanatory map showing the situation in 2010. To these future developments at the Maasvlakte some more attention is paid in this chapter.

Future container flows

Container flows at the Maasvlakte will be great in future without doubt; some forecasting models even predict several millions of containers to be handled yearly at the Maasvlakte in 2010! Depending on the scenario that is studied a considerable part of these containers will leave the Maasvlakte by the same transport mode as had brought them there: Marine vessels. This concept is called majority from sea to sea (MSS) and still is considered as an increasingly important concept in future.

Next to these sea-to-sea flows a great number of containers will be transhipped onto and from trains, even up to 850,000 a year, depending on the forecasting model that is considered. These flows of rail containers require the presence of a large rail terminal - larger than most contemporary rail terminals for containers.

3.2. Current situation

In the current 1994 situation only a relatively small rail terminal is present at the Maasvlakte, which is located directly behind the stack of the ECT-Delta marine terminal. This is shown in figure 3.1, which is based on the artist’s impression previously shown in Chapter 2.

Next to the ECT-Delta terminal, figure 3.1 shows the Delta-SeaLand terminal which is a marine terminal, too. The main difference between both marine terminals is the type of machines that are used for transport and transhipment of containers: ECT-Delta uses conventional machines like trailer systems and straddle carriers\(^1\), both manually operated, while Delta-SeaLand uses mainly newly developed machines like AGVs and automated stacking cranes (ASCs)\(^2\), both fully automatic and driver-less. This latter type of automatic equipment is believed to play an increasingly important role at future container terminal complexes.

This is a consequence of the increased complexity of the terminal system and of the increased number of containers that is to be transported and transhipped; using manually operated equipment would make the complex internal transport process uncontrollable and would induce extremely high personnel costs.

3.3. Future situation

Therefore some new marine terminals, terminals that are currently under construction at the Maasvlakte, will be equipped with highly automated machines, like the Delta-SeaLand terminal already is. These terminals, shown in figure 3.2, are part of the entire Delta 2000-8 project plan which provides a total package of solutions to

---

\(^1\)Refer to Chapter 2 for an example of a straddle carrier.

\(^2\)Automated stacking cranes (ASCs) are small, fully automated rail mounted gantry cranes.
manage the augmenting container flows at the future Maasvlakte. Especially the lower terminals, which are named 2000-I through 2000-III for convenience, play an important role in this. Refer to information of the Port of Rotterdam [30].

The rail terminal, or rail service centre, in the future Maasvlakte situation is shown at a different location in figure 3.2 than previously shown in Chapter 2. This is a consequence of the uncertainty that still exists about the precise, optimal location of this future rail service centre. The main issue is, however, that this terminal is expected to be unusually large.

The final two locations to discuss from figure 3.2 are DistriPark and MT-stack. The DistriPark is a large-scale industrial area in which containers are loaded, unloaded and repaired, and where other services related to marine container shipping are offered. The MT-stack is a separate stacking facility, independent of all marine terminals, at which empty containers (abbreviated to MTs) are stacked. Whether or not this MT-stack will actually be constructed depends on the outcome of various studies on the Maasvlakte terminal system.

- **ITT Internal transport system**

Between all marine terminals and other locations at the Maasvlakte a special internal transport system will be installed, which is called *inter terminal transport system (ITT-system)*. This ITT-system is to take care of all transport of containers between the various terminals at the Maasvlakte, and consequently will be rather complicated. Especially planning of flows of containers that are strongly correlated and all have different characteristics (like different priorities) is complicated\(^3\). Therefore, this ITT-system will inevitably be equipped with automated machines after some time, as it appears to be virtually impossible to manage such a complex transport system with manned equipment only. This will be illustrated by an example.

As all marine terminals, the DistriPark has a strong relation with all other locations at the Maasvlakte, like the rail service centre and the MT-stack, and therefore has a strong relation with the ITT-system. This shared interest in the ITT-system might yield, for instance, the situation that containers are to be transported from DistriPark to a marine terminal, while in the meantime several other containers are to be transported from some marine terminal to the rail service centre. As both transport movements claim and use capacity of the same ITT-system, these two movements are likely to *clash* in claiming transport capacity; transport capacity that is limited. A result of this clash could be that transport related to the rail service centre is postponed somewhat in favour of transport from DistriPark; which inevitably has consequences for the operations at the rail service centre.

This example readily shows that clashes that occur because of shared use of the same transport capacity by various container flows might influence operations throughout the entire Maasvlakte system; through the ITT-system all activities at the Maasvlakte are thus strongly interrelated and interdependent that only an *integral* approach, covering all activities at the Maasvlakte\(^4\) can yield valid results that predict future behaviour of the particular parts of the Maasvlakte. This will be a main issue for the model developed in this thesis.

---

\(^3\)This thesis therefore aims to give a structure that might be used to increase understanding of such a complex system and to determine the optimal configuration of it.

\(^4\)Possibly at different levels of detail.
Chapter 3. Current and future situations at the Maasvlakte

Figure 3.1: Current situation at the Maasvlakte (1994).

Figure 3.2: Future situation at the Maasvlakte (2010).
Chapter 4. Contemporary terminal design methods

4.1. Introduction

Planning new terminal complexes and rail service centres as discussed in the previous chapters is done by a few people only. This simply is a consequence of the relative small amounts of terminal complexes and rail service centres that is constructed, like there are very few places at which, for instance, knowledge on the design of nuclear plants is available. Concentrating knowledge on special subjects can, depending on the nature of the studied subject, have the unpleasant drawback that very little research for improved methods, like improved design methods of rail service centres, can be done. Nevertheless, some progress is made in developing these latter design methods.

This chapter gives an overview of the currently available methods to design rail service centres and then shows the drawbacks and advantages of these methods. This will be done by referring (1) to actual case studies as found in practice, and (2) to design tools that are already available to improve designing these rail service centres.

- Confidential information
Some information, especially on actual case studies as found in practice, is considered to be confidential and therefore can not be used explicitly in this chapter. Instead, it will be explained in general terms what the design process of rail service centres is like without giving too much details. In a number of cases it is not allowed to refer to confidential reports, which request will be respected.

4.2. Design cases

Information on the design process of rail service centres is hard to obtain. Normally, this kind of information is not provided freely for a number of reasons. One reason could be that this information contains data on future freight streams that needs to be kept secret; another reason could be that the knowledge on the design of rail service centres itself is considered as secret. But fortunately in some cases people are inclined to provide information on the design process of rail service centres. A survey of the most important available cases, on which information did become available, is given below.

Table 4.1 gives all four actual, important, design cases on which detailed information on the design process is available. Next to these major example cases information on a number of smaller, less interesting cases is available. In most of these cases the actual design process consists of two steps:
1. static analysis of at most five variants on terminal lay-out or operations;
2. dynamic analysis (i.e. simulation) of the most promising variant(s).

These two steps are always preceded by extensive studies on the numbers of trains that are to be handled at the rail service centre that is to be designed and on the timetables of these trains. As this preliminary step is considered inevitable it is given no further attention. The remaining two steps just mentioned, however, will be explained in more detail.

<table>
<thead>
<tr>
<th>Source</th>
<th>Rail service centre</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deutsche Eisenbahn Consulting</td>
<td>Bremen - Niedervieland</td>
</tr>
<tr>
<td>ECT</td>
<td>Rotterdam - Margriethaven Zuidzijde</td>
</tr>
<tr>
<td>ECT / Logitech / Tebodin</td>
<td>Rotterdam - RSC Maasvlakte</td>
</tr>
<tr>
<td>Deutsche Bahn</td>
<td>_1</td>
</tr>
<tr>
<td>SNCF / PEA Paris</td>
<td>Noisy-le-Sec (Commuter)</td>
</tr>
</tbody>
</table>

Table 4.1: Most important design example cases.

1Not related to one specific rail terminal; general design rules which are comparable to the first step to be discussed shortly.
Chapter 4. Contemporary terminal design methods

- Static analysis

The first step of the design process is always static analysis; this kind of analysis is based on conventional calculation techniques like rules of thumb and planning tables which give a rough insight in the numbers of tracks needed at the rail service centre and in the stacking capacity that is needed under the gantry cranes. Main aspect of analysis made during the first step is that all calculations can be written out on paper, either as formulas or as tables. This in contrast to the second step, as will be discussed shortly.

During static analysis normally one to five variants\(^2\) are examined by conventional calculation techniques, resulting in several performance parameters of these variants like the total number of tracks to install, or total costs for the rail service centre. Normally the number of variants is restricted to at most five as analysing each single variant requires several days and possibly up to some weeks. Therefore in practice it is mostly decided to divide static analysis itself into two sub-steps.

In this subdivision of static analysis a first "round" is used to find the most promising variants to analyse in more detail. Normally these are at most three variants and possibly only one. During this first round very simple and straightforward calculations are used to get a quick impression of the benefits of each variant. After the first round the second "round" starts. During this round the most promising variants are analysed in more detail, adding some side-effects like the interrelation between trains and equipment for internal transport. In contrast to calculations in the first round, which may last up to several days, these calculations in the second round may last even two or three weeks per variant.

In some cases the results of the static calculations are considered sufficiently detailed, which especially may hold for smaller rail terminals, and are taken as the final results of the terminal design process. In this case the second step in the design process, dynamic analysis, is not executed.

- Dynamic analysis

In those cases that dynamic analysis is executed a computer simulation\(^3\) model is built for the rail terminal that is under design. For this simulation model either a specialised simulation language like PROSIM\textsuperscript{TM} or SIMAN\textsuperscript{TM} is used, or a general computer programming language like PASCAL or C. In contrast to the first step the calculations made during this second step can not be made on paper; they are performed entirely by the computer, imitating the behaviour of the rail terminal in time. During this imitation random numbers often play an important role which virtually impedes to write calculations on paper.

Dynamic analysis is used to include a number of dynamic effects of the designed rail service centre that could not be included in the calculations made under static analysis. Examples of these dynamic effects are random disturbances of train arrivals relative to the timetables, random numbers of load units, and dynamic interrelation between equipment and trains. Most of the times these dynamic effects have a major impact on the analysis results, and therefore give a far better insight in the performance of the actual rail service centre, once it is built, than static analysis ever could.

Normally for each rail service centre that is to be analysed new simulation model is created, from scratch on. Depending on the differences between the variants that are to be examined, these variants can be analysed by the same simulation model by changing some well-chosen parameters, or they all require a (totally) different simulation model. Constructing a simulation model of a rail service centre takes, depending on the aimed level of detail and its complexity, from several weeks to several months.

In some cases even two different simulation models for the same variant are constructed. This was the case for analysis of the Commutor system in which a first, simple simulation model was used to get a better insight in the newly developed Commutor transhipment technique. After some experiments with the model an entirely new simulation model was created to analyse transhipment by Commutor in more detail. In total several months were involved with creating and testing these two simulation models.

---
\(2\) Which are different scenarios for terminal lay-out, timetables (if these may be changed), train operations, stacking rules and equipment.

\(3\) Refer to Appendix B for some details on computer simulation techniques.
4.3. Design tools

To speed up the second step in the design process, i.e. dynamic analysis, some tools have become available in the last few years. All of these design tools offer some general structure of a rail service centre or another type of terminal that can be modified to model the rail service centre or other terminal that is to be analysed. Each design tool has a different set of standard elements, a different point of view, and a different aimed level of detail.

This section gives only a short overview of the available design tools and their capability. More detailed information can be found in the literature that is referred to. Most of the design tools are available on a commercial basis.

<table>
<thead>
<tr>
<th>Company/Institution</th>
<th>Design tool</th>
<th>Field of application</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>JWD</td>
<td>GMTS</td>
<td>marine terminals</td>
<td>[12]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rail terminals</td>
<td></td>
</tr>
<tr>
<td>Dornier</td>
<td>Kombisim</td>
<td>rail terminals</td>
<td>[6]</td>
</tr>
<tr>
<td>DE-Consult</td>
<td>KLV-Simu</td>
<td>rail terminals</td>
<td>[2]</td>
</tr>
<tr>
<td>ISL</td>
<td>Scusy</td>
<td>marine terminals</td>
<td>[10]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rail terminals</td>
<td></td>
</tr>
<tr>
<td>University of Gdansk</td>
<td>MULTIMOD</td>
<td>bulk terminals</td>
<td>[14]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>marine terminals</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2: Available design tools.

All four design tools will be given some more discussion, although not exhausting.

♦ GMTS
The design tool GMTS is mainly used to analyse marine terminals; although rail terminals can be modelled too, they are only included in a rather coarse way in simulation. It runs under Windows and therefore offers good graphical representation. A major advantage of GMTS is that it offers the possibility to model almost any terminal configuration on-screen. A disadvantage of GMTS is that there are little possibilities to control the flows of containers through the modelled system thus affecting validity of the model. Nevertheless, GMTS gives the impression to be the most mature of all design tools mentioned in table 4.2.

♦ Kombisim
The design tool Kombisim is the oldest of all design tools and is less flexible than GMTS. In contrast to GMTS it focuses entirely on rail terminals and was actually used by the Deutsche Bahn. Although modelling rail terminals followed by analysis is simple with Kombisim, the options that are available are limited. It shows clearly that this design tool was especially created for Deutsche Bahn as all standard configurations for rail terminals used by Deutsche Bahn can be modelled relatively well.

♦ KLV-Simu
The design tool KLV-Simu is meant for the analysis of inland rail terminals that are serviced by road trucks. KLV-Simu allows to set numerous parameters that influence the transhipment process at rail terminals and the arrival of road trucks at the terminal's gate. It uses computer simulation for a 24-hour period that is divided into

---

4 One similarity between most of these design tools is their integral way of modelling rail terminals, which is, including as well rail operations as transshipment and internal transport movements. This in contrast to other computer applications like UX-Simu, PC-Simu as described in Sitzenmann [28], [27], Pierick [23], and Schwachaüller [26], which all cover rail infrastructure analysis only.
5 Abbreviation of: Jordan Woodman Dobson, Oakland, Canada.
6 Abbreviation of: General Marine Terminal Simulation.
7 Friedrichshafen, Germany.
8 Frankfurt/Main, Germany.
9 Abbreviation of: Institute of Shipping Economics and Logistics, Bremen, Germany.
10 Abbreviation of: Simulation of Container Unit Handling Systems.
equal time slices of ten seconds each. KLV-Simu is specifically suitable for the German type of railway operations.

- **Scusy**
The design tool Scusy is clearly designed in good co-operation with marine terminal operators, which shows from the quite pragmatic approach and the way parameters are specified. Scusy offers good possibilities to model most marine terminal configurations and equipment allocation structures, but analysis of rail terminals appear to be poor compared to Kombisim. Except from rail terminals, Scusy gives a more detailed impression than Kombisim does.

- **MULTIMOD**
The design tool MULTIMOD is not solely intended to model and analyse one terminal, but instead it is intended to model an entire logistic chain, based on multi-modal transport. By consequence the level of detail is rather coarse compared to, for instance, Scusy but it can be used for preliminary studies to establish the rough dimensions of almost any terminal. After these preliminary studies one of the more specialised design tools can be used for the actual design of the terminal.

### 4.4. Pros and cons of contemporary design methods

The methods described in the previous two sections give a "state-of-the-art" picture of the methods that are currently available to design terminals in general and rail container terminals in particular. From this discussion it followed that in a number of cases the design process stops after what was called the first step in the design process, which is static analysis. The reason for this premature stop was twofold: Either because the aimed level of detail could be reached by this first step (which is true for smaller terminals), or because the costs that would be involved with the second step, which is dynamic analysis, would become too high. Especially in the latter case the end of the design process is premature as not all analysis data that is inevitably needed for a sound decision pro or contra a variant of a future terminal's design will be present. This may yield a wrong or at least sub-optimal solution. Another drawback of applying only the first step of the design process is the lack of stimulus to investigate new technologies: Performing calculations on terminals that contain new techniques for transshipment or new rules for operations are even more arduous than calculations on terminals that contain conventional technology. In this way further investigation and development is impeded, which impedance could be eliminated especially for new operational rules by offering a flexible design tool like the ones shown in previous section. A clear advantage of the first step over the second step is the great insight that is obtained in the operations at the rail terminals that is designed; although not all effects can be included in this static analysis, it indisputably does increase insight of the terminal designer in the processes he actually is modelling. This advantage of the first phase will hardly ever be beaten by any computer aided terminal design tool.

From this reasoning the conclusion can be drawn that entering the second step, i.e. performing dynamic analysis, in some cases is too arduous and therefore too expensive. This drawback is mended by the design tools discussed shortly in the previous section which offer some standard frameworks to quickly model and analyse terminals of one or various types. Therefore these design tools are an important step forwards.

However, it also has been shown that all of the design tools have some drawbacks. Most of these drawbacks stem from the limited number of terminal types that can be modelled and analysed, or from the lack of handles to control flows within the modelled terminal sufficiently.

It is mainly this latter problem that is tackled in this thesis: It is aimed to offer a standard structure that is powerful enough the control flows through the modelled terminal in a detailed way, in the meantime offering the facility to analyse the same model at different levels of detail. Especially when considering a complicated terminal system as the future Maasvlakte in which various terminals of different types are related to each other through the ITT-system it is important to have a strict base structure that allows to control all individual flows through the entire terminal system. In the meantime the user of the model should be given automatically some advice on specific parameters like the number of machines to install. Following chapter gives a quick overview of the model.
Chapter 5. Proposal for a computer aided terminal design method

5.1. Introduction

"March 21 1998. Peter, presently we have only been able to analyse the first five of the twelve options for lay-out and for train operations as described in working paper MSV-102-A. This disappointing result is caused by a major breakdown of the main computer network. By consequence we are 1½ day back on our scheme and request you to increase our budget from 6 to 7½ days in order to complete full analysis of all twelve options. I am very sorry for this serious setback on our scheme and we are doing all that is in our power to speed things up.

Yours, Dave."

This could be a typical progress report composed by an operational engineer that is charged with the analysis of some future rail terminal; the engineer is to analyse in total twelve different options for lay-out and for train operations of the rail terminal in a six-day period only. Using the current methods for terminal design as described in Chapter 4 it is unthinkable that such a performance is demanded from anyone, as each different option for lay-out or for train operations requires either extensive calculations by hand or (partial) re-programming of some simulation model. Using these current methods for terminal design virtually impedes analysis of twelve different options in a period of six days only.

Another consequence of using the currently available methods for terminal design is that costs involved with each variant analysis are great: Both extensive calculations by hand and creating or updating a simulation model require much time and therefore induce high costs. By result only a limited number of variants for any terminal's design is analysed in detail in practice. These variants that are selected for detailed analysis are those variants that apparently are the most promising ones; which in most cases are the variants that come closest to the situation that is found at current, known terminals. By consequence, there is no stimulus to explore totally different operational techniques or lay-out for new terminals, as analysis costs would become too high.

To allow a more flexible and quick analysis of various options for a terminal's design some standard analysis tools should be available which permit to specify a terminal and all operations at that terminal by a number of standard parameters. In that way it becomes relatively simple to enter and analyse different options for lay-out or for train operations -- resulting in a progress report as shown in the beginning of this section.

This chapter gives a brief overview of a method to develop and analyse new rail terminals almost as quickly as described in the above progress report. This method is the main topic of this thesis and results in a generic model. After the model's structure and elements have been discussed briefly in this chapter they are elaborated in more detail at some other, separate chapters. But first of all some essential starting-points for the model must be defined. This will be done in the next section.

5.2. Starting-points

The character of the model to improve the design and analysis process of rail terminals is determined in a substantial way by the starting-points that are used for the model. These starting-points do not only include the type of terminals to model and the method to analyse them, but also aspects like the aimed users of the model and its point of view. All these starting-points are discussed sequentially.
Chapter 5. Proposal for a computer aided terminal design method

♦ Scope - Type of terminals to model
The model described in this thesis is developed first of all to speed up design and analysis of rail terminals, or more specific, for rail terminals that are served by so-called shuttle trains as described in Chapters 3 and 9. A consequence of this choice is the focus on railway aspects and detailed modelling of railway operations. As the model was developed first of all with the aim to use it for the future Maasvlakte situation in which the future rail service centre is to co-operate with a number of marine terminals and stacks, these latter terminals and stacks can be modelled and analysed too, even though at a lower level of detail. Any combination of any number of co-operating rail terminals and marine- or barge terminals can be modelled. In this thesis such a combination is called a terminal complex. Another consequence of the aimed primary use for the future Maasvlakte situation is the inclusion of complex structures for internal transport and of flexible definition of equipment types and assignment structures. These latter aspects will be discussed in detail in Chapter 7.

The elements that are found in most terminal complexes and therefore are included in the model are:
- multiple rail terminals;
- multiple marine terminals;
- multiple truck terminals;
- multiple stacks;
- multiple equipment exchange locations;
- complex internal transport.

Next to these physical elements some other, operational elements are included in the model like for instance high-level flow management for the entire terminal complex (refer to Chapters 6 and 7), equipment management (refer to Chapter 7) and train management (refer to Chapter 9). All of these aspects will be pinpointed briefly in this chapter and are discussed in more detail in the referred ones.

♦ Scope - Load units
Another aspect of the scope of the model is that it is restricted to terminals that are used for so-called standard load units only, like ISO containers and swap bodies. Refer to Chapter 2 for more details on this subject. These standard load units can be defined freely in the model based on following characteristics:
- design class, for distinction between different types of standard load units like containers and swap bodies.
- special class, for distinction between different special characteristics of load units like empties and reefers.
- height class, for distinction between different heights of load units like 8' and 9'.
- width class, for distinction between different widths of load units like 8' and 8'6'.
- length class, for distinction between different lengths of load units like 20' and 30'.
- weight class, for distinction between different weights of load units.

Any combination of these classes models a different type of load unit. In practice there is no need to model all types of load units that are handled at the actual terminal complex -- some simplifications are highly recommended. In a number of cases it is, for instance, sufficient to distinguish between 20' and 40' containers only.

♦ Scope - Clients
Strongly related to load units are the clients of these load units, which are extremely loosely defined in this model. If required, a number of different clients of the load units have to be defined, adding a "label of property" to each individual load unit. This aspect of property is only used in Chapter 8 when discussing block selection for a particular load unit in a stack by means of multi-criteria analysis; if this block selection on clients is not used, then there will be no need to define different clients of load units to make to model work properly.

♦ Point of view
A number of points of view could be selected for a computer aided terminal design method: Focus could be put for instance on civil engineering aspects, on mechanical engineering aspects or on logistic aspects. This latter point of view, i.e. on logistic aspects, has been chosen in this thesis as it is considered as one of the most elementary and hard-to-grasp aspects of terminals. Some example of these logistic aspects are train operations,

---

1This selection is mainly a consequence of the historical background of the project that initiated the here described work.
2In case of the future Maasvlakte system an example hereof can be found in the DistriPark.
3In case of the future Maasvlakte system an example hereof can be found in the Inter Terminal Transport system (ITT).
4The aspect of weight of load units is not used further explicitly in the model, but if needed it can be used to discern between two different types of load units that need to be handled differently.
Chapter 5. Proposal for a computer aided terminal design method

locomotive operations, transhipment and internal transport. Along with these logistic aspects some low-level civil engineering and mechanical engineering aspects are included like stack surfaces and equipment type definition. However, these aspects are only included where needed to obtain a valid logistic model.

Because of the chosen point of view a terminal complex will be modelled in a more abstract way than suggested by the fancy maps shown in Chapter 3. The future Maasvlakte system, for instance, will be reduced to model as shown in figure 5.1 in which only the essential location are shown, as "bulbs" that are connected by some lines. This is a consequence of modelling real-life systems and is almost inevitably. Therefore, a notation as shown in figure 5.1 will be adapted in this thesis.

![Figure 5.1: Abstract representation of a real-life terminal complex.](image)

An important aspect that can be seen from figure 5.1 is the system's boundary. This boundary plays a major role in the model described in this thesis and is defined as the imaginary boundary that runs over all locations (like terminals) over which load units enter and leave the modelled system. In other terms, all locations at which load units enter or leave the system are part of the system's boundary. By consequence only locations at which load units can be transhipped onto and from external means of transport (like trains, vessels or road trucks) can be located at the system's boundary; locations like stacks can not be located at the system's boundary.

- **Objective users**
  From the above description of the model's scope and point of view it can be concluded that the model covers a wide range of aspects that can be found at a terminal complex. This automatically yields the conclusion that the model should be applied by a great variety of persons, each one having expertise on part of this full range of aspects and working on a different level of abstraction. Therefore, the model should be usable and accessible for both conceptual and pragmatic experts on any aspect that is incorporated in the model. This wide variety of objective users has some consequences for the final implementation of the model: It should be understandable, logical and user-friendly.

- **Modelling method**
  Based on the aimed high level of accessibility, it has been chosen for to implement the model as a highly user-friendly and visual oriented computer programme under Windows™. This computer programme, which is called TORCMOD5, allows simple modelling of any system by means of pop-up menus and by means of double-clicking and dragging system elements. At any time a modelled terminal complex is guaranteed to be consistent, both through extensive checks and through the model's structure, as will be described briefly in this chapter. At all times a modelled terminal complex is ready to be analysed instantly.

- **Analysis method**
  The method selected to analyse a model of a terminal complex once it has been specified in the computer programme is simulation, both continuous and discrete simulation. Refer to Appendix B for some more details on this subject. Another approach could have been to analyse the modelled system by means of queuing theory or by other static calculations. However, these methods have not been chosen for as they provide less accurate and detailed information than simulation can (Brunner [2])6. Simulation of the modelled terminal complex

---

5 Abbreviation of: Transhipment Of Rail Containers Modelling tool.
6 Queuing theory is mainly suitable for those situations in which no detailed information on time-tables of trains is available, or in those cases where a time-table is reduced to a random distribution (refer to Wakob [33], p. 93) -- with loss of information on specific sequences of trains. This is a consequence of the preliminary assumption of queuing theory that "customers" (i.e. trains) arrive at an inter arrival time that is generated by a random (generally Poisson) process. Another drawback of queuing theory is the inability to model interrelation (continued on next page)
results in a number of output parameters that will be discussed at the end of this chapter, which can be used to determine the performance and costs of the modelled terminal complex.

♦ Analysis objective
Depending on the stage of the design process a different need for analysis results will exist: At the very beginning of the design process there is a need to get some rough ideas on the timing and size of flows through the modelled system only. At a later stage there is a need to perform quick experiments with a number of alternative scenarios and to get advice on aspects like, for instance, the number of straddle carriers to install at a certain location. At the final stage of the design process there is a need to fine-tune the modelled system and to determine its sensitivity for some highly detailed aspects like a block safety system for train operations.

This subdivision of the design process into three stages is incorporated in the model and will be discussed in the next section. In this way the user of the model is given an opportunity to gradually explore and expand the design of the modelled terminal complex.

♦ Definitions
In the remainder of this chapter following definitions will be used:

- system = real-life system that is to be modelled and analysed
- model = description of the real-life system (general)
- = description of the real-life system (at a specific level of detail)
- user = person converting the system into a model.

5.3. Subdivision into three levels of detail

When discussing the analysis objective of the model at the previous section it was shown in brief that in some cases there is a need to gradually expand the level of detail of the model. In this way there is an opportunity for the user of the model to perform a more detailed analysis when the terminal complex is in an early stage of design, and to perform more detailed analysis when it is in a further stage of design. Even more, fundamental changes in the model are more easily made and analysed when applying the analysis at a low level of detail than at a high level of detail, as the number of details that is interrelated and needs to be changed too is limited.

Based on this thinking it has been chosen in this thesis to divide the process of analysing terminal complexes into three stages or levels of detail. Each of these levels of detail has its own aim and therefore its own model elements from which a terminal model can be constructed, and its own aimed results after simulation. First of all these three levels of detail are described briefly, along with their coherence and aim. After that in a separate section per level of detail some more details will be discussed, like the model elements that are available at each level.

Note that before either one of these stages of the computer aided design process is entered, extensive data analysis should have been performed on aspects like future flows of load units to and from the terminal complex to analyse, costs of equipment, and scenarios for internal transport structures. Gathering these data requires a lot of time and can not be avoided by using a computer aided design method as normally these data are based on some vision of the board of the future rail terminal on strategies to follow and aims to reach, inspired by macro-economic studies.

♦ First level of detail: UCM
At the first level of detail only the basic structure of the modelled terminal complex is specified, along with data on flows of load units entering and leaving the system as will be discussed in detail at Chapter 6. This basic structure covers the presence and position of locations like stacks, rail service centres, marine terminals and

1Note that this subdivision is rather subjective. In an earlier stage of the project two other levels of detail than described here were considered, too. Both levels contained continuous models like UCM does, but added some more detailed aspects like pooling of equipment in a continuous way. Because of their limited added value these two levels were removed from the terminal design method.

2Abbreviation of: Unrestricted Continuous Model.
truck gates, and of connections for internal transport between these locations. In this way the first level of detail provides a "backbone structure" for the remaining two levels of detail.

The name of this first level of detail is UCM. It uses continuous simulation (refer to Appendix B) to give an impression of the size and timing of flows of load units through the modelled terminal system. Fundamental changes in the modelled terminal complex like adding or removing internal stacks is to occur at this level of detail.

**Second level of detail: SDM**
At the second level of detail a great number of aspects is added to the first level of detail; aspects that have been omitted purposely at UCM level of detail. Examples of these new aspects are full train operations, full equipment definition and full flow management controls. These latter controls are defined by "vague" preferences for particular stacks or for particular equipment pools as will be discussed in detail in Chapters 7 and 8.

The name of this second level of detail is SDM. It uses discrete simulation (refer to Appendix B) to give a rather detailed insight in the train operations, in the use of equipment and in the planning of flows of load units based on a number of preference settings. Even more, this second level of detail advises the user on the number of tracks to use at various important sites in the rail infrastructure, and on the number of equipment to install at various pools. Mid-level changes like changing operations in train operations or in equipment to use is to occur at this level of detail.

This second level of detail will be used more intensively for variant analyses than the other two, as in most cases it provides sufficiently detailed analysis results against a relatively small amount of effort.

![Diagram](image)

**Figure 5.2:** *Interrelation of all three levels of detail.*

**Third level of detail: DDM**
At the third level of detail the last details are added to the design of the modelled terminal complex: Block safety systems for trains and more detailed equipment routing. In most cases there will be no need to perform

---

9 Abbreviation of: Simplified Discrete Model.

10 Except block safety systems.

11 Abbreviation of: Detailed Discrete Model.
too many experiments with this level of detail as it requires relatively much effort to add the last details on these block safety systems.

The name of the third level of detail is DDM. Analogous to SDM level of detail it uses discrete simulation to give an insight in the effects of these last two details on the performance and costs of the terminal complex.

The subdivision into three levels of detail and their interrelation is given in figure 5.2. It shows the rapid expansion of the number of covered details when changing from UCM to SDM, in contrast to the more modest expansion when changing from SDM to DDM. This figure also shows that the user is free to change between all three levels of detail whenever he likes.

- **Design cycle**
  A typical usage of these three levels of detail is shown in figure 5.3. It shows the path between the initial step in the design process (at which data is gathered on future scenarios and different options to analyse) and the final step in the design process (at which all future scenarios and different options have been analysed). This final step can be reached from any level of detail, although it is most likely that it is reached either from SDM or from DDM level of detail. Figure 5.3 also shows some feedback loops between the different level of detail (each one is used for different types of modifications of the model) and some loops within each level of detail.

  These latter, small, loops require some more discussion, as these are most frequently used. When expanding these small loops figure 5.4 is obtained, in which a repeating loop of editing, simulation and analysing is shown. This loop is found at all three levels of detail. As they are used so frequently the implementation of the computer aided terminal design method would (and actually does) offer the possibility to go through them quickly and easily, thus stimulating research of new options and lay-outs.

- **Overview of elements per level of detail**
  Each level of detail has a different aim and therefore needs different elements to reach these aims. But, as shown in figure 5.2, each higher level of detail is based on the elements that are defined at a lower level of detail, thus creating a "chain" of different elements at different levels of detail that are strongly interrelated. The

---

12 That is, depending on the problem that is analysed any of these levels could provide sufficiently detailed analysis results.

13 In the meantime trusting that the user of the model is aware of the danger of over-experimenting in an ad hoc and impulsive way.
next three sections give a brief overview of all elements that are available at each level of detail, thus giving a quick impression of the abilities of the described terminal design method. Each of these elements requires more discussion than can be given in these next three sections, and therefore are elaborated in more detail in separate chapters that will be referred to.

After the elements have been discussed, it will be shown what type of analysis results are aimed at for each level of detail. Obviously, these aimed results will be strongly related to the available elements at those levels of detail.

### 5.3.1. Overview of elements at UCM level of detail

The lowest level of detail, UCM, is used to define the backbone structure of the entire model of the terminal complex that must be analysed. As stated earlier, before the model at this level of detail can be entered for analysis extensive data analysis should have been performed on aspects like future flows of load units to and from the terminal complex, on the types of load units that are expected to be relevant for the analysis, on the clients that are to be distinguished and on the system's boundary. In other terms, it should be clear what terminals are to be included in the analysis.

- **Physical elements and Flow data**
  
  The aim of the UCM level of detail is to obtain a model of a real-life terminal complex as shown in figure 5.5, which is based on figure 5.1. The elements shown in this figure are called physical elements as they are physically visible, can be shown in a straightforward way on a map, and define the geometry of the modelled terminal complex. This in contrast to flow data (or flow elements) that are not that obviously displayed on a map.

  ![](image)

  **Figure 5.5:** Example of a terminal complex at UCM level.

  Based on these two types of elements a simple model of a future terminal complex can be created at a low level of detail, allowing instant and quick simulation and analysis by means of continuous simulation as described in Appendix B. These elements will be discussed briefly in this chapter; a more extensive discussion can be found in Chapters 6, 7 and 8. Refer to the Index, too.

- **Physical elements**
  
  Physical elements cover all elements needed to model stacks, exchange points, marine terminals, barge terminals, rail terminals, truck gates and transport links between all these locations. In this way they determine the geometry of the modelled terminal complex. The names of these physical elements, as defined in this thesis, are:

  - bounding elements; used to model marine terminals, barge terminals and truck gates.
  - train handle areas; used to model rail terminals.
  - internal stacks; used to model stacks in detail\(^{14}\).
  - exchange points; used to model direct exchange between equipment other than at the above types of locations.
  - additional points; used to model any other location, that is without any function but defining the internal transport system.
  - physical paths; used to model transport links between any of the above locations.

  All of these elements will be given some more attention.

\(^{14}\) These details will not show through than at SDM and DDM level of detail.
Chapter 5. Proposal for a computer aided terminal design method

The first two model elements, bounding elements and train handle areas, are the only two model elements that are located at the system's boundary; all other elements are not. By consequence, only these two elements can be used to let load units enter or leave the modelled system. In other terms, only at these two elements load units can cross the system's boundary. This is illustrated in figure 5.5 in which both bounding elements ("Miersk" and "Nin-Jan") and the only train handle area ("Central RSC") are shown as grey circles, each one being connected to a pair of slant arrows. These slant arrows represent the flow of load units across the system's boundary. At bounding elements and train handle areas aspects like train services and barge services can be modelled, as will be discussed later on in more detail at the flow data.

Internal stacks are used to model stacking facilities in a rather simple way at UCM level of detail: To each load unit type that is defined a fixed stacking time can be assigned. Any load unit that arrives at an internal stack during simulation at UCM level of detail is stored for the specified amount of time at that stack, specified for that load unit's type. In this way a quick impression of stack sizes is obtained. The drawback of this approach is that in some cases the fixed stacking time is not known, or is different for various "flow-lines" through the modelled system. Therefore, at SDM level of detail fixed stacking times are avoided and flows of load units are controlled and timed in a different, more complex way. Also, at SDM level of detail a great number of details is added to internal stacks, compared to UCM level of detail. Figure 5.5 shows two different internal stacks: "RSC Stack" and "Nin-Jan Stack".

Exchange points are used to model direct interference between equipment at locations other than internal stacks, bounding elements and train handle areas. At these exchange points load units are transshipped directly between two different pools of equipment. Because of the rather simple way in which internal transport is modelled at UCM level of detail (which is still to be discussed), exchange points do not have any special influence on the simulation results and behaviour; actually, they are more or less ignored at UCM level of detail and become of importance not until SDM level of detail, as at that level they introduce a so-called equipment break as will be discussed later on. No example of an exchange point is given in figure 5.5.

Additional points are used to attach the network for internal transport too, only; no special actions are to occur at this type of locations. In this way they are only used to define the geometry of the modelled terminal system. Four additional points are shown in figure 5.5 which can be recognised by their lack of names in that figure. It will be clear from figure 5.5 that, although no special actions are to occur at these locations, they are essential to model the shown internal transport network.

Physical paths are the connections between the various locations shown in figure 5.5. These connections are used for internal transport from one location to the other at UCM level of detail. They also play an important role in the definition of one of the types of flow data that is still to be discussed: Closed tasks. These closed tasks will turn out to be a kind of sequence of physical paths. At UCM level of detail internal transport between the various locations is included in a rather simple way: To each physical path a fixed transportation speed is assigned that is used for all load units that are transported over that particular physical path during simulation. An aspect as transshipment should be included in a rough way in this fixed transportation speed, too. At SDM level of detail, on the contrary, internal transport will be modelled in much more detail.

Flow data
Based on the above physical elements a number of flow data is defined. These flow data determine the flows of load units across the system's boundary, and through the system itself. Five different types of flow data are discerned, as will be discussed briefly:

- closed tasks; used to define the course of flows of load units through the system.
- load schedules; used to define trains, marine vessels and barges.
- load volumes; used to define numbers of load units at these trains, marine vessels and barges.
- regular contra schedules; used to define marine vessels and barges (in a less stringent way than load schedules, as will be discussed).
- intensity contra schedules; used to define truck gates.

These different types of flow data will be explained by referring to following example.
**Figure 5.6:** Example of a "flow line" through the modelled system.

Consider figure 5.6 which is a stretched-out form of a connection between train handle area "Central RSC" and bounding element "Nin-Jan", taken from figure 5.5. At the left-hand side of figure 5.6 one train service named "Train Antwerp 1" is shown, and at the right-hand side three different marine vessels, docking at marine terminal "Nin-Jan" are shown.

Between these two locations at the system's boundary a connection can be defined, running over all physical paths between these locations. Such a connection is called a closed task in this thesis, and plays an important role when defining flows of load units across the modelled system. This role becomes most evident at SDM level of detail.

A closed task is always defined in two directions (bi-directional): It defines, when referring to the above example, both a flow from "Central RSC" to "Nin-Jan" and a flow from "Nin-Jan" to "Central RSC". Also, any number of parallel closed tasks can be defined between the same pair of locations. This aspect especially will become of importance at SDM level of detail, as it allows to model several flows with different characteristics between the same pair of locations at the system's boundary. In this way complex structures of parallel flows through the modelled terminal complex can be defined and analysed.

When a closed task is shown in a map instead of in the stretched-out form of figure 5.6 then a situation as depicted in figure 5.7 is obtained. This is the normal way to show closed tasks. Figure 5.7 also shows that only two pairs of slant arrows are valid when referring to a single closed task; these are the slant arrows that represent the flows of load units that run over that particular closed task, and by consequence are defined at both locations at the system's boundary between which the closed task runs. From this description it can be deduced that flows of load units are assigned to particular closed tasks; each flow can be assigned to a different, fixed closed task. This aspect will be discussed in more detail later on.

Note that the combination of two properties, namely (1) that closed tasks run between two locations at the system’s boundary only, and (2) that flows of load units are assigned to closed tasks only, yields an important conclusion: It is always guaranteed that load units that once enter the system at some time will leave the system again. In this way it is always guaranteed that the modelled system is automatically in a state of equilibrium, in which the number of load units entering the system equals the number of load units leaving the system, when considering a longer period of time. Such a system is called an automatically balanced system, which is of great importance for long-term simulations. This aspect will be discussed in more detail in Chapter 6.
A more complex situation than shown in figure 5.6 is given in figure 5.8. In this situation four marine vessels dock at bounding element "Nin-Jan" instead of three. Furthermore, in this situation two different closed tasks between bounding element "Nin-Jan" and train handle area "Central RSC" are discerned: The lower one runs over the stack of the train handle area, and the upper one does not. Important to note is that different marine vessels are related to these two closed tasks. In this way it can be controlled which ships are to be handled in relation to both closed tasks. This is an aspect that is included in load schedules and contra schedules.

Load schedules are used to model regular services for transport of load units across the system's boundary like trains, marine vessels and barges. These load schedules specify the timing aspects of these regular services like the moment at which they arrive at the system's boundary and the moment at which they leave again. They also cover aspects like random disturbance of this arrival time and the repetition interval of the service. Note that from this definition it is clear that the model is designed for systems at which services like trains and vessels arrive at a regular repetition interval; it is not designed for systems at which, for instance, barges arrive according to some "call-structures" at which barges are called at the moment that a sufficient great amount of load units is present at the terminal system for transport by these barges. Also note that load schedules do not contain any information on the number or type of load units that is to be handled; they only contain information on timing. Refer to Chapter 6 for more details.

Each load schedule contains a number of load volumes. In contrast to load schedules, these load volumes do contain information on the number and type of load units that are transported across the system's boundary. They also contain information on the closed tasks over which the load units flow and on the clients to which the load units belong. For more details it is advised to read Chapter 6.

The last two aspects to discuss at UCM level of detail are the contra schedules: Regular contra schedules and intensity contra schedules. The first type is used to model marine vessels and barges, and the second one is used to model truck gates.

The function of these contra schedules can be seen from figure 5.8 in combination with the previously discussed automatic balancing character of the model: At any time it must be guaranteed that the number of load units entering the system equals the number of load units leaving the system, when considering a longer period of time. Referring to figure 5.8 this would imply that the number of load units that is transported by the single train at "Central RSC" equals the sum of the numbers of load units that are transported by all four marine vessels docking at "Nin-Jan". It will be clear that it is not easy to specify all these numbers of load units by hand, in the meantime guaranteeing that the modelled system is in balance. Even worse, when one or more of these numbers of load units to be transported are specified as random distributions rather than as fixed numbers, then it is even impossible to specify these data by hand. To handle this problem the concept of contra schedules is introduced.

Contra schedules specify the timing of regular services as marine vessels and barges, or of mass transport as road trucks; they do not specify the numbers of load units that are transported. This in contrast to load schedules and their strongly related load volumes that specify both timing of regular services and the numbers of load units that are transported. In this way contra schedules become the more abstract counterpart of load schedules.

---

15 Especially for more complex terminal systems, covering more bounding elements and train handle areas, and more train services and marine vessels.

16 Which in fact is the chosen approach in the implementation of the here described computer aided terminal design method.
As suggested in the previous paragraph load units only flow from a load schedule to a contra schedule or vice versa, as shown in figure 5.9. At this point there is no difference between regular contra schedules and intensity contra schedules. Figure 5.9 also rightly suggests that load schedules are to be related or matched to contra schedules. This subject will be discussed in detail in Chapter 6; for the time it is sufficient to know that both types of schedules are matched dynamically to each other during simulation at SDM and DDM level of detail; they are not used at UCM level of detail which is a consequence of the use of fixed stacking times and fixed internal transport speeds at that level of detail, instead of so-called equipment routings and flow patterns as will be discussed at SDM level of detail.

As regular contra schedules are used to model the same type of regular services as load schedules do, this type of contra schedule is given no further discussion at this point. Intensity contra schedules on the other hand are used to model totally different types of transport: They are used to model mass transport like road trucks that arrive at a truck gate at the boundary of the modelled terminal complex according to some intensity pattern rather than according to some regular service. This will be illustrated by following example.

![Relative intensity vs Time period number](example_graph)

**Figure 5.10:** Example of an intensity contra schedule.

Consider figure 5.10 in which an intensity contra schedule is given, modelling the intensity of arrival of road trucks at a truck gate. This truck gate can only be modelled by means of a bounding element\(^{17}\) at which an intensity contra schedule is defined, as load units cross the system's boundary there. The horizontal time axis of the intensity contra schedule shown in figure 5.10 is divided into 24 separate periods, each one exactly one hour long. For each of these 24 one-hour periods it is shown how many trucks at average arrive at the truck gate, for two different directions: Import and export. Import refers to load units that are delivered by trucks and therefore enter the modelled system. Export, on the other hand, refers to load units that are picked up by trucks and therefore leave the modelled system.

The basic assumption in this definition is that each truck either delivers or picks up one load unit; and one load unit only. So, the model does not include detailed aspects like trucks coming to pick up two load units at a time that must be combined and loaded onto the same truck, nor aspects like trucks coming to deliver a load unit first and picking up another one afterwards\(^{18}\). By consequence, a more precise definition of intensity contra schedules would be a pattern describing the chance for system entry or -exit of load units for each one-hour period, rather than of road trucks.

\(^{17}\)To each bounding element only one type of contra schedules can be assigned, either regular or intensity contra schedules.

\(^{18}\)This simplification is a consequence of the focus on rail terminals in a large-scale marine terminal environment as found at the Maasvlakte.
Another important aspect of intensity contra schedules is that they may be different for all seven weekdays in two ways: They may be different in (1) the intensity pattern over all 24 one-hour periods, both for import and export, and (2) the total number of load units that cross the system's boundary during the entire weekday, irrespective of the patterns for import and export. These aspects are needed for a sufficiently detailed modelling of truck gates, and are discussed in more detail in Chapter 6.

### 5.3.2. Overview of elements at SDM level of detail

The mid level of detail, SDM, introduces a wide variety of new elements that are based on the structure provided by the elements at UCM level of detail. This variety of new elements adds many details that cover following three groups of aspects:

- rail operations;
- internal transport;
- stacking.

Although the length of this list is rather moderate, the number of elements and aspects that is covered by these three groups is substantial. All three groups will be discussed briefly in this section; more details can be found in Chapters 7, 8 and 9.

#### Rail operations

The first group of elements to discuss at SDM level of detail are the rail operations. As the model's primary concern is to facilitate design and analysis of rail terminals this is considered as one of the most important groups. The elements covered by this group are:

- rail infrastructure;
- train services;
- train concepts which define the behaviour of train services;
- light D-engines;
- light E-engines.

It would require too much discussion to describe all aspects covered by this group in this section in detail. However, some of them are described here in little detail to give a quick impression of the level of detail and of the coherence of the various elements.

The basic elements for any rail terminal are the rail infrastructure elements. At SDM level of detail three different types of infrastructure elements are discerned, analogous to reality: Rail tracks, rail bundles and rail points.

An example hereof is given in figure 5.11 showing one rail bundle that is connected to a number of track by some points. Various types of points, as found in the actual situation, are included in the model. In this way the routing of trains over the rail infrastructure can be guaranteed to be in correspondence with the real-life situation.

![Figure 5.11: Example of rail infrastructure.](rail_track_rail_bundle_rail_point)

All trains in the model behave according to a number of standard settings. These settings, which are called train concepts in this thesis, define operations of trains with respect to:

- locomotives, like coupling and uncoupling light electrical engines (E-locos) that are used at the public rail network and diesel-powered locomotives (D-locos) that are used at the rail terminal's local yard;
- system entry- and exit, like stationary brake tests and direct system entry for direct unloading of trains after they have arrived according to their timetable;
- handling, like the sequence of load units on-train and repositioning of the train at another location of the rail terminal's yard between unloading and loading;
- bundle selection, like finding the best bundle to unload the train at.

By altering one parameter that is part of such a train concept, the behaviour of all trains that refer to that particular train concept changes instantly. In this way quick variant analysis is allowed and manageable.

In this model trains are strongly related to the previously discussed load schedules: To each train service a number of load schedules can be assigned that determines the numbers of load units that are to be unloaded or loaded by these trains. In this way a sound coupling between UCM and SDM level of detail for train handle areas is guaranteed.
During simulation light engines can be used for local movements of trains within the range of the modelled rail terminal's yard. In some cases there is a need to find out quickly the number of locomotives to install at that rail terminal in order to achieve a certain service level for the trains that want to move, by guaranteeing a maximum waiting time for a train before a light engine is assigned to the train that wants to move. In order to facilitate this kind of analysis there is an option to specify a maximum waiting time for these trains, resulting in an advice on the number of locomotives to install at the rail terminal -- along with a detailed cost analysis.

**Internal transport**

Flows of load units through the modelled terminal complex are controlled in a totally different way at SDM level of detail than at UCM level of detail: At the latter level all stacking times and internal transport speeds were specified as constant, fixed values. By consequence, for all load units entering the system at a given time it was known on beforehand at what time they would leave again: The time at which they entered the system plus the stacking times of all stacks they are to run over, according to their closed task and their load unit type, plus the internal transport times of all physical paths they are to be transported over, according to their closed task and the transportation speeds and distances of these physical paths. At SDM level of detail things run different.

![Position at terminal complex](image)

**Figure 5.12:** Same "flow speed" for any two load units of the same type at UCM level of detail.

At SDM level of detail load units are registered both at load schedules and at contra schedules: For each individual load unit it is first determined at which time it is to enter or leave the system according to its load schedule, and then its matching contra schedule is searched at the other side of the load unit's closed task. This principle of matching load schedules to contra schedules has been shown previously in figure 5.9. In order to be able to find a matching contra schedule for a load unit some additional data must be known: The time it is to reside in the modelled. This time of presence is specified at SDM level of detail through the so-called equipment routings. Each of these equipment routings is related to one closed task, in one direction only. This aspect will be discussed in detail at Chapters 6 and 7, but for the moment it is important to know that at SDM level of detail for each individual load unit it is known at what time it enters and leaves the system based on a random distribution that specifies the duration of presence in the system, rather than based on fixed stacking times and fixed transportation speeds. The consequence of this different approach for both levels of detail is an increased freedom in planning of flows of load units through the system: At UCM level of detail all timing flows is fixed by fixed stacking times and fixed internal transport speeds, while at SDM level of detail all timing of flows is totally free and is controlled by some "vague" preferences that will be discussed shortly only.
Chapter 5. Proposal for a computer aided terminal design method

**Figure 5.13:** Different "flow speed" for two load units of the same type at SDM level of detail.

This aspect is illustrated in figures 5.12 and 5.13, showing the distribution of the "flow speed" for two different load units through the modelled system over the same closed task, for both levels of detail: UCM and SDM. From figure 5.12 it can be seen that at UCM level of detail any pair of load units that flow over the same closed task remain equally long in the modelled system, as they are stacked equally long at all stacks over which the closed task runs (provided that both load units are of the same type) and are transported equally fast over all physical paths that are included in that closed task.

From figure 5.13 it can be seen that at SDM level of detail any pair of load units that flow over the same closed task is likely to remain for a different amount of time in the system, as this amount of time is specified by random distributions that are related to equipment routings (and therefore to closed tasks). Based on this amount of time each individual load unit is to remain in the system the stacking times and the moments for internal transport are determined dynamically by the model during simulation. As mentioned before, this dynamic process of finding the best moments for stacking and for internal transport, given the times at which each single load unit is to enter and the leave the system is controlled by some "vague" preferences that will be discussed shortly. This difference in timing of flows of load units through the modelled system is shown in figures 5.14 and 5.15, too.

**Figure 5.14:** Controlling timing of flows of load units at UCM level of detail.

**Figure 5.15:** Controlling timing of flows of load units at SDM level of detail.

One of these "vague" preferences used to control the flow of load units through the modelled system at SDM level of detail are the so-called flow patterns. These flow patterns are, analogous to equipment routings, strongly related to closed tasks. They specify for each closed task, for both directions, the preference for the presence of load (units running over that closed task) to reside at each location covered by that closed task.
Consider for example the situation shown in figures 5.6 and 5.7 in which a closed task is defined between "Central RSC" and "Nin-Jan". A flow pattern would specify the preference for the presence of those load units at the various locations over which that closed task runs, for both directions over that closed task. Refer to table 5.1 for some example preferences of a flow pattern.

<table>
<thead>
<tr>
<th>Location</th>
<th>Preference for presence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central RSC</td>
<td>Direction &quot;up&quot;</td>
</tr>
<tr>
<td></td>
<td>Direction &quot;down&quot;</td>
</tr>
<tr>
<td>Nin-Jan Stack</td>
<td>++</td>
</tr>
<tr>
<td>Nin-Jan</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 5.1: Example of preferences according to a flow pattern.

Internal transport at SDM level of detail uses individual machines or equipment, this in contrast to internal transport at UCM level of detail that does not include the idea of separate machines at all. These machines at SDM level of detail are defined by their respective equipment type which can be compared to the previously train concepts for trains. These equipment types allow for a more flexible definition of equipment in correspondence with most real-life situations. In Chapter 7 all options for equipment pools and types are discussed in detail.

Through equipment networks that are spanned between the UCM-locations it is defined for which transport jobs equipment can be used. In this case too, closed tasks play an important role: Only equipment that is installed at a network connecting (part of) the locations over which a closed task runs can be used for transport of load units over that closed task. Giving more detail at this point would evince too much discussion, so it is referred to Chapter 7 again for more information.

Stacking

Stacking at SDM level of detail is, as just discussed, not subject to fixed stacking times. Instead, stacking times of load units at SDM level of detail are determined dynamically during simulation, based on the previously discussed "vague" preferences. A consequence of this approach is that a better insight of the actual demand for stacking capacity is obtained than at UCM level of detail.

All stacks at SDM level of detail can be subdivided into a number of separate stack blocks. Each of these stack blocks may have a different preference for use, based on various criteria. It may, for instance, be needed to distinguish between load units that are owned by a particular client, or between load units that are of some special type, or between load units that are destined for different marine terminals. All this kind of selection criteria may be assigned to a stack block, thus increasing the possibilities for system analysis.

5.3.3. Overview of elements at DDM level of detail

The final level of detail, DDM, adds a limited number of details to the model on two aspects only:
- rail operations;
- internal transport.

These final details are used to fine-tune the modelled system, partially based on analysis results obtained at SDM level of detail. This section gives a short overview of these final details.

Rail operations

The most important difference between SDM and DDM level of detail is the introduction of the so-called block safety system for trains at the latter level of detail. This block safety system has a significant impact on the planning of train movements across the rail terminal's yard, especially for those rail terminals serving a great number of trains. A short explanation of these block safety systems will be given shortly.

Another difference between rail operations at SDM and at DDM level of detail is the way rail bundles are defined: At SDM level of detail they consist of an unknown number of tracks, but at DDM level of detail they consist of a fixed, known set of tracks. This different approach of modelling rail bundles for both levels of detail has the consequence that:
1. experimenting with different lay-outs of rail infrastructure and with different scenarios for bundle grouping requires less effort at SDM level of detail than at DDM level of detail;

19These elements have not been implemented in TORCMOD yet.
Chapter 5. Proposal for a computer aided terminal design method

2. finding out the number of tracks to install at a rail bundle can be automated at SDM level of detail, and;
3. defining a block safety system, which must be defined in a detailed infrastructure environment in which each single track is discerned individually, is possible at DDM level of detail only as this level of detail discerns individual tracks within each rail bundle while SDM does not.

A block safety system, as described in more detail in Chapter 9, is commonly used at (parts of) railway yards to increase safety of train movements. Such a block safety system is based on the subdivision of the rail infrastructure into a number of blocks; access to these blocks is controlled by signals next to the tracks: a red signal indicates that an engine is not allowed to enter that track.

A consequence of the use of block safety systems is the reduction of the rail infrastructure capacity. As will be explained in Chapter 9 block safety systems cause a substantial amount of "dead time" in which rail is claimed or held for some particular train movement, and can not be used for some other movement. Infrastructure may even be claimed a considerable time for some train movement, ahead of the actual moment of arrival of that particular train at the claimed tracks. This amount of "dead time" depends on (1) the subdivision of the rail infrastructure into blocks (a better subdivision may substantially reduce the amount of "dead time") and on (2) the strategy for "claiming" the infrastructure: In total four different strategies for "claiming" are discerned in this thesis, which are:

1. ad hoc, stepwise claiming;
2. ad hoc, integral claiming;
3. planned, stepwise claiming;
4. planned, integral claiming,

each one having its own characteristics and field of application. Refer to Chapter 9 for more details on this subject.

● Internal transport

The difference in character between DDM and SDM level of detail, i.e. fine-tuning the modelled terminal complex, shows clearly through when considering internal transport: Two aspects are included additionally at DDM level of detail, compared to SDM level of detail:

1. inclusion of the actual paths for internal transport, according to the equipment networks modelled at SDM level of detail;
2. inclusion of alternative routing of equipment over these actual paths for internal transport, based on the current occupancy of these paths by other equipment.

These two aspects are strongly interrelated and will be given a little discussion. A main similarity between these two aspects is that they require no additional modelling at DDM level of detail, compared to SDM level of detail. Even more, the effects of both aspects of internal transport that are included at DDM level of detail can not be seen than through simulation and analysis only; they can not be shown through any new, visible element.

At SDM level of detail the distances for internal transport were determined in a rather simple way: They were determined either as the way over the physical path defined at UCM level of detail connecting the origin and destination of the equipment movement, or, if no such physical path is available, as the rectangular distance between this origin and destination. Especially for equipment networks within stacks and rail service centres this approach may introduce some noticeable simplifications of the model. At DDM level of detail these distances are determined in a more accurate way as a route over the equipment networks rather than as the rectangular distance between origin and destination. Such a route runs over a number of network links which can be compared to physical paths at UCM level of detail, having a more local, short-distance character.

---

20 Currently several new, more advanced safety control systems are developed. These new systems are not considered in this thesis as they are intended to be used for "high"-speed public networks only; they are not intended for use at yards of rail terminals.
21 In some cases, mainly on public rail networks, an additional system in the engines themselves automatically activate the brakes when the drives does not respond to the signal. The Dutch version of this system is called ATB.
22 Refer to Chapter 7 for more details on equipment networks and the way they are defined.
Using these routes over network links for internal transport at DDM level of detail also allows to take into account the aspect of alternative routing of equipment over these network links, which can be described as queue-avoiding behaviour of equipment. This is illustrated in figure 5.16 showing a truck that originally planned to travel from A to D over B, but "changes his mind" and selects a new route from A to D over C as another truck is blocking the route over B. Driving over B would reduce the speed of transport for that particular truck as the primary assumption is that the speed of transportation over a network link depends on the current density of trucks on that network link. This selection is to occur in an ad hoc way, so without any kind of planning. In order to be in correspondence with a wide variety of real-life situations it is freely defined (1) whether equipment may use re-routing, and (2) how far ahead equipment can look for a new route: In case of equipment that is driven and controlled by humans the options for re-routing will be limited as in such case the information on the occupancy of network links is available only at short notice. In case of equipment that is controlled by some central process unit for all equipment movements then routing of equipment can be determined far ahead, resulting in better performance of the terminal system.

Figure 5.16: Re-routing of equipment over location C.

A consequence of these additional details at DDM level of detail is a considerably reduced simulation speed, because of the much more complex routing processes of equipment. In a number of cases it may be decided that this extra effort does not pay off against the increase of accuracy of the simulation process.

5.4. Aimed results

Based on the above briefly described model elements most terminal complexes can be modelled. However, such a terminal complex is not just modelled "art pour l'art" but instead to perform some analyses on the modelled terminal complex. The type of data that can be used for such analyses depends on the model elements that are used and therefore on the level of detail that is considered. By consequence, the aimed results of the computer aided terminal design method are different for all three levels of detail. The next three sections give an overview of these aimed results for all three levels of detail, which for all three levels are in correspondence with the aimed level of detail to look upon the modelled terminal complex.

5.4.1. Aimed results at UCM level of detail

The aimed results for analysis at UCM level of detail are quite modest compared to those at SDM level of detail. This is a consequence of (1) the simple way in which stacking and internal transport are modelled, namely by means of fixed stacking times and fixed internal transport speeds, and of (2) the used approach to simulate the modelled terminal complex at UCM level of detail: Continuous simulation. These two reasons will be given some more attention after which the actual aimed results are given.

The simple way in which stacking and internal transport are modelled at UCM level of detail automatically yields the quite straightforward way of controlling flows of load units as discussed at figures 5.12 and 5.14, in which the course of each "individual" load unit through the modelled system is known once it crosses the system's boundary. This approach does not include individual machines that are used for internal transport, nor working hours of machines or any high-level planning of flows of load units. By consequence, analysis at UCM level of detail is restricted to obtaining an impression of the size of flows across the system's boundary and through the entire system.

Continuous simulation yields the consequence that at UCM level of detail individual load units are not discerned and therefore can not be manipulated nor traced individually. This limits the possibilities for detailed analysis results, as these results can only be related to entire flows; any insight in the behaviour of individual

23 These elements have not been implemented in TORCMOD yet.
load units lacks. That is one of the reasons why planning of flows of load units has been kept simple at UCM level of detail: A more detailed flow control in combination with continuous simulation, and therefore with continuous flows of load units could be considered as waste of effort as there would be no way to obtain sufficiently detailed information on the behaviour of this improved flow control. Even more, at SDM level of detail more advanced flow control mechanisms are available, combined with more detailed options for analysis.

The type of analysis data that is available at UCM level of detail therefore consists of one group only, namely:
• results on the size of flows of load units in time.
Although a modest list, it may result in some interesting information that can be used for two purposes: (1) Quick checking of the validity of the entered model and (2) rough insight in the timing and interrelation of flows. Both purposes will be explained.

**Quick validity checking**
Quick checking of the validity of the entered model is simply performed by providing summary information on the numbers of load units that have been transported to and from each location at the system's boundary, taken over a standard period of a week, month or year. These numbers of load units that have been transported can be split into flows per closed task, if required.

An example hereof is given in table 5.2 which is based on a situation as depicted in figure 5.5. This table shows the total of all flows that have entered and left the modelled system at all three locations at the system's boundary during a week, along with a more detailed subdivision of the import-flows into the locations to which the load units have been transported. Although table 5.2 does not discern between various closed tasks over which the load units flowed, this option could be used to obtain more detailed insight in the differences between these closed tasks.

The advantage of this quick validity checking is twofold. First, it is used to check whether all data on train services and vessels, specified as load schedules, has been entered correctly. Second, it is used to get a quick insight in the sensitivity of the size of flows through the system for various simulation runs, using different random seeds.

**Rough insight in timing and interrelation of flows**
A rough insight in timing and interrelation of flows is used to find out quickly which flows through the modelled terminal system are strongly related and may cause problems in further analysis at SDM and DDM level of detail. These strongly related flows are those flows that tend to occur simultaneously causing a highly peaked need for transport capacity. This will be illustrated.

![Graph showing strongly and less correlated flows](image)

**Table 5.2:** Example of summary information at UCM level of detail.

<table>
<thead>
<tr>
<th>Location</th>
<th>Import</th>
<th>Export</th>
<th>Sent to location</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Central RSC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Nin-Jan</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Miersk</td>
</tr>
<tr>
<td>Central RSC</td>
<td>200</td>
<td>120</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Nin-Jan</td>
<td>650</td>
<td>500</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>140</td>
<td>440</td>
<td></td>
</tr>
<tr>
<td>Miersk</td>
<td>260</td>
<td>490</td>
<td>210</td>
</tr>
<tr>
<td></td>
<td>210</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 5.17:** Example of correspondence and difference in timing of two flows of load units.
Chapter 5. Proposal for a computer aided terminal design method

Consider figure 5.17 in which two different flows of load units in time are shown, as a result of two different six-hour simulations. The left-hand side of figure 5.17 shows two flows that tend to show a simultaneous peak behaviour, while the right-hand side shows two flows that tend to show an inverted peak behaviour. If at SDM level of detail these two flows of load units were to be served by the same equipment pool, then the left-hand situation would yield a higher demand for equipment to install than the right-hand side situation would.

Analysis at UCM level of detail is used to find out quickly which flows tend to be strongly interrelated as shown in the left-hand side of figure 5.17 and which are not. This insight of the interrelation between various flows in the modelled terminal complex may be used either to consider changes in timing of trains and vessels (when allowed) or to consider changes in the stacking times of load units.

Analysis results on timing and interrelation of flows at UCM level of detail are shown as follows. All flows, between all locations and over all closed tasks are summed up for each single time-step of the continuous simulation. In this way an impression of the total intensity of all flows through the modelled system is obtained. In this total flow all peaks are signalled as those time-steps at which the total flow succeeds a threshold of about 70% of the maximum value of this total flow. Finally, for all these subsequent peaks it is analysed which flows are the most substantial ones, causing the greatest peak flow. By comparing these data for subsequent peaks a quick insight in the flows that are the main cause for peak load is obtained.

It is important to note, however, that the value of these results at UCM level of detail are limited. This is caused by the simple way flows of load units are controlled and timed at UCM level of detail: There is no difference in stacking time for various closed tasks, and there is no way to control the time load units are to reside in the modelled system in the way this can be controlled at SDM level of detail. Therefore, analysis at UCM level of detail will be used more for the purpose of quick validity checking than for the purpose of a rough insight in timing and interrelation of flows.

5.4.2. Aimed results at SDM level of detail

Compared to UCM level of detail, SDM level of detail provides very different types of results. This is a consequence of (1) the wide variety of additional elements at SDM level of detail, and (2) the use of discrete simulation for analysis instead of continuous simulation. As a result it is possible to discern individual load units and machines like trucks at SDM level of detail. All aimed results are divided into the same three groups that were found previously when discussing the model elements:

- rail operations;
- internal transport;
- stacking.

All of these groups cover both data on the usage of resources (like rail infrastructure and stack blocks) and on the costs that are involved with the usage of these resources.

♦ Rail operations

For analysis of rail operations at SDM level of detail a number of points of view could be adapted, each one resulting in a different set of results: (1) It could, for instance, be chosen for to put focus on overviews and schemes of the presence of particular trains at various locations; (2) or it could be chosen for to put focus on the internal service of locomotives to trains that want to move, resulting in a time that elapses between a request for a locomotive by a train and the actual assignment of a locomotive to that train; (3) or it could be chosen for to put focus on the delays of trains that leave the modelled terminal system, compared to their planned system exit time. This section gives only a description of part of al these options, which are those options that are implemented in TORCMOD at the moment this thesis is written. In this way it is believed to give a good

---

24 That is, if there is no or limited freedom to shift either of both flows in time, thus avoiding the simultaneous peak in demand for equipment capacity.

25 Another approach would have been to combine the characteristics of the current UCM with those of SDM level of detail into a new definition of the UCM level: It would be preferred to (1) specify at UCM level of detail the time load units reside in the modelled system through equipment routings, completely analogous to SDM level of detail, combined with (2) fixed stacking times as found in the current UCM level of detail that are, however, determined dynamically during simulation, based on the time these load units are to reside in the system and on the previously mentioned "vague" preferences for the presence of load units at the various locations over which their closed task runs, as specified at the flow patterns that are used at SDM level of detail.

26 Only these aspects are currently implemented in TORCMOD.
impression of the currently available possibilities for system analysis, which are likely to be extended in near future.

The currently available possibilities for system analysis with respect to rail operations can be divided roughly into two categories, which are:
- results on the usage of rail infrastructure;
- results on the usage of locomotives.

Strongly related to the first category is information on routing of individual trains over the rail infrastructure. Both categories will be discussed in more detail.

All train movements that are made during simulation at SDM level of detail are recorded individually at all infrastructure elements over which the movement runs. This recording covers both timing of the movement (i.e. the start- and end time) and ownership of the movement (i.e. either a light engine or a particular train of a particular service). By recording all train movements that occur, detailed analysis after simulation can be performed. It can, for instance, be shown on a map of the modelled system where a particular train of a particular service has driven -- and has been handled. Or it can be shown where all trains of one particular train service have been handled: Have they been handled at totally different bundles, or is there a regular pattern? Or it can be shown where all light engines have driven.

Based on this detailed information on train movements also information on tracks, points and bundles can be obtained: It can be shown which trains have been present during which periods of time, and for rail bundles an advice on the number of tracks to install can be obtained. This advice is based on a histogram as shown in figure 5.18, giving the percentage of time (relative to full simulation time) that 1 to 4 tracks were needed at that particular bundle. Combined with a so-called service level that specifies the percentage of time that the (free) demand for rail capacity must be satisfied the advised number of tracks to install can be determined. Refer to Van Zijderveeld [35] for more details.

When to each rail infrastructure element (that is, tracks at handle bundles, tracks at non-handle bundles and normal tracks, along with all available types of points) fixed and variable costs are assigned, a quick impression of the required rail infrastructure investment costs can be obtained. This costs analysis is also based on the previously discussed number of tracks to install at rail bundles, as a function of the required service level.

More information can be obtained on the operations of locomotives. It can, for instance, be inquired what number of D-locs is to be installed at the terminal complex in order to guarantee a maximum waiting time for trains between the moment they request for a locomotive and the moment they are actually assigned a locomotive. Another aspect on which information can be requested is the time that is spent to various actions by locomotives. In total four different types of actions are discerned, which are:
- idling while being uncoupled from any train;
- idling while being coupled to a train;
- driving while being uncoupled from any train; this only occurs either when the light engine has to drive to pick up a train, or when the light engine has to be stored at some storage bundle;
- driving while being coupled to a train.

---

27Important to note when using this service level at SDM level of detail is, that it is related to the totally unrestricted demand for rail infrastructure capacity, that is, without the effects of a block safety system. By consequence, in many cases the top 10% of the requested rail infrastructure capacity can be ignored, as not offering this capacity will not highly impede rail operations. Through improved planning of rail movements, imposed by the presence of a block safety system, this top 10% of requested rail infrastructure capacity will turn out to be of limited value. Therefore, in most cases the service level can be set to about 90%.
The results of these times are highly dependent on (1) the selected train operations with respect to light engines and (2) the selected strategy to find the best route and to find the bundle to obtain a requested locomotive from.

Also information on the numbers of trips that have been made and on the numbers of locomotives that have been requested by trains can be requested. Concerning this latter type of information some additional data is available: It can be asked which percentage of the made requests has been answered instantly by the assignment of a locomotive, and which percentage has had to wait, along with the average waiting time. Finally, the number of requests for locomotives that has been left unanswered at the end of simulation is given, which is very important indeed to signal a substantial short of locomotives. Such a short of locomotives inevitably results in delays of trains and by consequence in increased required infrastructure capacity.

To all actions of locomotives (idling, driving, coupled, etc.) costs can be assigned, along with personnel costs and (imaginary) fines for waiting times of trains, that had to wait for the assignment of a locomotive.

All data on light engines is split into data on E-locs and in data on D-locs, improving to make a clear distinction between locomotives that are used internally only (D-locs) and locomotives that enter the modelled system from the public rail network (E-locs).

**Internal transport**

Compared to analysis of rail operations, the analysis of internal transport is less detailed. It provides information on the number of machines that was needed during simulation, based on the results of the equipment optimisation method named equipment hopping (refer to Chapter 7) to find the optimal short- or mid-term planning of the use of equipment. This information is presented in the form of histograms, analogous to the previously discussed demand for rail infrastructure capacity. Based on this information an advice on the number of machines to install at the various pools can be given. All data on the use of equipment can be split into separate closed tasks or equipment routings.

Based on data on the number of required equipment an impression of the involved costs can be given. These costs cover both investments and personnel costs. A better insight in the personnel costs is obtained by discerning between two type of personnel costs: (1) personnel costs that are involved by the working schemes that were specified for the people manning the equipment and (2) personnel costs that are involved by the times the equipment was actually used. In those cases were a peak load of equipment found, this will result in a major discrepancy between both types of personnel costs.

**Stacking**

Analysis data on stacks is rather straightforward: For all stack blocks time-graphs can be obtained that give an overview of (1) the numbers of load units that have been stored, (2) the numbers of load units that have been un-stacked, and (3) the numbers of load units that have been stacked. Based on these data and a required service level an advice on the size per stack block, expressed in numbers of generalised load units can be given. This method is completely analogous to the advice on numbers of tracks per rail bundle as discussed previously at the rail operations. If required, these data can be split into numbers of load units, sorted per closed task, per load unit type or per client. Also a total for the entire internal stack, summed over all stack blocks, can be given.

If the advice on the size of the stack blocks is adapted and applied for the internal stack, then a static cost analysis is performed based on following three parameters:

- costs per square meter stacking surface (i.e. surface actually used for stacking);
- costs per square meter stacking surface, per layer. This option is used for stacks that are implemented as a scaffolding rather than as normal, conventional stacks; these costs are only calculated from the second layer on;
- costs per square meter driving path (i.e. surface used for equipment to drive over, and can not be used for actual stacking).

In this way a quick insight in the costs involved with a particular stack lay-out can be obtained.
5.4.3. **Aimed results at DDM level of detail**\(^{28}\)

There is no difference between the available analysis parameters at SDM and at DDM level of detail. The only difference is their value that may be different for both levels of details, because of the addition of new aspects at DDM level of detail. In this way a good insight in the effect of these newly added aspects is obtained. Especially for rail operations, to which the aspect of *block safety systems* is added, this effect will be obvious.

---

\(^{28}\)These elements have not been implemented in TORCMOD yet.
Chapter 6. Load schedules and contra schedules

6.1. Introduction

Large-scale terminal complexes require a lot of data to model them properly: Data on lay-out, data on equipment, data on stacking facilities, and data on load units that are to be loaded and unloaded at the system's boundary. It is this latter group of data that is discussed in this chapter. It will be shown that this group of data is one of the most important, but in the meantime one of the most arduously modelled groups. It will turn out that modelling flows at a large-scale terminal complex is thus complicated that a special structure is needed to model them properly without too much effort. This special structure will be based on load schedules, contra schedules and equipment routings as discussed at Chapter 7.

The problems encountered when modelling flows of load units entering and leaving the system will be illustrated by discussing two figures, figure 6.1 and figure 6.2. The first one shows a simplified map of the modelled system in which various locations can be recognised: Two stacks, one rail service centre and two marine terminals named "Nin-Jan" and "Miersk" respectively. Furthermore some additional locations are shown which are of no great importance at this point, and some connections between all these locations. Finally, three slant double-headed arrows are shown, indicating the locations at which load units may either enter or leave the system. These latter arrows represent the flows of load units at a high level of abstraction that will be elaborated in detail in this chapter.

The second figure shows the trains and marine vessels that "dock" at all three locations at the system's boundary and need to be loaded and unloaded there: Two trains are to be handled at the rail service centre, four marine vessels are to be handled at marine terminal "Nin-Jan" and three other marine vessels are to be handled at marine terminal "Miersk".

Figure 6.1: Example situation of a terminal complex.

Figure 6.2: Docking marine ships and trains in the example situation.

Figure 6.2 also suggests that several combinations of "origin" and "destination" of flows of load units are allowed: Load units could flow either from the rail service centre to one of the marine terminals, or from one

---

1Note that in the future situation at the Maasvlakte the number of locations will be definitely greater than shown in this figure, thus yielding a more complex model.
Chapter 6. Load schedules and contra schedules

marine terminal to the other marine terminal. Furthermore, using the stack of the rail service centre appears to be optional, as this particular stack is skipped by one of the slant connections.

Note that figure 6.2 does not show the numbers of load units that are to be loaded and unloaded, nor the moments in time at which these load units are to be loaded or unloaded. In other terms, we need a lot of additional information to model the flows of load units entering and leaving the system properly. However, suppose that this additional information is known through a thorough market analysis and is as shown in table 6.1.

<table>
<thead>
<tr>
<th>Service</th>
<th>Load</th>
<th>Unload</th>
<th>ETA²</th>
<th>ETD³</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train 1</td>
<td>40</td>
<td>45</td>
<td>8:30=10 min</td>
<td>12:14</td>
<td>Daily</td>
</tr>
<tr>
<td>Train 2</td>
<td>36</td>
<td>31</td>
<td>9:06=05 min</td>
<td>18:25</td>
<td>Daily</td>
</tr>
<tr>
<td>MS Josephine</td>
<td>12</td>
<td>25</td>
<td>15:15=10 min</td>
<td>18:25</td>
<td>Daily</td>
</tr>
<tr>
<td>MS Gartmore</td>
<td>120</td>
<td>89</td>
<td>Thu 18:25=35 min</td>
<td>Fri 17:20</td>
<td>Weekly</td>
</tr>
<tr>
<td>MS Xantippe</td>
<td>452</td>
<td>578</td>
<td>Sat 11:00=50 min</td>
<td>Sun 7:40</td>
<td>Weekly</td>
</tr>
<tr>
<td>MS Seabell</td>
<td>90</td>
<td>106</td>
<td>Mon 5:00=25 min</td>
<td>Mon 17:35</td>
<td>Weekly</td>
</tr>
<tr>
<td>MS Paster</td>
<td>67</td>
<td>92</td>
<td>12:15=10 min</td>
<td>16:15</td>
<td>Daily</td>
</tr>
<tr>
<td>MS Pride</td>
<td>216</td>
<td>180</td>
<td>Fri 22:10=30 min</td>
<td>Sun 20:20</td>
<td>Weekly</td>
</tr>
<tr>
<td>MS Hong Kong</td>
<td>90</td>
<td>56</td>
<td>Thu 14:05=25 min</td>
<td>Thu 19:00</td>
<td>Weekly</td>
</tr>
</tbody>
</table>

Table 6.1: Example of specifications of flows of load units that are to be loaded and unloaded.

Table 6.1 shows for each train and vessel how many load units they will load and unload, at what time they will arrive including their disturbances, at what time they are to leave the terminal complex, and after what interval they arrive anew.

Supposing that the above referred to market analysis also found that all vessels and all trains are to be unloaded as soon, and loaded as late as possible, then all data needed to model the flows at the system's boundary appears to be present. Furthermore, it might seem as if all data needed to construct the flows through the system is available, too. So, it might seem as if we are able to find important aspects like required stack sizes and required numbers of equipment to install at various locations, based on the above data. However, it will turn out that this data is only a small part of the total amount of data needed to model flows through the system properly.

What additional data is needed? First of all, there is a need to relate flows entering at one side of the system to flows leaving at another side of the system. In other terms, there is a need to relate for example flows that enter at "Nin-Jan" to flows that leave either at the rail service centre or at "Miersk". This should result in some overall percentage of all load units entering at "Nin-Jan" that are to leave again at the rail service centre and some (complementary) percentage of all load units entering at "Nin-Jan" that are to leave at "Miersk".

Second, sometimes a more accurate relation between flows entering and leaving the system is required: In some cases, for example, a certain number of load units transported to the system by Train 1 must be transported out of the system again by "MS Gartmore" -- and not by some other ship docking at marine terminal "Nin-Jan". Or, for example, all load units entering the system by "MS Pride" and "MS Paster" must leave the system through either "MS Xantippe" or "MS Seabell", as these four ships belong to the same shipping company which uses the Maasvlakte system as a hub for pure sea-to-sea transport. This type of accurate data is especially used when modelling the present or future situation.

Third, there is a need to safeguard the equilibrium of the system, ensuring that, taken over a longer period of time, the total amount of load units entering the system will leave the system again. In the example shown in table 6.1 this equilibrium is not safeguarded at all as it specifies that every week 1584 load units are to enter the Maasvlakte system and that 1716 load units are to leave it again. Therefore, every week 132 load units are removed from the Maasvlakte system though they have never been brought to it. This aspect is of special importance when the analysis of the modelled system through simulation covers a rather long period of time.

²Abbreviation of: Expected time of arrival; includes some random disturbance.
³Abbreviation of: Expected time of departure.
like two weeks, which is a necessity for a thorough analysis in which the effects of various random disturbances in the system are included. Although in this example the undershoot of 132 load units a week seems to be small, it is about 8% of the total flow and therefore rather important. Furthermore, the system modelled through figure 6.1 and table 6.1 is relatively small and simple, compared to the actual future situation at the Maasvlakte. By consequence, it will be definitely more difficult to guarantee the equilibrium of flows for this actual, future situation.

*Fourth*, in some case different types of load units, flowing between the same locations at the system's boundary, are characterised by definitely distinct times they reside in the system: For instance, empty containers flowing from the rail service centre to "Miersk" definitely reside longer in the Maasvlakte system (possibly up to five weeks) than loaded containers do. Furthermore, these empty containers may be stored at separate stacks, separate from the loaded containers, and therefore follow a different "route" through the Maasvlakte system.

*Furthermore*, all these data should be entered such that it remains relatively simple to specify, alter and understand the modelled system and its flows, and that it is guaranteed at any time that the modelled system is consistent, without logical errors like a lack of equilibrium in flows of load units.

So, in short, we find that all flows of load units entering an leaving the system must be modelled such that (1) all flows are guaranteed to be in equilibrium at any time, (2) both high-level and detailed matching of incoming and outgoing flows through various transport services like trains and vessels are allowed, in any mix, and (3) differences in the times load units reside in the system and in the "route" they follow through the system, can be modelled.

**First approach**

One approach to model all these data is to link separate services as trains and vessels to each other in a rather rigid way, specifying for instance the number of load units that is to flow from a certain train to a particular vessel, that is to enter the system at some time. Differences in "routes" for various load units and in times they reside in the system are modelled through separate stacks that are reserved for, say, empty containers and through fixed times these load units are to remain stacked there. This approach would result in a huge input file for the simulation model specifying all individual flows that are to be simulated.

<table>
<thead>
<tr>
<th>From service</th>
<th>Load units</th>
<th>To service</th>
<th>Load units</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Train 1</td>
<td>uniform [10..20]</td>
<td></td>
<td>&quot;MS Seabell&quot;</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td>&quot;MS Seabell&quot;</td>
<td></td>
</tr>
</tbody>
</table>

*Table 6.2: Inconsistency when using the first approach.*

Drawbacks of this approach are (1) that the data entered is very complex as each relation between two services like trains and vessels is to be specified in a highly detailed way, even in those cases where such details are not needed at all for a reasonable system analysis. (2) Furthermore, it is very hard to guarantee the system's equilibrium as it takes a great amount of effort to match the sums of all detailed flows entering or leaving at all services to each other by hand. As a matter of fact, it is considered as virtually impossible to match these sums for a system as large as the future Maasvlakte system by hand, for several scenarios. (3) In those cases where all or some flows are modelled by means of random disturbances rather than by fixed number so load units, which indeed is advised for a good system analysis, another problem pops up. This problem will be illustrated by referring to table 6.2 which shows how a certain amount of load units flowing from Train 1 to vessel "MS Seabell" is specified as a uniform random distribution of 10 to 20 load units. If all flows of load units were to be matched in the rigid way of this first approach, then a problem would occur when specifying the flows entering at "MS Seabell": The same number of load units stemming from Train 1 should be included in the flows received by "MS Seabell", in the meantime safeguarding that the total amount of load units received by this vessel equals 90 as specified in table 6.1. Now realise that a score of these random disturbances may be used in the system, thus complicating matching of flows tremendously. (4) Fixed stacking times omit a great amount of
the liberty of the model in planning the use of equipment and therefore in planning of flows of load units: Fixed stacking times result in a rigid model that is not in the position to explore alternative scenarios for stacking and for timing of flows, possibly resulting in alternative, better solutions in terms of the number of equipment needed to install at the terminal complex. This aspect is elaborated in more detail at the so-called flow patterns and at the equipment optimisation model.

In summary, it can be said that this approach yields a great amount of (manual) work on a complex type of data, while in many cases the high level of accuracy reached by this first approach is not needed at all. Especially when modelling situations that are to occur after about 10 years, it will be of no great interest whether the flow from Train 1 to "MS Seabell" will equal 13 or 18 load units in a terminal system that is not known in detail yet and thus complex that any uncertainty of flows stemming from Train 1 will be compensated by uncertainties of flows stemming from Train 2, "MS Pride", "MS Hong Kong", or some other service.

♦ Second approach
Another approach to model all these data is to specify only the total flows of load units entering and leaving the system through the various services as shown in table 6.1 along with percentages of load units being exchanged between all these services, and various fixed stacking times for various types of load units to model differences between empty load containers and loaded containers.

Drawbacks of this approach are (1) the problems as described at the first approach in case of random flows of load units, (2) the need to determine the sum of all incoming and outgoing flows to match them for equilibrium in each scenario that is analysed and (3) the same problems as found at the first approach concerning fixed stacking times, thus removing all liberty of the model to find more efficient stacking times for load units.

The benefit of this approach is that the relation between the various services can be modelled in a quite simple and powerful way, increasing the insight in the system that is modelled.

In summary, it can be said that, irrespective of the benefit of simple modelling of relations between various services, the matching of flows is too complicated, especially in case of random disturbances, to be done by hand. Furthermore, due to the way the course of flows through the system is specified, i.e. in a rather "loose" way through fixed stacking times, the possibilities to obtain a thorough insight in the interference of flows after simulation are limited at this second approach. Obviously, this also holds for the first approach. This problems could be solved by introducing some "backbone structure" that increases the possibilities to manipulate individual flows, if desired, as well as the possibilities for a thorough system analysis after simulation.

♦ Chosen approach
The approach chosen in this thesis is quite different from both approaches described before. The aim of the chosen approach is to create a powerful structure that is capable of modelling all aspects mentioned in the beginning of this section, that in the meantime offers a sound basis for system analysis after simulation, thus offering better facilities to examine and improve the modelled system. This approach heavily relies on the closed tasks as described in Chapter 5 and adds some more elements to model and control flows of load units.

The first problem tackled by the chosen approach is the problem of balancing flows through the system. This is done by discerning two types of schemes modelling services: load schedules and contra schedules. Both types of schemes have been described in a rather brief way in Chapter 5.

6.2. Load schedules

Load schedules come quite close to the services as meant in table 6.1 as they specify the number of load units that are to be loaded or unloaded and the time at which these services enter and leave the system. However, these load schedules contain some more parameters to specify the flows of load units up to the desired level of detail. All parameters that define a load schedule are discussed in two parts: The first part deals with all parameters that specify the behaviour in time of each load schedule as a whole, while the second part deals with all parameters that specify the behaviour of all "load packages", i.e. all "sets of load units" that are incorporated in each load schedule.
Chapter 6. Load schedules and contra schedules

- **Parameters of load schedules - 1/2**

  First of all the parameters specifying the timing of services as mentioned in table 6.1 when modelled through load schedules are given. These parameters are:

  - time of system entry of the first occurrence of the service, specified as a fixed time;
  - disturbance on the system entry time of any occurrence of the service, specified as a random distribution;
  - time of presence after the disturbed entry time of any occurrence of the service, specified as a random distribution;
  - repetition interval of the service, specified as a fixed time.

  From these parameters it can be found that in the chosen approach all load schedules represent regular services, services that are to occur at a regular repetition interval. Obviously, in some cases services are to occur only once and not at a regular interval; fortunately, this can be modelled through an extremely long repetition interval, making the service to occur only during the period of time that is simulated for system analysis.

  Figure 6.3 will help understanding the above parameters. It shows the presence in the system of some service that is modelled through a load schedule, as a function of its fixed system entry time, its random disturbance on this entry time and its random presence in the system, from the disturbed system entry time on.

  ![Figure 6.3: Timing of load schedules.](image)

  There is one major difference between the system exit time modelled at load schedules and the system exit time as specified at table 6.1: While load schedules are specified by the duration of their presence in the system, table 6.1 only specifies the end time of the presence of the service. Therefore, according to table 6.1 a service should be present for a smaller amount of time than intended whenever some disturbance on its system entry occurred. But we just found that services modelled through load schedules are present for some known amount of time, irrespective of their disturbance on system entry. In short, the duration of presence is not influenced by any disturbance on system entry for load schedules, but for services as specified in table 6.1 it is. This difference has some major consequences on timing of load schedules when the disturbance on system entry is great.

  Which of both methods is best? A fixed time for system exit or a "fixed" duration for system presence? This choice heavily depends on the type of service that is modelled through the load schedules: Barges behave different from sea-going vessels, and they in their turn different from trains. First of all, it is important to realise that the size of disturbance on the system entry time may differ for various load schedules: Whenever the size of this disturbance is great then the system exit time will be influenced greatly, too. On the other hand, whenever the size of this disturbance is small then the system exit time will be influenced less.

  For sea-going vessels the disturbance on system entry time may differ significantly between two occurrences: At some times the delay may be great (several hours) and at other times it may be small (several minutes). However, one important aspect of sea-going vessels is that they are thus expensive that they are to be loaded and unloaded as soon as possible, thus having a more or less constant duration of presence, regardless of their disturbance on system entry. However, in some cases problems may occur as found with barges.

  For barges the disturbance on system entry may be great compared to their duration of presence. In some cases this may result in a reduced duration of presence, thus forcing the terminal complex to load and unload the barges faster than normal. However, in some cases the normal loading speed of barges is quite low due to the low priority these barges sometimes have. Although these effects on timing of barges may be serious in specific cases, they are not considered as too important in the model described in this thesis, as it focuses on rail terminals rather than on barge terminals.

---

4Note that any distribution mentioned in this thesis could be specified as a "distribution" returning at all times some constant value. If required, this constant value could equal zero and thus eliminate any effect of any disturbance modelled through such a distribution.
Chapter 6. Load schedules and contra schedules

For trains being handled at rail terminals the planned ETD must be considered as fixed as trains have to leave the system in time to catch their reserved "rail path" on the rail infrastructure. This is a consequence of the "sequential" character of movements over rail infrastructure and of the safety rules that apply for rail movements. Therefore, when load schedules model trains instead of vessels, the ETD as given in table 6.1 should be considered as fixed and may not be delayed by any disturbance on system entry. During simulation of trains at UCM level of detail this error in timing of trains is ignored due to the limited importance of analysis at this level of detail and its additional major simplifications. During simulation of trains at SDM and DDM level of detail, however, this error is corrected for as at these levels of detail trains are modelled through so-called train services to which a number of load schedules can be assigned as described in Chapter 9, rather than through plain load schedules; in contrast to plain load schedules, these train services do have a fixed system exit time.

♦ Drawback of the chosen approach

One drawback of the use of load schedules and contra schedules is, that it does not support a situation in which services like barges only are invoked whenever enough load units are present at the concerning barge terminal. This in contrast to barges that arrive at a regular interval. However, this situation can be approached by modelling a number of regular contra schedules, representing the times at which these barges could arrive, possibly combined with a rather great disturbance on their system entry times. Although a more fancy solution could be created this has not been chosen for at the moment.

Note that all random distributions used in this thesis return sample value that may equal or exceed zero, but may never be less than zero. By consequence, the disturbance on system entry as shown in table 6.1 in column ETA can not be modelled by such a random distribution as its size was specified as ±10 minutes. This difficulty can be helped, if required, by lowering the specified ETA by 10 minutes, by increasing the random disturbance to the range [0..20] minutes instead of [-10..10] minutes, and by increasing the average value of the disturbance to 10 minutes. Note that in practice in most cases only delays are of importance, as vessels that arrive too early in a number of cases do not need to be serviced immediately, and therefore do not influence the process of the terminal complex.

♦ Parameters of load schedules - 2/2

The second part of parameters of each load schedule specifies the behaviour of all "load packages", i.e. all "sets of load units" that are incorporated in the load schedule; until now it has only been spoken of the behaviour and timing of the load schedule as a whole without even mentioning the numbers of load units that are to be loaded or unloaded. However, this aspect is discussed here.

Each load schedule contains one or more so-called load volumes that can be considered as packages of load units of a certain type and of a certain quantity that belong to some client and are to be loaded or unloaded at that location at which the load volume is specified. In the example shown in figure 6.4 this location is "Central RSC". In other terms, the load volume shown in the example of figure 6.4 specifies that all load units that are included in that load volume are to be loaded or unloaded at the rail service centre. Obviously, we also need to specify to which location any load unit that is unloaded according to these load volumes is to flow. Therefore, we need to specify the opposite location of the rail service centre for each particular load volume.

In some cases this opposite location, which is marine terminal "Nin-Jan" in the example of figure 6.4, can be reached over several closed tasks. Therefore, we also need to specify the closed task over which the load units are to flow between the rail service centre and marine terminal "Nin-Jan". Such a closed task in its turn specifies the stacks, exchange points\(^5\) and other locations over which the load units are to flow. The selected closed task

\(^{5}\)Refer to Chapter 5 for an introduction on exchange points.
is also used to manipulate the flows of the load volumes during simulation by means of *flow patterns* and *equipment routings* as discussed in Chapter 7.

Additional to this information on load units and closed tasks, each load volume contains information on the *action* that is to be done with these load units; in other terms, each load volume contains information on the *flow direction* of the load units (are they loaded or are they unloaded) and on the *timing* of loading or unloading of these load units, relative to the load schedule to which the load volumes belong. This timing is modelled by two random distributions. The first one specifies the moment at which loading or unloading starts, relative to the start of the load schedule. This distribution is called the *left offset*. The second one specifies the moment at which loading or unloading ends, relative to the start of the load schedule. This distribution is called the *right offset*. An example of these offsets is given in figure 6.5.

Some examples of the use of these two offsets are given in table 6.3. Each of these examples is characterised by totally different parameters: there are differences in (1) the times the load schedules to which the load volumes belong are present, (2) the left-offsets and (3) the right-offsets. From table 6.3 it can be seen that both the left- and the right offset are specified as a percentage, relative to the total duration the load schedule is present in the system. There is a clear reason for this choice: In this way the speed at which load units are to be loaded or unloaded is dependent on the (random) time of presence of the load schedule. Whenever the load schedule is present during a smaller period of time than normal, for whatever reason, then the load units are to be loaded and unloaded faster than normal.

These percentages, determining the offsets, are random distributions themselves. However, in most cases these random distributions will be specified as a constant value, which can be modelled by means of a random distribution, too. Some standard settings for the offsets are unloading in the first 50% of the time the load schedule is present, and loading during the remaining 50% of this time. It will be clear that some special rules apply to the random distributions specifying the offsets. *First*, they ought to be in the range [0…100%] as an offset can not be greater than the time the load schedule is present. *Second*, the sum of both offsets should never equal or exceed the total time the load schedule is present in the system. These rules are to be safeguarded by the terminal design aiding tool.

In short, the second part of parameters for each load schedule consists of:

- a list of load volumes, each load volume specifying:
  1. number of load units to load or unload, specified as a random distribution;
  2. type of the load units;
  3. client (owner) of the load units;
  4. direction of the flow, relative to the location at which the load schedule is specified; this direction can be either *load* or *unload*;
  5. opposite location, opposite of the location at which the load schedule is specified;
  6. closed task, connecting the opposite location to the location at which the load schedule is specified;
  7. offset of start of loading or unloading relative to the start of the related load schedule, specified as a random distribution;
  8. offset of end of loading or unloading relative to the start of the related load schedule, specified as a random distribution.
Even though it might seem to be complex to specify all this data, in most cases it is not. In a number of cases the person modelling the real-life system will not bother about the clients of load units, and therefore can be left to some default value. In some other cases the type of load is important for the model due to different behaviour of the model for different types of load units (like different priorities and stacking times), but these differences do not necessarily need to be modelled by specifying the type of load unit for each individual load volume: These differences can also be modelled by creating different closed tasks with different characteristics, to which load units of different types can be assigned.

This is shown in figure 6.6, in which two different closed tasks between the rail service centre and marine terminal "Nin-Jan" are depicted. To each closed task totally different characteristics can be assigned, modelling the differences in load unit types6. By consequence, load volumes that specify flows of empty load units could be related to the lower closed task, while load volumes that specify flows of loaded load units could be related to the upper closed task. In that case the actual type of load units specified in the load volumes can be left to some default value. However, if desired it is also possible to specify differences in types of load units in the load volumes.

Finally, the random distributions specifying the offsets can easily be modelled by some standard settings, facilitating modelling of large numbers of load volumes.

6.3. Contra schedules

In some respects contra schedules are very much like load schedules: They are used to model services like the ones shown in table 6.1, or possibly to model truck gates at which external trucks enter and leave the system. In other respects contra schedules are totally different from load schedules: They do not contain exact information on numbers of load units to load or unload. So, contra schedules are partially equal to load schedules, and partially not.

Load schedules and contra schedules are linked by means of closed tasks. As figure 6.7 rightly suggests, load schedules and contra schedules are always used in pairs: Any load flow stemming from a load schedule flows towards a contra schedule, and vice versa. In other terms, any load unit that is unloaded at a load schedules is loaded at the opposite side of its closed task at a contra schedule. And vice versa, any load unit that is loaded at a load schedules is unloaded at the opposite side of its closed task at a contra schedule. So, each load unit is registered both at a load schedule and at a contra schedule at the opposite side of the terminal complex, linked by a closed task.

---

6Note that this induces the need for a safety check that guarantees that only flows of load units of the proper type, i.e. the type for which the closed task was intended and created, can be assigned to that closed task. In this way modelling errors are prevented.
Figure 6.8 shows the way load schedules are linked to contra schedules over the closed task(s) connecting the location at which they are defined. The "opposite location" mentioned in this figure is the location at the opposite side of the closed task, opposite to the location at which the load schedule is defined.

That is why contra schedules do not contain exact information on numbers of load units to load or unload: Registering load units once at a load schedule automatically results in registering the same load units at some contra schedule at the opposite side of the terminal complex. This immediately gives rise to the question to which particular contra schedule these load units are assigned, or in other terms, how load schedules and contra schedules are matched. This subject will be discussed shortly. But first another problem will be tackled.

In figure 6.9 the same closed tasks from figure 6.7 are repeated, with the addition of two words: "push" and "pull". These two words indicate the direction of the flow of load units relative to the load schedule they stem from: Whenever they are unloaded at their load schedule, that is, when the direction of the load volume they are specified at is unloading, then these load units are actually pushed into the system by their respective load schedules. Opposite, whenever load units are loaded at their load schedule, then these load units are merely pulled out of the system by their respective load schedules.

Especially for pull flows it will be clear that a problem arises: When are these load units, that are pulled towards their load schedules, to start moving towards the location at which these load units are to be loaded, the location at which their load schedules are defined, in order to arrive in time over there? Even more, at what time are these load units to enter the system at the opposite side of their closed task, in order to arrive at the right moment at their load schedules? This problem is solved at SDM and DDM level by introducing the so-called equipment routings as discussed in Chapter 7. Each of these equipment routings is related to exactly one closed task, and to only one direction over that closed task. Therefore, each closed task yields precisely two equipment routings, for each direction one.7

To each equipment routing, and therefore to each flow direction over a closed task, a random distribution is associated that determines the duration of the flow over that equipment routing. In other terms, that random distribution determines the difference in time between system entry of a load unit (either at a load schedule or a contra schedule) and system exit of that load unit (either at a contra schedule or a load schedule). As this duration is defined by a random distribution, its value may be totally different for two load units that stem from the same load schedule.

An example of these random distributions for two equipment routings stemming from the same closed task is given in figure 6.10. Note the, purposely exaggerated, great difference between these two distributions. In this way, for instance, it can be modelled that load units flowing from a truck gate to a marine terminal reside definitely shorter in the modelled terminal complex than load units flowing in the opposite direction do.

---

7Please do not confuse these two directions with the push and pull direction, as these latter two express the relative direction of a flow of load units, relative to its load schedule, while the first two do not.
Another remarkable consequence of the chosen structure is that fixed stacking times are not needed and therefore are not used at SDM and DDM level of detail. The actual stacking times at the various locations over which a closed task runs and at which load units may be stacked, are left to the simulation model to find out, along with the required stack sizes. This especially is important when closed tasks run over several locations at which load units may be stacked. If required, stacking times still can be manipulated by manipulating the settings of the flow patterns and the preferences of the equipment routings as discussed in Chapter 7. In this way the total of flow characteristics of the model can be altered and analysed for different scenarios, while in the meantime the flows across the system's boundary remain unchanged! So, by specifying a duration per equipment routing rather than fixed stacking times we have created a situation in which the internal flows, which are influenced by the internal structure of the terminal complex that must be optimised, are separated from the external flows which are imposed by schedules of trains, sea-going vessels and barges. In other terms, we have created a strict distinction between available degrees of freedom that need to be explored on one side, and limiting conditions that need to be respected on the other side.

**Flows at UCM level of detail**

As stated previously equipment routings are not used at UCM level of detail, and therefore flows at UCM level of detail can not be controlled through the duration of flows attached to these equipment routings. By consequence, some other mechanism is needed to control these flows. This mechanism has been chosen to be simple fixed stacking times at stacks, an approach that carefully has been avoided at SDM and DDM level of detail. The resulting structure is given in figure 6.11 showing some fixed stacking time at a stack over which the shown closed task runs. As internal transport between the various locations at UCM level is modelled as a fixed transport time on each part of the closed task, it is simple to determine at what time pull flows are to start at the opposite side of the terminal complex, opposite of the location at which the load schedule is defined.

The way flows and transport times are modelled at UCM level of detail does not require the presence of contra schedules at the opposite side of the terminal complex, as it is fully determined at what time each load unit is to enter the system by the required exit time (specified at the pulling load schedule) and by the fixed times for stacking and internal transport. By consequence, contra schedules are not used at all during simulation at UCM level, even though they are defined at that level. This seemingly illogical structure has been chosen for as it allows to model contra schedules and load schedules at the same level of detail, thus stressing their relative coherence.

Before the selection process of the proper contra schedule for each load unit can be given another aspect needs to be discussed. This aspect is the difference between regular contra schedules and intensity contra schedules.

### 6.3.1. Regular contra schedules

Regular contra schedules are used to model services as mentioned in table 6.1, which are all services with a batch-wise character. This means that the services occur at a regular interval and contain some relatively great amount of load units that are either loaded or unloaded. This in contrast to single load unit transport like trucks as will be modelled by intensity contra schedules. But first of all regular contra schedules will be discussed.

As regular contra schedules model the same type of services as load schedules do, some part of their parameters is equal to those of load schedules; this is the part that specifies timing of the contra schedules. Another part of the their parameters is different from those of load schedules; this is the part that specifies numbers and types of load units.

So, the first part of all parameters of regular contra schedules that model a regular service are, completely analogous to those specified for load schedules:

- time of system entry of the first occurrence of the service, specified as a fixed time;
- disturbance on the system entry time of any occurrence of the service, specified as a random distribution;
- time of presence after the disturbed entry time of any occurrence of the service, specified as a random distribution;
• repetition interval of the service, specified as a fixed time.
Refer to table 6.1 for some example of services that could be modelled in this way.

The second part of all parameters of regular contra schedules contain rough information on the numbers of load units that can be transported by the modelled services. Although is might seem odd that it is spoken here of information on numbers of load units, that are modelled by contra schedules, it is not. Indeed, in contrast to load schedules, contra schedules do not contain exact information on the numbers of load units that are to be transported; this exact information is present at load schedules only. But nevertheless, contra schedules do contain some data on the numbers of load units to be transported, even though it is not as detailed as at load schedules.

Consider for example the services as shown in table 6.1 and find that vessel "MS Seabell" transports at average 5.3 times as much load units as vessel "MS Josephine" does. Supposing that both vessels are modelled as regular contra schedules, it seems to be reasonable to demand that the number of load units that is assigned during simulation to contra schedule "MS Seabell" actually does equal about 5.3 times as much as the number that is assigned to contra schedule "MS Josephine". If this demand is not made, then the results that were obtained after simulation of the modelled system would not be valid due to a great discordance between the actual and the simulated flows. So, we need to specify some capacity, specifying the relative size (or importance) of a regular contra schedule, relative to all other regular contra schedules that are defined at the same location at the system's boundary.

So, supposing that all vessels docking at marine terminal "Nin-Jan" are to be modelled as regular contra schedules, then their respective capacities could be as shown in table 6.4. These capacities are simply determined as the sum of the numbers of load units that are to be loaded and unloaded according to table 6.1. Note that the absolute values of these capacities are of no great importance; what is important are their relative values, relative to all regular contra schedules that are defined at the same location, that is, at marine terminal "Nin-Jan".

<table>
<thead>
<tr>
<th>Service</th>
<th>Capacity</th>
<th>Relative importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;MS Josephine&quot;</td>
<td>37</td>
<td>2.5 %</td>
</tr>
<tr>
<td>&quot;MS Gartmore&quot;</td>
<td>209</td>
<td>14.2 %</td>
</tr>
<tr>
<td>&quot;MS Xantippe&quot;</td>
<td>1030</td>
<td>70.0 %</td>
</tr>
<tr>
<td>&quot;MS Seabell&quot;</td>
<td>196</td>
<td>13.3 %</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1472</td>
<td>100 %</td>
</tr>
</tbody>
</table>

Table 6.4: Total capacity per service.

The capacities specified at contra schedules are expressed in numbers of so-called generalized load units in order to eliminate the effect of different types of load units. These generalized load units are also used in Chapter 7 when specifying the so-called carrying capacities of various equipment types. Obviously, these capacities play an important role when matching load schedules to contra schedules.

A nice shortcut in specifying these capacities of regular contra schedules is to base them on the duration of those contra schedules. In some cases, especially for analysis of long-term scenarios, this approach is sufficiently accurate. The basic assumption is that the average speed of loading and unloading is the same for all regular contra schedules docking at a particular location. In this way a large vessel is planned to remain docked for a longer period of time than a small vessel is; the large vessel is assumed to carry an amount of load units that relates to the amount carried by the small vessel as their respective duration do. Even more, as this duration is specified as a random distribution it may differ for each single vessel originating from that contra schedule, thus resulting in a different number of load units transported by each individual vessel. This in contrast to the method discussed at table 6.4 which results in the same amount of load units for all vessels of the same regular contra schedule.

As the capacities of regular contra schedules are all take relative to all other regular contra schedules docking at the same location, the choice for automatic capacities based on the duration of the contra schedule must be applied for all regular contra schedules installed at a particular location. In this way problems with matching capacities expressed in numbers of generalized load units with capacities expressed in time are avoided.

**Relating regular contra schedules to closed tasks**
In some cases it is required to specify that some regular contra schedules may only load and unload load units that are related to a limited number of closed tasks. In this way it can be modelled, for instance, that specific load schedules are related to specific regular contra schedules by means of a limited number of closed tasks.
An example of such a situation is given in figure 6.12 that is based on figure 6.2. This figure shows a situation previously discussed in which the Maasvlakte system is used by some shipping company as a hub for sea-to-sea transport. The four ships owned by this company are indicated in figure 6.12 by four double-headed black arrows.

As there is a need to distinguish between flows that are related to this particular shipping company and all other companies that use the Maasvlakte system, a separate closed task is created for particular shipping company. So, this closed task can be said to be "private" for that shipping company as shown at the top of figure 6.13. If, for instance, both vessels that dock at marine terminal "Nin-Jan" are modelled as load schedules, and both vessels that dock at marine terminal "Miersk" are modelled as regular contra schedules, then these latter contra schedules would be related to the "private" closed task belonging to that shipping company only. Analogous, the load schedules modelling both vessels at "Nin-Jan" ought to contain only load volumes that are related to that "private" closed task.

By default, a regular contra schedule can be used for any closed task. Only when deliberately specified, they can be related to a limited number of closed tasks.

*Contra schedules are not used at train handle areas*

Regular contra schedules are *not* used to model trains, that is, they are not used to model trains up to the level of detail that is reached at train handle areas. As can be found in Chapter 9 trains that are handled at train handle areas are modelled as so-called *train services* rather than as plain load schedules or as contra schedules. The reason for this is the number of details that is attached to these train services at SDM and DDM level of detail; details that are not included for sea-going vessels and barges, as the model described in this thesis focuses on rail service centres.

By consequence, as all services that are handled at a rail service centre are train services, it is not allowed to model contra schedules at train handle areas. This does not only hold for regular contra schedules, but also for intensity contra schedules. The result of this approach is that all flows of load units that are either loaded or unloaded by trains that are to analysed and simulated in a detailed way, and for that reason are modelled at a *train handle area*, are specified by means of load schedules. In those cases where the precise course of trains is not considered of great importance, it can be chosen for to model the rail service centre as a bounding element instead of as a train handle area. Then contra schedules *can* be used to model train services, but without all details that would have been added to the train services at SDM and DDM level of detail if they were modelled at a train handle area. This is purely a consequence of the primary objective of the model described in this thesis to be used for detailed analysis of rail service centres.

---

8. ...and therefore can only be used at bounding elements.
9. Refer to Chapter 5 for an introduction on bounding elements.
6.3.2. Intensity contra schedules

Next to marine terminals and rail service centres, a large-scale terminal complex may contain one or more truck gates. At these truck gates external road trucks pass the system’s boundary to deliver or pick up load units. Analogous to regular contra schedules, intensity contra schedules can only be defined at bounding elements, and therefore can not be defined at train handle areas.

Truck gates as found at terminal complexes are different from marine terminals in at least two aspects: First, truck gates are characterised by so-called continental time schedules instead of marine time schedules, which means that external trucks do not arrive around the clock as marine vessels do, but instead more or less during continental working hours. Road trucks tend to bring load units to the terminal complex at the end of the day, and pick them up at the beginning of the day. This effect is caused by the strong relation between these transports by road trucks and the "biorhythm" of the industries, industries that prefer to receive raw materials at the beginning of the day (needed for their production) and to send finished products at the end (needed to reduce their stocks and stocking costs). Second, transport at truck gates can be characterised as single load unit transport in contrast to batch-wise transport as found at marine terminals. Single load unit transport means that a great number of individual, independent trucks are to arrive according to some intensity-pattern rather than a limited number of regular services as found at marine terminals. Furthermore, each of these individual, independent road trucks represents transport of only one (or at most two) load units, this in contrast to regular services as found at marine terminals that may represent transports of up to several hundreds of load units at a time.

Even at the Maasvlakte system, which can be characterised as a mainly marine system, the effect of truck gates on the operations is great due to the tendency of single load unit transport by trucks to be peaked due to their strong interrelation with the "biorhythm" of the industries, and therefore with continental time schedules. The actual strength of this interrelation is determined by numerous factors like the type of industries that is served by the trucks, and the transport distances for the trucks. The future Maasvlakte system will be influenced by a new truck gate at the DistriPark. As this DistriPark may have a major impact on the operations of the future rail service centre at the Maasvlakte, truck gates and their single load unit transport characteristics are to be included in the model described in this thesis.

An important characteristic of load unit transport at truck gates is the great amount of individual transports (or trucks) that call at the truck gate. This amount of individual transports is considered too great to model them in a highly detailed way by means of regular contra schedules. In other terms, the number of transports is too great to model all of them as separate trucks, each truck having a known arrival time and being related to a known, particular load unit that is either to be delivered or to be picked up. Even more, when deciding to model these single load unit transports in such a detailed way then the problem of guaranteeing an equilibrium of the system that is being modelled would pop up again; the problem that has been avoided carefully by introducing contra schedules. Therefore, another method for modelling and scheduling individual trucks should be adopted, using some kind of transport intensity which is to be interpreted as the chance that a truck will enter the system in order to load or unload a container during some known period of time. This transport intensity should be different (or at least potentially different) for trucks that deliver load units (which are import flows) and trucks that pick up load units (which are export flows), as well as for the various weekdays: Trucks may show a different pattern of arrival for Mondays and Sundays, or they even may

---

10Whenever the number of load units to be transported by trucks would be defined by non-constant random distributions, then it would be even impossible to determine fixed schedules of trucks entering and leaving the system on beforehand, guaranteeing the system's equilibrium in the meantime.
show a different total number of trucks that shows up during each weekday. All these data should be included in the intensity contra schedules that model truck gates.

An example of intensity contra schedules modelling at truck gate at the system's boundary is given in figures 6.14 and 6.15, showing a different behaviour for two weekdays: Mondays and Sundays. Two different aspects of intensity contra schedules can be seen from these figures: The pattern defining the relative intensity of each hour of the day, relative to all other hours of that day, and the day-importance defining the total number of transports that occur in each day, relative to all other days in the week.

Figures 6.14 and 6.15 show two major differences between Mondays and Sundays: First of all the sequence of import and export is more or less reversed for both weekdays, which is caused by the special character of the Sunday in this example: It is mainly used to transport raw materials to industries in preparation of the next Monday. Second, the total number of load units transported through the truck gate is definitely greater at Mondays than at Sundays. In this way intensity contra schedules can be used to include a number of external effects in the model.

All numbers shown in numbers 6.14 and 6.15 must be considered as relative numbers rather than as absolute numbers, expressing either the relative importance of each hour of one day compared to all other hours of that day, or the relative importance of one weekday compared to all other weekdays. All these numbers may differ for both directions: Import and export. In summary, the data modelling intensity contra schedules can be specified as follows:

- for each weekday, and both for import and export direction:
  1. importance of that weekday, relative to all other weekdays
  2. importance of each one-hour period of that weekday, relative to all other one-hour periods of that weekday.

Intensity contra schedules do not provide any mechanism to relate specific closed tasks to specific road trucks, in contrast to regular contra schedules that do allow assign specific regular contra schedules to certain closed tasks. The reason for this can be found in the mass-character of intensity contra schedules: Intensity contra schedules do not offer any possibility to refer to a specific road truck, and therefore can not be used to relate a specific road truck to a specific closed task. Nevertheless, if required such a special situation can be approached by creating for each truck gate found in the real-life system two (or even more) different bounding elements: To each of these bounding elements, which are all located at the same position, different intensity contra schedules and different closed tasks can be assigned. By consequence, each of these bounding elements will receive load units from different closed tasks and in this way different intensity contra schedules for different closed tasks are included in the model. An example of such a situation is given in figure 6.16, showing how the truck gate of DistriPark is modelled by means of two different bounding elements, each possibly having different intensity contra schedules. Note that the two double-headed arrows represent two different sets of closed tasks that connect the gate of DistriPark to marine terminal "Nin-Jan".

This "trick" can only be used to model different intensity contra schedules for different closed, tasks, but also to model any mix between regular and intensity contra schedules at any bounding element: Just define

---

11 This example is not based on figure 6.1.
Chapter 6. Load schedules and contra schedules

several bounding elements at the same location; all these bounding elements will be connected to the remaining system elements by means of different closed tasks, and therefore each of these closed tasks will refer to different regular or intensity contra schedules.

Although there are a lot of differences between intensity contra schedules and regular contra schedules, there is one important correspondence: They can only be modelled at bounding elements and not at train handle areas. As discussed before, this is a consequence of the first concern of the model described of this thesis: Modelling and analysing rail service centres.

6.4. Matching load schedules with contra schedules

Now that we have tackled the problem of push- and pull flows and have learned about the random distributions that specify the total duration of flows and about the difference between regular and intensity contra schedules, we can focus on the selection of contra schedules. In other terms, we can focus on the way load schedules and contra schedules are matched. Obviously, the total duration of flows as specified at the equipment routings will play an important role in this\textsuperscript{12}.

Both for regular contra schedules and for intensity contra schedules, this total duration of flows will turn out to be a mere indication of the time that elapses between system entry and system exit of a load unit than a strictly defined and respected time span. The reason for this is quite straightforward, and can easily be derived from the example situation shown in figure 6.17.

![Figure 6.17: Problems when relating load units to (regular) contra schedules.](image)

This figure shows a situation in which a push flow of load units, stemming from some load schedule at the rail service centre runs to some regular contra schedule at marine terminal "Nin-Jan", that is used to model a sea-going vessel. Suppose that the duration during which a particular load unit, running over the related closed task, is to reside in the system is known to be the duration as shown in figure 6.17. This duration is the time that elapses between the moment the load unit enters at the rail service centre and the moment it is requested (or planned) to leave the system again.

In other terms, the planned duration the load unit is to reside in the system automatically results in a known time for system exit for that particular load unit at marine terminal "Nin-Jan". Obviously, removing that load unit from marine terminal "Nin-Jan" should occur by some regular contra schedule that models a vessel that docks precisely at "Nin-Jan" at that specific, requested, time.

But, unfortunately, figure 6.17 shows that not a single regular contra schedule is found that could be used to remove the load unit from marine terminal "Nin-Jan" at the requested time. It only shows that two other regular contra schedules are available, the first one being a little early and the other one a little late. Now the problem arises which of these two contra schedules, both modelling sea-going vessels, should be selected to match this load unit with. Or, in other terms, the problem arises which of these two contra schedules should be selected to remove this load unit from the system. The same type of problem arises for load units that are to be matched to intensity contra schedules instead of to regular intensity contra schedules.

The next two sections will show how load units are to be matched to regular and to intensity contra schedules, taking into account (1) the duration these load units are to reside in the system, (2) the timing of contra schedules and (3) their respective capacities.

\textsuperscript{12}As contra schedules are not used during simulation at UCM level of detail, the aspect of matching load schedules and contra schedules is not included at that level of detail.
6.4.1. Regular contra schedules

First of all, matching of load units to regular contra schedules is discussed. As can be seen from figure 6.17 the chosen structure of load schedules and contra schedules induces the unpleasant aspect of finding the best contra schedule for a load unit once that load unit’s system entry time is known. In this way it will never be guaranteed that the random distribution modelling the duration each load unit flowing over a particular equipment routing resides in the system, will be found for one hundred percent correctly in the simulation results. There always will be some discrepancies between these random distributions and the actual times that load units have resided in the system during simulation, after simulation has ended.

But fortunately these discrepancies do not affect the validity of the simulation results. As can be seen from figure 6.17 the particular load unit that could not be matched to a vessel docking precisely at the required system exit time, can be matched to a vessel that docks somewhat earlier, or to a vessel that docks somewhat later then desired. This aspect is not in contradiction with the modelled real-life system, or with the other two approaches to model flows of load units as discussed previously at the Introduction of this chapter, as will be argued in the next few paragraphs.

In the real-life system load units are booked to be transported by specific vessels or trains based on a number of aspects like the load unit’s final destination (Germany or Brazil), various shipping companies, and prices. As this selection process is too complex to model in a satisfactory way, some simplifications are made in any model of any terminal complex to approach this real-life situation closely enough. This results in some standard to specify in the model which vessel or train is used for transport of specific load units.

As showed in the Introduction of this chapter, at least three different standards (or approaches) exist to model the real-life situation, each having its own advantages and disadvantages. One of the advantages of the first two approaches (which were not adapted) is the strict and clear relation between specific vessels and trains, increasing insight in the flows. Unfortunately, these approaches also render the disadvantage of highly detailed (manually generated) input data and of complicated system balancing checks. By using the concept of load schedules and contra schedules the amount of effort to model a specific situation has decreased, but it has introduced the problem of matching contra schedules to load schedules.

However, when we realise that in most situations the relation between vessels an trains is not thus rigid as suggested by the huge amount of input data of the first two approaches, and that in many cases this relation is not known at all with any reasonable certainty for future situations, then the relative "loose" matching of load schedules and contra schedules is found to be sufficiently detailed for most situations, while in the meantime greatly reducing the amount of required input data.

♦ First selection of regular contra schedules

Now, returning to our matching problem, consider the situation shown in figure 6.18 which is based on figure 6.17. The only two modifications are the addition of the lower bound and of the upper bound of the random distribution that specified the duration of the flow between system entry and system exit, relative to the known system entry time as specified by its load schedule. The duration of the equipment routing is not that rigid as figure 6.18 may suggest: Depending on the value that was returned by sampling the random distribution modelling the duration the load units are to reside in the system, these load units potentially could have been found to leave the system.

![Figure 6.18: Boundaries to search a matching regular contra schedule in.](image)

Note that this figure depicts a push flow; in case of a pull flow the entire figure should be flipped horizontally.
anywhere between this upper- and lower bound. Any regular contra schedule that could potentially be used to match the load unit with must lie within these boundaries. So, these boundaries are the first aid to find a matching regular contra schedule, offering a first selection criterion to limit the choice. Another, second selection criterion will turn out to be needed to find the single, best regular contra schedule out of those being selected by this first selection criterion.

Even though it might be expected that any contra schedule is to be considered a potential match for the load unit whenever any part of that regular contra schedule lies within the boundaries that the load unit might reside in the system, it is not. A regular contra schedule is only selected if a particular part of the regular contra schedule must lie within the boundaries shown in figure 6.18, a part that depends on the direction in which the load unit that is to be matched flows, relative to its load schedule: For pull flows a different part of the regular contra schedule must lie within the boundaries shown in figure 6.18 than for push flows.

Referring to figure 6.9 it can be seen that any load unit that is to be loaded at a contra schedule flows in push direction, while any load unit that is to be unloaded at a contra schedule flows in pull direction. Depending on this direction either the first or the last half of the regular contra schedule must lie within the boundaries shown in figure 6.18 to make the contra schedule an interesting match for the load unit: For load units that are to be loaded at the contra schedule, the last half of the regular contra schedule must lie within these boundaries, and for load units that are to be unloaded at a contra schedule, the first half of the regular contra schedule must lie within these boundaries. The basic assumption of this subdivision of each regular contra schedule is that load units are unloaded during the first half of presence of the contra schedule, while load units are loaded during the second half. These two halves of each regular contra schedule are shown in figure 6.19; the first half is called the estimated range of the load unit's system entry time, while the second half is called the estimated range of its system exit time.

Figure 6.19: Estimation of ranges for system entry- or exit time of a load unit.

Note that the thus found ranges for system entry- or exit time are a first, rough estimation of the actual system entry- or exit time of the load unit. These estimations will be replaced later on by a more accurate estimation.

In those cases where not a single regular contra schedule is found in this first part of the selection process to match the load unit with, an error in the modelled situation has been encountered. This error could be fixed by temporarily expanding the upper bound of the random distribution that specifies the duration of the load unit's presence in the system. This new value for the upper bound should be specified by the user of the terminal design aiding tool, and will be used for future occasions in which these matching problems occur. Some reasonable value of this new upper bound would be 24 hours more than the current upper bound. Based on this new upper bound the first step in the selection process is repeated, until at least one matching regular contra schedule is found.

- **Second selection of regular contra schedules**

Now that a number of regular contra schedules have been found that match the boundaries for the load unit's presence in the system as given in figure 6.18, either on the first or the second half of their own presence as shown in figure 6.19, we need to find the best regular contra schedule from these selected ones. For this selection a special technique as described in Appendix A is used, called random selection.

Random selection of a matching regular contra schedule for a particular load unit is based on two parameters, each one expressing a different aspect of the likeliness of the contra schedule to be chosen. The first parameter is the distance in time between the requested system entry- or exit time, requested by the load unit that is to be matched. The second parameter is the relative occupancy of the regular contra schedule, relative to its capacity. These parameters will be discussed in detail.
The distance in time is taken between the requested system entry- or exit time for the load unit as shown in figures 6.17 and 6.18 on one hand, and an estimation of the granted system entry- or exit time if a particular regular contra schedule would be selected. This estimation of the granted system entry- or exit time is a single moment in time, in contrast to the ranges for system entry or exit as shown in figure 6.19 that were used for the first selection of regular contra schedules. Analogous to these latter ranges of time, the estimation of the granted system entry- or exit time depends on the direction in which the load unit flows, relative to its load schedule: In case of a pull flow the granted system entry- or exit time is estimated to be the mid of the left-hand period of figure 6.19 that represents unloading, and in case of a push flow the granted system entry- or exit time is estimated to be the mid of the right-hand period of figure 6.19 that represents loading. These two moments in time are shown in figure 6.20 and equal \( \frac{1}{4} \) and \( \frac{3}{4} \) of the period the contra schedule is present, respectively.\(^\text{14}\) Supposing that this estimated time, either for system entry or system exit, equals \( T \) and that the requested time, requested by the load unit, equals \( t \), then the distance in time equals:

\[
\text{Distance in time} = | t - T | + 0.01 \text{ [hr]}
\]  

(6.1)

in which a small constant value of 0.01 hour is added in order to prevent division by zero in formula (6.3) in case the estimated time equals the requested time.

The relative occupancy of the regular contra schedule relates the number of load units that already has been assigned to a contra schedule to its capacity, which is either specified as a number of generalised load units, or automatically determined on basis of the contra schedule's duration. The number of load units that is already been assigned to a contra schedule is determined for one direction only, i.e. either for loading or for unloading. The direction used depends on the flow direction of the load unit that is to be matched. Supposing that the number of load units already matched to a particular contra schedule \( cs \) for the right direction \( d \), depending on the flow direction of the load unit equals \( n_{cs,d} \), and that the contra schedule's total capacity equals \( c_{cs} \), which is independent of the direction, then the relative occupancy of the regular contra schedule would equal:

\[
\text{Relative occupancy} = \frac{n_{cs,d} + 0.1}{c_{cs}}
\]  

(6.2)

in which a constant value of 0.1 load units is added both to the numerator in order to prevent division by zero in formula (6.3) in those cases where no load units have been assigned to the regular contra schedule yet. As the number of already assigned load units \( n_{cs,d} \) is a natural number, the relative error of this addition is very small (less than 10%) in the worst case only one load units is assigned to the load schedule yet. In combination with the rather vague assignment of load units to contra schedules, this simplification is considered acceptable.

Now the score\(^\text{15}\) that is used as a basis for random selection of the best regular contra schedule to match the load unit with is found as:

\[
\text{Score} = \frac{1}{\text{Distnc. in time} \times \text{Relat. occupancy}} = \left( \frac{1}{| t - T | + 0.01 \text{ [hr]} \right) \left( \frac{c_{cs}}{n_{cs,d} + 0.1} \right)
\]  

(6.3)

The settings to use for random selection are free to choose by the user of the model, each different group of settings resulting in different matching results. However, it seems to be reasonable to apply quadratic random selection in combination with a relatively low skip level of 20 to 50 percent. In this way a rather "weak" random

---

\(^{14}\)Especially for regular contra schedules modelling large sea-going vessels that may dock up to one or two full days it is important to distinguish between both flow directions of load units for the first, rough estimation of the system entry- or exit time.

\(^{15}\)Refer to Appendix A for more details on scores and random selection.
selection behaviour is obtained, resulting in a fair division of load units over all available contra schedules. This weak random selection behaviour is believed to be in correspondence with the real-life situation that is modelled, in which the actual rules for matching load units to vessels or trains are quite vague and hard to model. Therefore, a more constrained random selection behaviour, that searches a contra schedule in a more narrow band, could suggest an accuracy and resoluteness that in reality do not exist.

Note that this way of matching does not guarantee at all that the actual number of load units that is matched to a regular contra schedule is in correspondence with its specified capacity. This is a consequence of the limited time range in which a matching contra schedule must be found, and of the inclusion of the distance in time between the requested system entry- or exit time, and the time offered by the contra schedule in the score of the contra schedule. However, as there is no experience yet with the actual behaviour of this matching algorithm during simulation, it may turn out that more likely results are found when the relative occupancy is included as a quadratic in the total score. In that case formula (6.3) would turn into:

\[
Score = \frac{1}{\text{Distnc. in time} \cdot \text{Relat. occupancy}} = \left(\frac{1}{|t - t_r| + 0.01 \text{ hr}}\right) \left(\frac{c_{ts}}{n_{ts, d} + 0.1}\right)^2
\]

(6.4)

### 6.4.2. Intensity contra schedules

Matching flows of load units to intensity contra schedules is quite analogous to matching them to regular contra schedules: The same structure as found at regular contra schedules, using random selection, distance in time and relative occupancy is used. Nevertheless, there are some differences.

**First selection of intensity contra schedules**

First of all, each one-hour period of a regular contra schedule is to be considered as a single, separate regular contra schedule, lasting exactly one hour and having a capacity that equals the specified intensity of the one-hour period\(^\text{16}\). As each one-hour period covers precisely one unique hour-period, and each day contains exactly 24 of these one-hour periods, a kind of "fence" consisting of 24 regular contra schedules per day covering exactly all 24 hours of each day is obtained. All 24 "boards" in the "fence" fit perfectly, resulting in a "fence" without any gap, as shown in figure 6.21. Based on the system entry time of the load unit to match (in case of a push flow) and the boundaries of the random distribution that specifies the duration the load unit is to reside in the system, it can be found which one-hour periods are covered and are interesting to match the load unit to as shown in figure 6.21 (compare with figure 6.18). A one-hour period is only then selected as interesting to match the load unit with, if the mid of that period, i.e. its centre, lies within the found boundaries.

**Second selection of intensity contra schedules**

Second, there is a need to establish the relative occupancy of each one-hour period that is selected as interesting to match the load unit with. As we have seen at the regular contra schedules, this factor is determined as the fraction of the number of load units already assigned to that one-hour period (for the direction that is considered, so for import or export) divided by the capacity of that one-hour period. Compared to regular intensity contra schedules, a problem is encountered at this point: While all regular contra schedules had a capacity that was expressed either in numbers of generalised load units or in the time they reside at their

---

\(^{16}\)At least, at this point it seems to be. Later on it will be shown that some special "trick" must be played with these capacities in order to prevent scaling problems and to include the total day-importance of each weekday.
common bounding element, a unit of capacity that was equal for all regular contra schedules docking at the same bounding element, no such common unit of capacity exists for intensity contra schedules. Even more, there are some intentionally modelled differences in capacity between two subsequent weekdays (refer to figures 6.14 and 6.15) that should be respected, and there are some non-deliberately modelled differences in capacity hidden in the intensity contra schedules. This latter aspect will be clarified by a small example.

Consider the situation that the intensity contra schedules for two different weekdays (Wednesday and Thursday) are entirely equal, but happen to be modelled by different absolute values, as shown in table 6.5. This non-deliberate difference in capacity for different weekdays is to be left out of the matching process, while the intentional difference in importance per weekday as a whole as shown in figures 6.14 and 6.15 should be respected and included in the matching process.

<table>
<thead>
<tr>
<th>Period</th>
<th>Wednesday</th>
<th>Thursday</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute</td>
<td>Relative</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>9%</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>18%</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>55%</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>18%</td>
</tr>
<tr>
<td>Total</td>
<td>110</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 6.5: Same intensity schedule modelled by different absolute values.

So, the capacity of a one-hour period that is to be used in the matching process is determined as follows. Suppose that the sum of the user-specified intensities of all one-hour periods in one weekday equals $n_{day}$ for the flow direction of the load unit. Refer to the second and fourth columns of table 6.5 for some examples of these user-specified intensities. Also suppose that the importance of that weekday as a whole equals $I_{day}$. Refer to figures 6.14 and 6.15 for some examples of this total day-importance, for a Monday and a Sunday respectively. Then the capacity $c_i,match$ of a certain one-hour period $i$ that is to be used in the matching process and was originally modelled to equal $c_i^{17}$, equals:

\[
c_i,match = \frac{c_i}{n_{day}}
\]  

(6.5)

An example will be given by discussing figure 6.22 which shows four one-hour periods that have been found suitable to match a particular load unit with; this is the result of the first selection criterion. Therefore, the centres of all four time periods must lie within the range for system exit or entry time of the load unit, as shown in figure 6.21.

It appears that these four time periods cover two different days: The last two one-hour periods of a Sunday and the first two periods of a Monday. The total day-importance is likely to differ between these two days, and indeed it does: The total day-importance of Sunday appears to equal 2 and that of Monday 10. So, $n_{Sun}$ appears to equal 2 and $n_{Mon}$ appears to equal 10. In that case it is obvious that the actual capacities of all four one-hour periods that will be used in the matching process will be somewhat different than the capacities shown in figure 6.22 might suggest.

\[17\] Refer to both columns "Absolute" in table 6.5 for some examples of this capacity.
If we suppose that the sum of all capacities at Sunday equals 225 and is called $n_{Sun}$ and that the sum of all capacities at Monday equals 176 and is called $n_{Mon}$, then the capacity of each of the four one-hour periods that is to be used when matching the load unit can be determined according to formula (6.5) as shown in Table 6.6. Note the major impact of the differences in the total day-importance $I_{day}$ of Mondays and Sundays on the capacity $c_{i,match}$ that is to be used in the matching process.

<table>
<thead>
<tr>
<th>Period</th>
<th>$c_i$</th>
<th>$n_{day}$</th>
<th>$I_{day}$</th>
<th>$c_{i,match}$</th>
<th>Relative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sun-23</td>
<td>10</td>
<td>225</td>
<td>2</td>
<td>0.0889</td>
<td>5%</td>
</tr>
<tr>
<td>Sun-24</td>
<td>25</td>
<td>225</td>
<td>2</td>
<td>0.2222</td>
<td>12%</td>
</tr>
<tr>
<td>Mon-01</td>
<td>15</td>
<td>176</td>
<td>10</td>
<td>0.8523</td>
<td>46%</td>
</tr>
<tr>
<td>Mon-02</td>
<td>12</td>
<td>176</td>
<td>10</td>
<td>0.6818</td>
<td>37%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 6.6: Matching capacities for the four one-hour periods from figure 6.22.

Now that the capacity to use in the matching process for each one-hour period within the range as shown in figure 6.21 is found, the single one-hour period to match the load unit with can be determined. Completely analogous to the method used at the regular contra schedules, this actual match is based on random selection and on the scores of each one-hour period according to formula (6.3) or (6.4). Important to note is, however, that the estimated time $\bar{t}$ for system entry or exit for each one-hour period as used in formula (6.1) is set to the centre of the one-hour period, as is shown in figure 6.21.

| Period  | $\bar{t}$ | $|\bar{t} - t|$ | $n_{cs,d}$ | $c_{i,match}$ | Score$^9$ | Chance |
|---------|-----------|----------------|-----------|---------------|-----------|--------|
| Sun-23  | Sun-22:30 | 01:45          | 18        | 0.0889        | 3.4       | 0.0%   |
| Sun-24  | Sun-23:30 | 00:45          | 9         | 0.2222        | 53.1      | 30.7%  |
| Mon-01  | Mon-00:30 | 00:15          | 51        | 0.8523        | 64.1      | 50.0%  |
| Mon-02  | Mon-01:30 | 01:15          | 12        | 0.6818        | 44.7      | 19.3%  |
| Total   |           |                |           |               | 100.0%    |        |

Table 6.7: Scores and chances of occurrence of all four time periods of figure 6.22 and Table 6.6 in case of quadratic random selection using a skip level of 20% in combination with the zero-based option.

Suppose, for example, that the load unit that is to be matched is requested to leave the system on Monday at 00:15 hours. In that case $t$ equals Mon-00:15 and the time distance of each of the four one-hour periods is determined as the absolute difference between this requested exit time and the period’s centre. These time periods are shown in Table 6.7. Supposing that the numbers of load units $n_{cs,d}$ that already have been assigned to these four time periods are as shown in Table 6.7 too, and that the capacities $c_{i,match}$ are as found at Table 6.6, then we can find the scores per period by using formula (6.3), in which $c_{cs}$ is replaced with $c_{i,match}$. These results are displayed in the one but rightmost column of Table 6.7. If we suppose that the right time period is selected by quadratic random selection using a skip level of 20% in combination with the zero-based option$^{20}$ then the chances for each period to be selected are as shown in the rightmost column of Table 6.7.

When matching intensity contra schedules to load units one special case could occur: The case that not a single one-hour period is found within the allowed boundaries that has a non-zero capacity $c_i$. In that particular case a "trick" is played completely analogous to the one played at the regular contra schedules: The user of the terminal design aiding tool is requested to specify a wider range to search a valid one-hour period in. This process is repeated until at least one valid period is found.

Completely analogous to regular contra schedules, this way of matching does not guarantee at all that the actual number of load units that is matched to an intensity contra schedule is in correspondence with its specified pattern. This is a consequence of the limited time range in which a matching one-hour period must be found, and of the inclusion of the distance in time between the requested system entry- or exit time, and the time offered by the contra schedule in the score of the contra schedule. Analogous to our discussion at the regular

\(^{18}\) Periods in this table are coded by (1) their weekday and (2) their one-hour period within that weekday. So, period "Sun-23" represents the twenty-third one-hour period of a Sunday, which is the one-but-last one-hour period of that weekday.

\(^{19}\) Multiplied by 1,000.

\(^{20}\) Refer to Appendix A.
contra schedules, it may turn out that more likely results of the matching algorithm are found when the relative occupancy is included as a quadrate in the total score. In that case formula (6.4) would be used instead of (6.3).

### 6.4.3. Timing of load units

Until now timing of flows of load units across the system's boundary has been discussed in a rather rough way: It has been shown to which regular contra schedule, or to which one-hour period of an intensity contra schedule a load unit is to be matched, but it has not been shown at what precise time that load unit is to enter or leave the system. On the other hand, it has been suggested a few times that the time at which a load unit is to enter or leave the system according to its load schedule is known in a rather precise way -- even though it has not been shown yet what that precise way is.

This section describes timing of load units passing the system's boundary in a detailed way; it will turn out that there is a difference between timing of load units with respect to their load schedule, and timing with respect to their contra schedule. This difference is mainly caused by the more accurate way load schedules are defined -- and are intended to be used.

In total four parts can be discerned in this section, each part describing a different type of timing. The four types of timing, each one corresponding to a different part, are:

- timing of load units relative to their load schedule, when defined at a bounding element;
- timing of load units relative to their load schedule, when defined at a train handle area;
- timing of load units relative to their regular contra schedule, always defined at a bounding element;
- timing of load units relative to their intensity contra schedule, always defined at a bounding element.

From the description of the first two types it can be deduced that load schedules defined at bounding elements are handled somewhat different than load schedules defined at train handle areas. In fact, this conclusion is true for SDM and DDM level of detail only, which is a consequence of the previously predicted modelling of trains at these two levels by means of so-called train services rather than by means of plain load schedules. In contrast to load schedules, these train services do have a fixed system exit time (which is a consequence to some common operational rules that apply for railways) resulting in a slightly different timing behaviour than found at plain load schedules.

**Part I: Timing relative to load schedules - at bounding elements**

Timing of load units relative to their load schedules, when defined at bounding elements, requires little discussion. As discussed at figure 6.5, which is repeated in figure 6.23, load volumes are the model elements at which the actual (random) numbers of load units that are to enter or leave the system are specified, along with additional data like the load units' client, closed task and direction of flow, and also along with information on timing of the entire load volume, relative to its encapsulating load schedule. This latter type of information is modelled by means of the left- and right offset of each load volume, relative to its load schedule.

Now, precise timing of load units relative to their load schedules is quite simple, and will be discussed in three steps. **First of all** the disturbed system entry time and duration of the total load schedule are determined, automatically resulting in the system exit time of the load schedule. **Second**, the left- and right offset of each load volume in the load schedule are determined, which are mere fractions of the total load schedule's duration than absolute offsets. **Third**, the number of load units registered at each load volume is determined and spread evenly over the entire load volume's duration. Note that all these load units, stemming from the same load volume, are totally equal: They are of the same type and of the same client, they are to flow in the same direction to the same opposite location, by using the same closed task. There is, however, one slight difference between these individual load units at this point: The precise time they are to enter or leave the system at their load schedule.
Table 6.8:  Example calculation for precise timing of push-flow load units stemming from one load volume (___ = source data, □□□ = derived data).

Table 6.8 shows an example calculation for load units stemming from a push-direction load volume. This load volume is timed, relative to its load schedule, as is shown in the upper row of table 6.3. All load units stemming from this load volume are evenly spread over the load volume's duration. Note especially that each first load unit in a load volume is timed at the very beginning of that load volume, while the last load unit in a load volume (provided that at least two load units originate from that load volume) is timed at the very end of the load volume; all other load units are spread equally in between.

- **Part II: Timing relative to load schedules - at train handle areas**

  The second type of timing is timing of load units relative to their load schedules, when defined at train handle areas. This type of timing is somewhat different for UCM level of detail, compared to SDM and DDM level: At UCM level of detail timing of load units, relative to their load schedules when defined at train handle areas is exactly the same as when they were defined at bounding elements. Therefore, at UCM level of detail the previously discussed method for bounding elements can be used.

  At SDM and DDM level of detail on the contrary, load schedules are encapsulated by so-called train services as will be discussed in more detail in Chapter 9. This encapsulation of load schedules by train services has some major consequences for timing of load units: Load units that are transported by trains are timed totally different at SDM and DDM level of detail, that at UCM level of detail where actual "trains" as used at SDM and DDM level of detail do not exist. As a matter of fact, all data specified at the load schedules and load volumes can be ignored if desired at SDM and DDM level of detail. Refer to Chapter 9 for a detailed description hereof.

- **Part III: Timing relative to regular contra schedules**

  Now it is known at what times load units enter or leave the system according to their respective load schedules. Based on these times and on the (random) duration of their respective equipment routings they are to flow over, it can be found at what contra schedule these load units are registered at the opposite side of the terminal system, as discussed earlier in this chapter. This results either in a particular regular contra schedule or in a particular one-hour period of an intensity contra schedule.

  In case of a regular contra schedule, all load units that have been registered at that regular contra schedule are spread evenly over that contra schedule's duration, such that the total deviation of all load units' requested system entry or exit time (refer to figure 6.17) from their granted system entry or exit time is minimised; this granted system entry or exit time is determined by the order in which all these individual load units are loaded or unloaded at the regular contra schedule.

  The procedure to minimise this deviation can be divided into two separate steps: In the first step all load units are sorted into two lists, one list for load units to unload and one list for load units to load at the regular contra schedule. In the second step all load units are spread over the available time for loading or unloading, such that

  21 Because of their total resemblance, any of the load units originating from one load volume can be chosen to be the "first", the "last", or whatever.
the total deviation between requested and granted times is minimal. Both lists of load units, i.e. the list for load units to unload and the one for load units to load, are handled separately. This process will be illustrated by figure 6.24.

In figure 6.24 an "available period" is shown, together with a number of load units (named "A" through "F") that need to be timed somewhere in that period. As all load units are to be spread evenly over the available period, in total six different positions in time are discerned in the available period: One at the very beginning of the period, one at the very end of the period, and four others evenly spread over the period as indicated by the small vertical ticks. Note that this procedure appears to presume that the total number of load units to spread over the period is known, as well as the precise starting and ending time of that period. In fact, this is a basic assumption that is always satisfied as will be shown shortly.

The procedure to assign load units to the various time positions in the available period is as follows: Only those load units whose requested time lies outside the available period are actually assigned; all others will be assigned later on by repeating precisely the same assignment procedure, using a different time period as will be discussed shortly. Those lying outside the available period are assigned in two assignment sequences or directions, as shown in figure 6.24 as the "first" and "second" direction. The "first" direction assigns the load units that have a requested time before the start of the period to the leftmost available time periods, locating the one with the smallest requested time at the very beginning of the available period (so, at the very first time position), and next to that one the one with the one but smallest requested time, and so on, until all load units that have a requested time before the start of the period are assigned. The "second" direction is completely mirrored with respect to the "first" direction.

The result of the assignment of the load units having a requested time that lies outside the available period as shown in figure 6.24 is shown in figure 6.25. This latter figure also shows in the lower half a "new" available period and the two load units that have not been assigned yet. This "new" available period simply consists of the time positions that have not been used during the first time the matching algorithm was executed. Now the remaining two load units can be assigned in exactly the same way as in the first assignment round. If necessary, this assignment procedure can be repeated a number of times, until all load units have been assigned.

The total assignment procedure is executed twice for each regular contra schedule: Once for the load units that are to be unloaded at the contra schedule, and once for the load units that are to be loaded at the contra schedule. The size of the "available period" from figure 6.24 depends on the number of load units that is to be handled in load- or unload direction, and on the total duration of the contra schedule. If, for instance, 20% of the load units registered at the contra schedule are to be unloaded, then the first 20% of that contra schedule's duration is reserved for unloading of load units; the remaining 80% of the contra schedule's duration is then reserved for loading of the remaining 80% of the registered load units.
It is important to realise that, in order to execute this assignment procedure properly, it must be known precisely which, and therefore how many, load units are to be unloaded or loaded at a regular contra schedule. Otherwise, the results of the assignment procedure would change each time a new load unit is known to be handled at the contra schedule, resulting in changes in timing of all or some load units that have been registered before at the contra schedule. In order to avoid this effect as far as possible, both for push and for pull flows, some special demands must be imposed on the simulation model. However, it will turn out that this effect can not be eliminated totally, which actually is in correspondence with the idea of a fixed knowledge horizon for all flows of load units. In this way uncertainty in information is included in the model in a rather rough way.

**Demands for the simulation model**

In Chapter 7 it will be discussed that the simulation model that is used to analyse the model includes a so-called knowledge horizon which plays an important role in the planning of flows of load units: All load units that enter the system before the planning horizon are known to the simulation model, and all load units that enter the system after that horizon are not. Is seems as if this aspect has some impact on load schedules and contra schedules, too.

Consider figure 6.26 in which the current simulation time is shown along with the (fixed) knowledge horizon that has been specified by the user of the model. All load units entering the system before that horizon are known to the system, all other load units are not. Figure 6.26 also shows for some particular load unit the time of system entry and of system exit, the first being defined by its load schedule. By consequence, the load unit appears to flow in push direction. From these data we can derive that the load schedule from which the load unit originates must start before the knowledge horizon. This demand can be satisfied easily by creating during simulation all load schedules whose undisturbed system entry time is less than the current knowledge horizon. As soon as the disturbed system entry time of these load schedules is less than the current knowledge horizon, too, then all load units that created and are registered at that load schedule can be matched to contra schedules and thus can be timed at their respective contra schedules. By consequence, all load units that already had been registered and timed at these contra schedules, and were created by other load schedules, will receive a change in timing.

While figure 6.26 showed a push flow, figure 6.27 shows a pull flow; instead of a flow running from a load schedule to a contra schedule, figure 6.27 shows a flow running from a contra schedule to a load schedule. This has some major effects on the timing of load units and on the time the simulation model has to "look in advance" in order to create load schedules in proper time. Note, by the way, that in figure 6.27 the system exit time of the load unit is determined by the load schedule and the system entry time by the contra schedule (and therefore is marked as "granted"), while in figure 6.26 this situation is completely reverse.
Figure 6.28: Determining the additional time span between load schedules and contra schedules in case of pull flows.

In order to safeguard that load units are registered at their respective contra schedules at the time they are actually to enter the system, load schedules need to be created some time in advance, in advance to the current knowledge horizon. The amount of time these load schedules are to be created in advance, relative to the current simulation time plus the knowledge horizon, is determined by the (random) duration of all equipment routings to which the load schedule is related. As each load schedule contains a known number of load volumes, each one referring to precisely one closed task, and as each closed task results in two equipment routings, it is simple to find the maximum value for this additional time, over all equipment routings. This procedure is shown in figure 6.28 in which a particular load schedule contains load volumes for three different closed tasks, each possibly yielding a different maximum duration for transport between this load schedule and the contra schedules.

After each load schedule has been created in advance, its load units can be created too in advance, and can be registered at their respective contra schedules. However, if this would be done immediately at the time the load schedule is created (which may last up to 24 hours) then a number of load units is created too early, thus cheating on the requested knowledge horizon. This especially hold for load units that are to be created at the very end of the load schedule's duration. Therefore, load units for pull flows are not created immediately after their load schedule has been created, but only when their load volume needs to be created. This moment in time can be determined easily at the same way as just described for an entire load schedule, with one difference: Instead of the maximum duration of all equipment routings related to all load volumes that are in the load schedule, as shown in figure 6.28, the maximum duration of the single equipment routing that is related to the single load volume that is to be created, is used.

Note that the uncertainty in information on flows of load units, which results in changes in timing of load units, is only present at contra schedules that start after, or very close to, the current knowledge horizon. As this knowledge horizon will under all circumstances be at least a few hours, the effect on flows of load units close to the current simulation time is limited.

As previously found, in some cases a problem might occur when matching load units to contra schedules: In some cases not a single contra schedule can be found within the range that is allowed by the (random) duration of a load unit's equipment routing. In such cases a matching contra schedule is searched outside this range, resulting in an earlier system entry time at a contra schedule than requested. Even more, whenever a valid contra schedule is found within the valid range, then still in some odd cases the eventually granted system entry time at that contra schedule may be located before the start of the range. As shown in figure 6.29 this is the case when the first, rough estimation of the load unit's system entry time does lie in the required range, but the eventually granted system entry time does not. This special case is mainly found at contra schedules with a long duration, such as contra schedules modelling large sea-going vessels.

Figure 6.29: Special case of erroneous matching of a load unit to a contra schedule.
Chapter 6. Load schedules and contra schedules

♦ **Drawback**
In order to fix this problem some safety margin must be added to the maximum duration of all equipment routings as was found according to figure 6.28. This safety margin must be somewhat greater than the 24 hours previously mentioned at page 55 and is used both when creating load schedules in advance and when creating individual load units in advance. A drawback of this safety margin is that it violates the knowledge horizon that was specified by the user.

♦ **Fixing the drawback**
This drawback can be avoided by dividing the matching process of load units to contra schedules in two separate steps as previously indicated: The first step in the matching process searches a valid matching contra schedule, but does not yet register the load unit at that contra schedule; it is only remembered which contra schedule is to be selected in future. This first step is executed more in advance than the second step will be; the time this first step is executed in advance does include the mentioned safety margin.

The second step in the matching process determines the precise time at which the load unit is to enter the system at the contra schedule that was found in the first step, and actually registers the load unit at the contra schedule. This second step is executed somewhat later than the first step; at maximum a time span equal to the safety margin after the first step, depending on the contra schedule that has been found: If a valid contra schedule has been found in the normal range of the related equipment routing's duration, then a time span equal to the safety margin lies between the first and this second step. If a valid contra schedule has been found outside the normal range, then the second step is executed immediately after the first step.

These two steps are illustrated in figure 6.30. So, during simulation all load schedules that have a disturbed system entry time before the current simulation time plus the knowledge horizon plus the maximum duration of all equipment routings of that load schedule plus the just mentioned safety margin, are to be created in advance.

![Diagram](https://via.placeholder.com/150)

**Figure 6.30:** Separating matching to a contra schedule from registering at a contra schedule.

For each load volume that is registered at that load schedule the load units are created at most this amount of time in advance and are immediately matched to their contra schedules; but they are not registered yet at these contra schedules. The amount of time these load volumes are created in advance is less than the amount of time used for load schedules, as for all these load volumes it is known to which closed task and equipment routing they are related, and therefore the previously mentioned maximum duration of all equipment routings can be replaced by the maximum duration of the single equipment routing that is related to the load volume. After the safety margin has elapsed, the load units of the load volume are actually registered at their contra schedules. During registering, the actual precise time of system entry is determined of these load units.
Figure 6.31: Timing of creating load schedules, of creating load volumes, of matching load units to contra schedules, and of registering load units at contra schedules.

One of the load volumes in the load schedule is to start immediately after the load schedule has been created, as it has a zero-valued left offset and as it runs over the equipment routing that resulted in the maximum duration. All load units in this first load volume are matched immediately to their contra schedules. During this matching it turned out that all load units could find a valid contra schedule in the normal range of their equipment routing's duration, and therefore they are registered at these contra schedules only after the safety margin has elapsed.

The second load package also turned out to have a zero-valued left offset, but runs over a faster equipment routing. Therefore its load units are matched somewhat later to their contra schedules than those originating from the first load volume were matched. But, during this matching, it turned out that no valid contra schedule could be found in the normal range of the used equipment routing's duration, and therefore all load units are immediately registered at these contra schedules, too. There is no safety margin between matching and registering load units of the second load volume.

The third load volume appeared to start later, either because of a fast equipment routing or of a greater left offset, and did not encounter any problems in finding a valid contra schedule in the normal range of the used equipment routing's duration. Even this third load volume is to start at least an amount of time that equals the knowledge horizon before the current simulation time.

From this example it can be deduced that, whenever only one load unit from a load volume is matched to a contra schedule that lies outside the normal range of the used equipment routing's duration, then all other load units stemming from the same load volume are registered more in advance, too. This simplification is accepted to keep the model from growing too complex.
Part IV: Timing relative to intensity contra schedules
Timing of load units that are registered at intensity contra schedules is completely analogous to timing of load units that are registered at regular contra schedules. The only "difference" is that each one-hour period of an intensity contra schedule is to be considered as a single, individual regular contra schedule that lasts exactly one hour and merges perfectly with the fictitious regular contra schedule of its preceding and of its succeeding one-hour period. In this way a kind of "fence" of 24 small regular contra schedules a day is obtained, a "fence" without any gap between the "boards". Compare with figure 6.21 in which this approach is used, too.

6.5. Some examples of the use of load schedules and contra schedules

A number of situations can be modelled by means of load schedules and contra schedules. Some situations can be modelled quite satisfactory, but other situations can not. This is mainly a consequence of the first concern of the model described in this thesis: Modelling and analysing rail service centres.

In order to help understanding the possibilities and restrictions, a few examples will be given, each having some special characteristics. Three different situations are described: train-to-sea transport, sea-to-sea transport and sea-to-truck transport.

6.5.1. Train-to-sea transport

One of the most important types of transport encountered at the Maasvlakte, relative to the future rail service centre, is train-to-sea transport. This section gives an example of how such a situation could be modelled by means of load schedules and contra schedules.

Figure 6.32: Example of train-to-sea transport.
Figure 6.32 shows a possible future situation of the Maasvlakte rail service centre, related to marine terminal "Miersk", along with four different closed tasks as could be defined in this situation. Two of these closed tasks run over the stack of the rail service centre, and the other two do not\(^{22}\). Figure 6.32 also shows that from each of these two pairs of closed tasks one closed task is reserved for loaded containers, and the other is reserved for empty containers. In this way different characteristics for both types of load units can be specified.

All trains that are to be handled at the rail service centre are to be modelled by means of load schedules. During simulation at UCM level of detail these load schedules that are defined at a train handle area are used in precisely the same way as they are used at bounding elements. During simulation at SDM and DDM level on the other hand, these load schedules are encapsulated by so-called train services as described in Chapter 9.

All marine vessels docking at marine terminal "Miersk" are modelled by means of regular contra schedules; the numbers of load units that flow through the system are therefore not defined at this marine terminal, but only at the rail service centre. This is in correspondence with the aimed level of detail for train handle areas of the model.

### 6.5.2. Sea-to-sea transport

This section gives an example of modelling sea-to-sea transport. A problem that arises in this case of transport, i.e. transport between a number of regular services, is to choose which regular services are to be modelled as load schedules, and which are to be modelled as regular contra schedules. This problem will never arise in case of train-to-sea transport, as all trains that are defined at a train handle area are defined by load schedules; they can not be modelled by contra schedules. By consequence, all vessels that are to be related to train-to-sea transport must be modelled by regular contra schedules; it is not possible to model them by load schedules and then match them to train services -- which are modelled by load schedules, too.

Even though this problem might seem to be serious, it is not that serious because of the aim of the model described in this thesis: Modelling and analysis of rail terminals -- and not of marine terminals. Marine terminals and, by consequence, regular contra schedules were only introduced to be able to include the effects these marine terminals have on internal transport to and from the rail service centre in the model. They were not included for detailed analysis of marine terminals.

So, if we consider the situation for sea-to-sea transport as shown in figure 6.33, based on figures 6.1 and 6.12, then one or more closed tasks are defined between both marine terminals. Suppose that marine terminal "Nin-Jan" is selected to get all vessels modelled by means of load schedules, and that marine terminal "Miersk" by consequence is to get all vessels modelled by means of regular contra schedules, then any sea-to-sea flow can be modelled. Some of these contra schedules might also receive load units from a train, if modelled, but the load schedules may not.

If required, any mix of load schedules and load volumes at both marine terminal can be used. However, it may become a little "fuzzy" which load schedules are related to which contra schedules. This "fuzziness" can be reduced by creating different closed tasks and assigning just a limited number of ships to each of these closed tasks, as described at figure 6.12.

---

\(^{22}\) Each closed task defined in any system must run over at least one stacking location with unlimited stacking capacity. This yields the conclusion either that the rail service centre allows local stacking next to the handle bundles, or that marine terminal "Miersk" allows local stacking. Refer to Chapter 7 for more details on this subject.
Note that sea-to-sea transport at the same marine terminal can be modelled too, by playing the same "trick" as described at figure 6.16 in which two separate bounding elements were created for one real-life truck gate. These two bounding elements were then connected by some small closed tasks. Note that this kind of "short-loop" situations only needs to be modelled when it has effects on the remainder of the system. This is the case when (1) the equipment pools needed for internal transport at these locations are used for internal transport at other locations in the system too, and are expected to be a bottleneck, or (2) the limited stack capacities at these locations are expected to influence the remaining system.

6.5.3. Sea-to-truck transport

The last example concerns transport between a marine terminal and a truck gate. Obviously, transport at the side of the truck gate is to be modelled by means of intensity contra schedules, leaving no other choice for transport at the side of the marine terminal to be modelled by means of load schedules.

An example of this situation is given in figure 6.34. Note that this figure does not specify at all how load units are transhipped and transported by internal transport; it only shows how a marine terminal and a truck gate are related by means of one (or possibly more) closed tasks. Each of these closed tasks runs over the "private" stack of the marine terminal, too.

Apart from the load schedules and specifying the numbers of load units to be transported through the system, this some additional data is needed for this example. This data concerns the patterns of the arrival of the trucks, both for import and export direction. An example herewith, repeated from figures 6.14 and 6.15, is given in figure 6.35 which shows both the intensity patterns for Mondays and Sundays at the truck gate, together with the total number of trucks that is to arrive at average at those two weekdays.

![Diagram of sea-to-truck transport](image)

Figure 6.34: Example of sea-to-truck transport.

![Graph of intensity contra schedules for two different weekdays](image)

Figure 6.35: Example of intensity contra schedules for two different weekdays.

---

23Refer to Chapter 7.
Chapter 7. Equipment

7.1. Introduction

"...the air was heavy with the sound of containers bumping around. Sometimes I feared they would collide, but fortunately they appeared to be masters in hopping away just before being splattered by a fellow container. It was a beautiful sight, all those containers bumping and bouncing happily around at the terminal yard. I sat down in the sunset, on a soft lawn, and enjoyed the scene...". Here ends the dream of the terminal operator. In reality load units are no living creatures that are able to move around the terminal yard on their own; in reality they are huge and heavy things that need to be moved around by machines. These machines (or equipment) need to be controlled and directed such, that all containers arrive in time at the right location. This requires a lot of planning, especially in case of a large-scale terminal complex. That's why in the dream all containers could move of their own.

Now, what is the point of this introduction? This introduction is used to draw attention to the problem of internal transport, which is a very broad and important aspect of a terminal's design; and not of a rail terminal's design only.

♦ Definitions

Normally, internal transport is defined as all transport from one location at the terminal yard to another location at the terminal yard by means of transport equipment that is owned by the terminal. By consequence, internal transport is all transport that does not cross the system's boundary. Although this is a commonly accepted definition of internal transport, in this thesis a broader definition of internal transport is adopted: Internal transport is defined as the total of all activities that are required to transport load units from one location at the terminal yard to the other. This definition does not only include all pure transport movements, but also all transshipment movements, all stacking movements, all equipment relocation movements, and all planning of all these movements. This broad, but complex definition opens the path towards an integrated approach of the problem of internal transport at a large-scale terminal complex, including most of the activities that occur in real-life internal transport.

The aim of this chapter is twofold. First, this chapter gives some elements that can be used to model almost any situation of internal transport at a large-scale terminal complex, both at SDM and DDM level; there is no need for such elements at UCM level due to its coarse level of detail. Of course, it is not possible to tackle all situations that might occur in the real-life system in a general model as described in this thesis, as the variety in organisational forms of internal transport is almost infinite. Therefore, the model must be detailed enough to tackle most of the situations at a reasonable level of detail. Note that this definition is rather subjective. Second, this chapter gives a model structure to be used at SDM level that optimises, and gives advice on, the number of equipment to install at various locations at the terminal complex. This model should take into account complex routing structures, knowledge horizons, various working hours, required flow patterns and overall costs. This model will be elaborated in detail, and will be used during simulation at SDM level.

The aim of this chapter is not to give a method to find automatically the best organisational form of internal transport at a large-scale terminal complex. As stated in Chapter 5, this question is too complex to tackle at the moment. The method described in this section only optimises a structure of internal transport activities, once that structure is given.

In the remainder of this chapter following definitions will be used:

- system = real-life system that is to be modelled and analysed
- model = description of the real-life system (general)
- model = description of the real-life system (at a specific level of detail)
- optimisation model = model to determine to optimal long-term assignment of equipment for internal transport to transport jobs
- user = person converting the system into a model
- machine = single piece of equipment.
7.2. Elements for internal transport

Defining a general model for internal transport at a large-scale terminal complex is a rather challenging task. Such a model should cover a great number of aspects as found in the real-life system at a reasonable level of detail. Otherwise, the model would not be valid nor of great use. When considering the real-life world a number of demands that are imposed on a model for internal transport can be found, which are:

1. the model should include flexible definition of equipment types. This property is needed to model the great variety of equipment types that can be found in the real-life world. Some examples of equipment types are rail mounted gantry cranes (RMGs), straddle carriers (SCs), automated guided vehicles (AGVs) and multi-trailer systems (MTSs). Refer to Chapter 2 for a picture of AGVs and SCs.

2. the model should include a number of essential operational aspects like limited service ranges, limited working hours and preferences for use of equipment on specific transport jobs.

Following sections discuss these aspects in more detail and show how they are included in the proposed terminal design method.

7.2.1. Equipment types

It has been indicated just before that a general model for internal transport should include a flexible definition of equipment types, in order to model the large variety of these types in the real-life world. It appears that a limited number of parameters is sufficient to define almost any type of equipment at SDM level of detail. The equipment type definition given in this section will be used for DDM level of detail too.

These parameters can easily be found when considering some of the real-life equipment types. Consider, for instance, following list of equipment types:\footnote{Refer to Chapter 2 for a description of these equipment types.}

- fixed gantry cranes (FGC)
- rail mounted gantry cranes (RMG)
- rubber tyred gantry cranes (RTG)
- automated guided vehicles (AGV)
- terminal trailers (TT)
- multi-trailer systems (MTS)
- reach stackers (RS)
- straddle carriers (SC).

Based on the above list of real-life equipment types a limited number of parameters that define these different equipment types can be found. These parameters are:

- can move ?
- is self-propelled ?
- can pull ?
- can carry ?
- can tranship ?

All of these parameters are included in the model and will be discussed one by one. Not all combinations of abilities are allowed, either for reasons that are derived from the real-life world or for reasons that are imposed by the equipment optimisation model that is to be discussed after a while.

\# Equipment type parameters - qualitative

Most equipment types found in real-life terminal systems are able to move; only a few exceptions are found in real-life, like fixed gantry cranes. For equipment types that can move there is a need to specify the locations that can be served by that equipment, or, in other terms their service ranges. This aspect will be discussed in more detail at the so-called equipment networks.

Some movable equipment types are self-propelled and others are not, which is the main difference between trailer-type equipment and self-driven equipment. Some equipment types that are self-propelled are able to pull other, non-self-propelled movable equipment. In the model discussed in this thesis each trailer-type of equipment is to be associated with one pulling type of equipment. An example hereof is the association of MTS-trailers with MTS-trucks. Therefore, an MTS-trailer can only be pulled by an MTS-truck, while an MTS-truck...
may be able to pull several trailer-types of equipment. Any set of pulling equipment and trailer-type equipment that is to remain coupled always should be modelled as one single equipment type that incorporates both the pulling equipment and the trailer-type equipment.

Equipment that can move may also be able to carry load units and may thus be used for actual internal transport movements. The model described here does not allow pulling equipment to carry or to transship load units, although in the real world such a kind of equipment might be developed in future. This restriction has been included in the model to obtain a well-defined equipment optimisation model concerning pulling equipment that is not unnecessarily complex. This restriction is not expected to be a great obstacle in modelling any real-life system.

The ability to transship load units is independent of the setting of almost all other parameters. Only equipment types that are allowed to pull are not allowed to transship load units, which is a consequence of the chosen structure of the equipment optimisation model. In reality it will hardly ever occur that such an equipment type is used.

The aspect of operators controlling the machines, although very important in real-life indeed, is not included in the equipment type definition. The single reason for this is that the model always assumes that operators are available and at their position to control the machines. Nevertheless, operators are discerned when analysing the results of the simulation of the internal transport system for two reasons: First to include labour costs in the analysis results, and second to get a rough insight in the number of required operators.

Figure 7.1 gives some examples of real-life equipment that can be defined by the parameters just discussed. It is good to note that not all combinations of parameters are currently used in practice and may give rise to some new ideas about new equipment types. Furthermore, some combinations of parameters are crossed out as they either conflict with one of the previously given rules of result in a rather impractical type of equipment. The latter is the case for the rightmost columns of figure 7.1 showing an equipment type than can neither move nor transship.

![Equipment Type Definition Diagram]

**Figure 7.1:** Examples of equipment type definition.

A special role is played by gantry cranes installed (1) over handle bundles at train handle areas or (2) at stack blocks of internal stacks. As will be explained in more detail when discussing the equipment optimisation model these two kinds of gantry cranes are modelled as non-movable equipment instead of normal movable equipment as might be expected. This special arrangement has been made for the implementation of the here described terminal design method, allowing it to find those pools of gantry cranes that are to be installed over rail bundles or at stack blocks. For the moment this aspect will be given no further attention.
Equipment type parameters - quantitative
Of course the above parameters only define the ability of equipment to perform certain tasks. Several additional parameters are required to quantify these abilities. For instance, the ability to be self-propelled induces following four additional parameters (for coding and indexing of variables refer to Appendix D):

- \( et_{uv} \) = driving speed (unloaded, average) [m/s]
- \( et_{lv} \) = driving speed (loaded, average) [m/s]
- \( et_{uat} \) = acceleration time (unloaded, average) [s]
- \( et_{lat} \) = acceleration time (loaded, average) [s]
- \( et_{udt} \) = deceleration time (unloaded, average) [s]
- \( et_{ldt} \) = deceleration time (loaded, average) [s]

The letter \( l \) or \( u \) in the variable name indicates the type of transport movement: \( l \) indicates a loaded transport movement which means that the machine is pulling a non-self-propelled piece of equipment or is carrying load units. In all other cases the letter \( u \) is used, indicating an unloaded transport movement.

The average driving speed is the nominal driving speed including curves in the tracks, but excluding acceleration and deceleration from the loading- or unloading position. The average acceleration time is the time needed to reach the average driving speed, or to decelerate from the average driving speed. The parameters \( et_{uv} \) and \( et_{lv} \) are an abbreviation of equipment type velocity; the parameters \( et_{uat} \) and \( et_{lat} \) are an abbreviation of equipment type acceleration time; the parameters \( et_{udt} \) and \( et_{ldt} \) are an abbreviation of equipment type deceleration time.

The ability to pull induces the need to specify which other types of equipment can be pulled, along with the time needed to couple to and uncouple from these types of equipment. These aspects are modelled as follows: Each movable, non-self-propelled equipment type refers to exactly one movable, self-propelled equipment type that may be used to pull it. In other terms, each trailer-type of equipment refers to exactly one pulling type that may be used to pull it. This referred pulling type is stored in the parameter \( et_{ptr} \) of the trailer-type equipment. The time needed to couple to and uncouple from these types of equipment are the same for all these equipment types. These values are stored in following two parameters:

- \( et\_{cpt} \) = time needed to couple to a machine [s]
- \( et\_{uct} \) = time needed to uncouple to a machine [s]

The ability to carry induces following parameter specifying the capacity in so-called generalised load units:

\[ et\_{ce} \] = carrying capacity in generalised load units [gen-LU]

Generalised load units are defined as a kind of imaginary average of all load units that are considered at a specific point. Generalised load units do not exist in the real-life world, just as if "the average man" does not exist. If, for instance, an MTS-trailer can be loaded either with one 20' container, two 20' containers or one 40' container, and each of these load schemes would have equal chance of occurrence, then its capacity in generalised load units might equal 1. The parameter \( et\_{ce} \) is an abbreviation of equipment type carrying capacity. Most equipment types that are able to carry load units have a carrying capacity of one. Some good examples hereof are straddle carriers and fork lift trucks. Some equipment types, however, have a carrying capacity that exceeds one and therefore are called multiple load unit equipment types, abbreviated to MLU-equipment types. Some examples hereof are the Multi Trailer System (MTS) and double stacking trailers.

Generalised load units are used to express capacities of equipment in numbers of load units in an somewhat abstract way, abstaining from complex load schemes. So, instead of specifying a great number of varying load schemes for some type of equipment (as in the above example) it is sufficient to specify just one single number which expresses the capacity of equipment in a rough way. The disadvantage of this approach is reduction of detail (which level of detail is not required at all at SDM level). The advantage of this approach is a much more clear and easy formulation of the problem, resulting in a model that can easily be handled and understood.

The ability to transship load units induces four parameters specifying the time required for several steps in the transshipment process. The first step in the transshipment process is preparation for transshipment. This includes time needed to position the transshipment equipment relative to the load unit, and time to attach the transshipment equipment to the load unit. The second step in the transshipment process is actual lifting of the load unit. The third step in the transshipment process is putting down the load unit at its destination; the time needed to move to
this destination is not considered as a part of the transshipment process in this model. The fourth step in the transshipment process is releasing the load unit and any other post-transshipment activity. These four parameters are defined as follows:

\[
\begin{align*}
et_{at} &= \text{time to attach} & [s] \\
et_{ut} &= \text{time to get up} & [s] \\
et_{dt} &= \text{time to put down} & [s] \\
et_{rt} &= \text{time to release} & [s]
\end{align*}
\]

In the real-life world these parameters have following meaning. Consider, for instance, a container that is transshipped reach stacker between a stack and a truck. In that case the reach stacker would first drive to the stack position of the container (not included in the transshipment process), then it would position at the stack in front of the container, lower its spreader and attach the spreader to the container (et_at). Then the reach stacker would lift the container (et_ut), it would move to truck (not included in the transshipment process), and it would position in front of the truck and put the container on the truck (et_dt). Then it would release the spreader from the container, lift the spreader, and drive a little backward (et_rt).

Of course these four parameters could have been combined into one single parameter. The reason that this option has not been chosen for is twofold. First, the current definitions put more stress on the various stages of the transshipment process thus clarifying the difference between various alternative types of equipment. Second, the current definitions allow the use of four different random distributions, which is more accurate than condensing these four random distributions into one single distribution.

### 7.2.2. Equipment pools

Now that parameters defining different equipment types are known it is time to focus on the organisational aspects of internal transport. These organisational aspects are modelled by introducing three more elements, being equipment pools, equipment networks and equipment routings. This section will treat the equipment pools.

Equipment pools as defined in this thesis are very much like equipment pools found in the real-life world. In fondo, equipment pools are simply collections of equipment of one type that functionally need to be considered as a whole. This rather cryptic definition gives the user the freedom to group equipment such as he thinks needed to model a complex equipment assignment structure. This will be clear after having learned about equipment networks and equipment routings, where the user is allowed to specify which machines (i.e. which equipment pool) can be used for certain transport movements, at a certain level of preference. From now on any single piece of equipment in a pool is called a machine. Therefore, all machines in a certain pool are of one single type.

Equipment pools are not solely defined by the equipment type of all its machines. Equipment pools also contain information on working hours, fixed valuations, pool capacities and operators. These parameters will be discussed one by one.

**Working hours** are defined in a quite straightforward way: Working hours simply are time schedules specifying the availability of an equipment pool in time. Actually, there are no restrictions at all in specifying these time schedules, although the user should guard that they represent the actual situation. Once a set of working hours for each weekday is specified and stored in a so-called work schedule, it can be assigned to any equipment pool. Specifying working hours is facilitated a lot if all working hours can easily be set equal for all weekdays.

**Fixed valuations** are discussed in more detail at the equipment routings. For now it is sufficient to know that, if an equipment pool is given a fixed valuation, it will always have the same preference to be used for any transport movement, i.e. at any equipment network, at any section of any equipment routing. All these elements will be discussed later on.

**Pool capacities** are simply the number of machines that is installed at an equipment pool. In some cases it may be desirable to specify a fixed, limited number of machines, but in most cases it will be more convenient to leave the precise pool capacity open to discussion and let the equipment optimisation model determine during simulation the best number of machines to install at a pool. In that case the pool capacity is said to be unlimited. For each equipment pool an insertion point must be defined which is the location at which newly created
equipment may be inserted during simulation whenever required. After the so-called equipment networks have been discussed, it will be clear why this equipment insertion point must be an UCM location that is connected to any equipment network at which the equipment pool is installed.

Operators, as stated before, are solely modelled for purposes of cost analysis of the system. Therefore, any freely definable type of operator can be assigned to any equipment pool. At model analysis costs can be associated to each operator type, thus giving an impression of the number of operators required at the system along with the costs they involve.

### 7.2.3. Equipment networks

Next to discussing equipment types and equipment pools, the service ranges of these equipment pools should be discussed. In other terms, it should be discussed which locations at UCM level as discussed at Chapter 5 can be served by which equipment pools. This is done by introducing **equipment networks**.

*Equipment networks* relate equipment pools to locations and closed tasks at UCM level. Equipment networks, which are known from SDM level on, are strongly related to locations defined at UCM level. This strong relation is used to determine automatically which locations can be served by which equipment pools.

Equipment networks consist of two elements: **network pins** and **network links**. Network pins are simply the “nodes” in an equipment network between which the network links are spanned. Equipment (i.e. machines) can drive over these network links from one location to the other, or even within a location like within an internal stack.

An example of an equipment network is given in figure 7.2. This figure is based on an example previously given in Chapter 5. The links of the equipment network are shown in thick lines, along with the pins of the equipment network. Two different types of network pins are shown in this figure: The first type is shown as a circle, and is freely locatable at any position; this is the normal type of network pin. The second type is shown as a circle with a small triangle, and is attached (pinned) to a location at UCM level; this is the pinned type of network pin. A pinned pin is not freely locatable at any position, but its position is determined by the location at UCM level to which the network pin is pinned. Through these pinned pins it is possible to relate an equipment network to the underlying locations at UCM level, and, by consequence, to the closed tasks running over these locations. This aspect will turn out to be crucial when discussing the so-called **equipment routings**.

![Figure 7.2: Example of an equipment network](image)

Of course it must be possible to transport load units over equipment networks through equipment pools, and therefore it is possible to assign any equipment pool to any equipment network. Thereafter these pools are available for internal transport over the networks. Although it might seem as if any combination of equipment pools and equipment networks is valid, there are some restrictions to the equipment pools that are assigned to equipment networks, depending on the capabilities of the types of equipment of these pools.

These restrictions, imposed to any equipment network, are as follows. **First**, at least one equipment pool should be self-propelled, otherwise internal transport will not be possible at all. **Second**, whenever a trailer-type equipment pool is installed at a network then at least one pool of the appropriate, required pulling type must be installed at the same network, too. Trailer-type equipment can only be pulled by pulling equipment that is
installed at the same equipment network. In this way it is possible to easily separate various sets of pulling and trailer-type equipment pools: Just create two or more identical networks and install any desired pair of pulling and trailer-type equipment that should be discerned at a different network. Third, whenever two equipment networks are pinned to the same location and load units are to be transferred from one equipment network to the other, then naturally always one of the two equipment pools involved in the transfer should be able to tranship.

All equipment networks can be divided into two types: private networks and public networks. The main difference is the way these networks are created when using the terminal design aiding tool: Private networks are created automatically by the terminal design aiding tool itself, while public networks are created manually by the user of the terminal design aiding tool. These private networks are automatically created at SDM model level for all locations at the system's boundary and for all internal stacks; any internal transport within these locations is to use solely these private networks. Actually, the terminal design aiding tool only creates a so-called king-pin for each private network and automatically pins it to the location owning the private network. Any expansion of this private network, more than the king-pin, must be made manually by the user of the terminal design aiding tool.

No restrictions at all are imposed on overlapping equipment networks, which implies that in some cases more than one equipment network can be available for internal transport between two locations, making these equipment networks "compete" with each other. In the next section it will be explained that this is equal to internal transport on a specific section of a specific equipment routing. Normally such competing equipment networks are equipped with totally different equipment pools -- normally there should be no overlap in equipment pools at overlapping equipment networks. Nevertheless, such a situation could occur, leaving the question to find out which single network should be used then. This problem is solved as follows: The equipment optimisation model to be discussed below only takes into consideration the equipment network that has the highest so-called preference (to be discussed at the equipment routings) for the overlapping equipment pool.

7.2.4. Equipment routings

In the real-life system there are preferences to use specific equipment pools at specific transportation jobs. For instance, on a specific transportation job it is allowed to use a certain type of straddle carrier but it is preferred to use a reach stacker. These preferences have a major impact on the actual assignment of transportation jobs to machines.

This aspect is modelled by introducing equipment routings. These equipment routings are strongly related to closed tasks as discussed at Chapter 5. In fact, each closed task, which is always bidirectional according to Chapter 5, yields exactly two equipment routings, one for each direction over the closed task. This will be clarified by expanding an example given in Chapter 5 which is repeated in figure 7.3. This figure shows a single closed task running from "Central RSC" to "Nin-Jan" and vice versa over one internal stack.

![Figure 7.3: Example of a closed task (repeated).]
Based on the above definitions two equipment routings named R1 and R2 respectively can be discerned in figure 7.3 as shown in figure 7.4. All additional points included in the closed task have been removed from the figure as they are of no particular interest when discussing equipment routings; only those locations that introduce a so-called equipment break are of any interest at this moment.

An equipment break is defined as a discontinuity in internal transport over a closed task; an equipment break can only be introduced by an internal stack, an exchange point, and by the first and the last location at the closed task. Taking into account the definition of closed tasks, these latter two locations are always located at the system's boundary.

Another important element, strongly related to equipment routings and equipment breaks, are the so-called sections. These sections represent an indivisible transportation job. There are two types of sections: The first type of section is the way between two subsequent equipment breaks in a specific equipment routing. This type of section is called a normal section. An example herof is "Section 2" in figure 7.5 running from "Central RSC" to "Nin-Jan Stack". The second type of section is the local area within a location at a specific equipment routing. This location must either be an internal stack or the location at the beginning or the end of the equipment routing (and by consequence at the system's boundary). This type of section is called a local section. An example herof is "Section 3" in figure 7.5 at internal stack "Nin-Jan Stack". The subdivision of routings into sections comes close to the real-life system, although in the real-life system there may be a less clear distinction between equipment routings.

Internal transport on a section is only then possible when one or more equipment networks are available at that section, along with some equipment pools that are installed at those equipment networks. An equipment network is only then available for internal transport at a certain section when the network is pinned to both locations that mark the beginning and the end of the section through an equipment break, and when there is either a direct or an indirect connection over the network between both locations that does not run over any other location that introduces an equipment break. An example will clarify this.
Figure 7.6 displays a possible configuration of equipment networks related to one closed task, based on the situation shown in figure 7.2 and 7.3. In this configuration four different equipment networks are discerned: One public equipment network named ITT (Inter Terminal Transport) and three private equipment networks. Transport within internal stack "Nin-Jan Stack" always runs over the that stack's private network. Transport between this stack and "Nin-Jan" itself can run either over ITT or over the private network of "Nin-Jan". Figure 7.6 also shows that all additional points that are included in the closed task are of no importance at all, whether they would have been pinned to or not. Note that two of the private networks only consist of the automatically created king-pin, while the third private network, the one at "Nin-Jan", has been expanded manually by a pinned pin at internal stack "Nin-Jan Stack" and by a link between these two pins.

Any equipment pool that is installed at any equipment network as shown in figure 7.6 can be assigned to any transportation job that runs over a section that is covered by that network. This introduces the problem of defining preferences to use specific equipment pools at specific transportation jobs. These preferences are modelled by using human assessment according to the REMBRANDT scale. Refer to Appendix A for a description of various methods to model human assessment and of the REMBRANDT scale. In this context a transportation job is defined as a transportation job on a specific section, over one of the available equipment networks at that section, using one of the equipment pools at such an equipment network. An example hereof is given in figure 7.7.

![Diagram](image-url)

Figure 7.6: Example of equipment networks.

Figure 7.7: Example of preferences for equipment usage (only pool B and D shown).
Chapter 7. Equipment

Figure 7.7 shows a "box" in which an example is given for the assignment structure of equipment to specific sections, covered by equipment networks. Figure 7.7 displays a situation in which four equipment pools named A through D, which stem from four different equipment types, can be assigned to any of the four equipment networks from figure 7.6. It will be clear that equipment can only be assigned to those sections from figure 7.5 that are covered by these networks. The possible combinations between sections and networks are shown as hatched planes at the right-hand side of the "box". For reasons of clarity, figure 7.7 only shows the preferences for equipment pools B and D; the preferences for equipment pools A and C are not shown.

In this example the equipment pools are assigned as follows to the equipment networks:
- Equipment pool A to private networks at "Central RSC" and "Nin-Jan Stack"
- Equipment pool B to public network ITT
- Equipment pool C to private network at internal stack "Nin-Jan Stack"
- Equipment pool D to private networks at "Nin-Jan Stack" and "Nin-Jan".

These assignments are shown as crosses at the top side of the "box". In this example equipment pool D has a so-called fixed valuation of "++", which means that equipment pool D is always assessed as "++" regardless of the section of the equipment network. Note that a fixed valuation is assigned to the equipment pool and not to a routing or a section. Fixed valuations are important for two reasons. First, they facilitate modelling a complex situation of internal transport as the valuation of an equipment pool only needs to be specified once; thereafter it is guaranteed that the equipment pool always is assessed the same way. Second, they facilitate modelling a situation in which, for instance, a special (and therefore expensive) equipment pool is used as a backup for cases of emergency and is preferably not used at all. Such a safety pool can be modelled easily by assigning a low fixed valuation to it.

In some cases equipment networks though available for internal transport at a specific section through the pinned pins, may not be used for transport at that section. Suppose, for instance, that in the example of figure 7.6 the equipment network of ITT may be used for any section of the displayed closed task except the section between "Nin-Jan" and "Nin-Jan Stack" on the closed task of figure 7.6. In that case the equipment network of ITT may be blocked for use at that section, which means that it can not be used for internal transport at that section although it physically connects the locations between which the section is spanned.

Equipment networks may overlap as mentioned in the previous section. This means that several equipment networks may be used for transport at a certain section, which is fine except when at both (or even more) equipment networks the same equipment pool is installed. In that case the equipment pool overlaps, too. This induces an ambiguous situation in which it is not defined which preference should be adapted for use of the overlapped equipment pool, as each of the overlapping networks could specify a different preference. This question especially becomes of importance when discussing the model for equipment optimisation. This problem is solved by selecting the equipment network with the highest preference for the equipment pool, as this approach would come as near as possible to real-life behaviour (although in the real-life world normally no such things as competing equipment networks exist).

Another function of equipment routings is to specify the duration of the flows running over that equipment routing. This is modelled by associating a random distribution with each equipment routing, and therefore with each direction over each closed task, that specifies the total time that load units running over that equipment routing remain in the system. In other terms this random distribution specifies for each single load unit the difference between the time it will enter the system and the time it will leave the system again. This time is used to find a matching contra schedule for each single load unit. Note that (1) the duration meant here may differ for both directions over the closed task and (2) the load unit may run either from the (abstract) contra schedule to the (specific) load schedule, or vice versa; therefore, for each load unit the contra schedule may start either before or after its load schedule.

A major difference exists in the way the total period each single load unit remains in the system is specified at UCM and at SDM level: At UCM level this period of time is simply determined by specifying at each internal stack the time each load unit of a certain type should be kept in store. The sum of these fixed stack times, together with the simplified transportation speeds over the connecting physical paths, determine the total time that each load unit remains in the system. At SDM level, on the other hand, this period of time is specified through the above described random distribution only, leaving open to the model to determine the speed distribution of each load unit on its way through the system. In other terms: It leaves open to the model whether
a load unit remains a long time at the first stack that it comes across on its closed task, or that it moves as rapidly as possible to the last stack where it will leave the system again. Obviously, this aspect is most interesting when load units can be stacked at more than one location over which their closed task runs. In this way the model can determine in a much more accurate way the required stack sizes and speed distributions, as the result of a delicate equilibrium between various costs involved with these flows of load units. This aspect will be given more attention when discussing the so-called flow patterns and the equipment optimisation model.

7.3. Options for internal transport

In the real-life system there are a number of aspects that either influence the course of internal transport, or are essential to be included in a general model as described in this thesis to be valid. These aspects will be described in this section. In some cases, namely the case of local internal transport and of local stacking, it will turn out that with a limited number of options the character (and by consequence the level of detail) of the model can be changed considerably.

First, the so-called flow patterns will be discussed. These flow patterns determine at a high level of abstraction the flow of load units through the system in time. Second, the aspect of local internal transport will be discussed, which turns out to be essential to model a terminal system at a reasonable level of detail. Third, the aspect of local stacking, which is strongly related to local internal transport, will be discussed. Finally, the aspects of transhipment and equipment relocation are explained.

7.3.1. Flow patterns

Essential for the course of the load units through the system are the so-called flow patterns. These flow patterns can be said to specify the desired speed pattern for the flows of load units through the modelled system, per equipment routing. Note that this is only the desired speed pattern\(^2\); there is no absolute guarantee that this desire is met, as the speed pattern is also influenced by the costs of assignment of equipment to internal transport.

Flow patterns are used to model at a high level of abstraction the (human) preferences for the presence of load units running over specific equipment routings at specific locations. An example thereof is given in figure 7.8, showing the preferences for the presence of load units at both equipment routings from figure 7.4. These preferences are taken from the REMBRANDT scale. Refer to Appendix A for a description of various methods to model human assessment and of the REMBRANDT scale. In the example of figure 7.8 the preference for load units at the internal stack "Nin-Jan Stack" is definitely higher than the preference for load units either at "Central RSC" or at "Nin-Jan". This is a quite normal phenomenon for internal stacks, as they are meant to store load units.

Flow patterns are mainly used to model situations in which human beings controlling the flows of load units (especially those controlling the flows of load units at one terminal of the terminal complex only) have some distinct ideas on the preferred presence of load units: In most cases it is highly preferred to get the load units as soon as possible at the location where they will leave the system again. This is done for reasons of safety: Whenever a load unit is already present at the stack of that maritime terminal where it will have to leave the system again within one or two days, it is relatively simple, and therefore safe, to organise and control the final transhipment movement from the terminal's stack to the terminal's sea-quay. If, on the other hand, this load unit is still stacked at another, remote, maritime terminal, then it will require at lot more organisational effort to get that load unit in time at the proper sea-quay. This effort normally is too great for human-controlled systems.

\(^2\)This speed pattern is not related directly to the transport speed of equipment, but is a more high-level and abstract type of speed.
Therefore, it is a good idea to apply so-called flat flow patterns (i.e., flow patterns that return "Indifferent" for all locations for all equipment routings) to study the effect of these human preferences on the number of equipment to install, along with the costs that are involved by such human preferences. These human preferences, causing non-flat flow patterns, are caused by the incapability of human beings to survey and control a complex terminal system; in those cases where the terminal complex is planned by a computerised system there is no need to introduce non-flat flow patterns.

It is important to bear in mind that the preference for the presence of load units at bounding elements and at train handle areas is only then taken into account when local stacking (which is to be discussed shortly after) is allowed at these locations. Otherwise, the setting of the preference at these locations is ignored and has no effect whatsoever.

Most of the cases the flow patterns over both equipment routings that stem from one and the same closed task will have mirrored flow patterns. This is also shown in figure 7.8 and means that the flow pattern in one direction over the closed task is reversed with respect to the flow pattern in the opposite direction over the closed task. Normally, this will be the case, except when the pattern of internal stacks on the closed task can not be mirrored.

The preferences for the presence of load units at the various locations is used at the load unit sub-model of the equipment optimisation model to be discussed later as the stacking costs per load unit at these locations. This is where flow patterns influence the flows of load units through the system. These stacking costs are to be balanced with the equipment costs, i.e., the costs of the usage and the presence of equipment in the system.

7.3.2. Local internal transport

Local internal transport is defined as all internal transport within a location. This location can either be an internal stack, a train handle area or a bounding element. There is a slight difference between the way local internal transport is handled at these various types of locations, which will be explained in this section.

At internal stacks, local internal transport is used to model all equipment movements required to stack and to unstack load units. At bounding elements, local internal transport is used to model transport along a sea-queue in a highly simplified manner in those cases where this transport is expected to be essential for the accuracy of the model. At train handle areas, local internal transport is used to model transport to and from the rail bundles. These options will be elaborated one by one. Note that, next to local internal transport, there also is an option of local stacking at bounding elements and train handle areas. This will be discussed in the next section.

♦ Internal transport at bounding elements

For bounding elements local internal transport is optional and will mainly be used to model internal transport on a sea-queue. Due to the aimed level of detail of bounding elements it will be clear that this type of internal transport does not need to be modelled in detail. Therefore, it is modelled simply as a random distribution specifying the distance for internal transport. In most cases this type of internal transport will be totally independent of the internal transport system connecting the bounding element to other locations in the modelled system, especially when the option of local stacking, to be discussed in next section, is used. Note that local internal transport does not require the bounding element's private network to be expanded: Local internal transport at a bounding element does not occupy any space, does not actually relocate any machine or load unit, and therefore does not need any infrastructure for machine movement. An example of the application of local internal transport at a bounding element is given in figure 7.9, consisting of two parts: The left-hand part shows the "actual" local internal transport between a quay-crane unloading a sea vessel and the local stack of the bounding element. The right-hand part shows how this local internal has been condensed in a standard transport distance. It should be obvious that any dynamic coupling and therefore dynamic influence between the quay crane and the local internal transport equipment is not included in the model -- unless incorporated in the random distribution specifying the distance for local internal transport. The reason for this is quite straightforward: This type of
influence is considered as too detailed to be included in a model to design terminals for rail transhipment. Note that local stacking, which is also depicted in figure 7.9, is not discussed until the next section.

![Diagram](image)

**Figure 7.10: Example of local internal transport at a train handle area.**

*Internal transport at train handle areas*

For train handle areas local internal transport is compulsory and will be used to model two different aspects of internal transport at a train handle area: Internal transport at the handle bundles of the train handle area, and internal transport between these handle bundles and the gate of the train handle area. This is illustrated in figure 7.10, showing in the left-hand side the plan of a train handle area consisting of two yard bundles and two handle bundles (only the latter two bundles are of any importance at this point) and in the right-hand side the three private networks that automatically result from the plan shown at the left-hand side: One private network for each handle bundle, and one private network for the train handle area as a whole; the latter private network is called the main private network of the train handle area, the first private networks are called the handle private networks of the train handle area.

Note from figure 7.10 that both handle private networks consist of two king pins instead of one, and of a network link that is spanned between these two king pins. These elements are all created automatically and can not be removed by the user of the terminal design aiding tool than by deleting the related handle bundle. In contrast to normal king pins, these king pins are pinned to rail points rather than to UCM locations. The network link of each handle private network is used to move equipment along the related handle bundle.

Both types of private networks at a train handle area have a different function: The handle private networks are used to model handling of trains at the associated handle bundle. Or, in other terms, they are used to model transhipment of load units to and from trains at the handle bundle at which extreme points the private network is pinned. They also are used to model stacking of load units at the handle bundles themselves. The main private network of the train handle area is used to model internal transport between the handle private networks and the gate of the train handle area. This functional division between both types of private networks is fixed. Therefore, the main private network may be expanded to connect the train handle area through its own private network to other locations, but the handle private networks do not need to be expanded.

There is no need to connect the main private network to the handle private networks. Transport between the gate of the train handle area and the handle bundles is done via imaginary, rectangular, network links between the king pin of the train handle area and the nearest of the two king pins of each handle bundle.

Giving a glimpse in advance of the way transhipment and local stacking are organised at train handle areas, it can be said that all these activities occur at the private networks of the handle bundles. The main private network of the train handle area is solely meant to model internal transport between the gate and the handle bundles. If the same equipment pools are installed at the main private network and at the handle private networks, then a situation of so-called closed internal transport is created, which is to be discussed at figure 7.37.

---

3 Which can be train handle areas, bounding elements, internal stacks, exchange points or additional points.
Internal transport at internal stacks

For internal stacks local internal transport is compulsory and will be used to model internal transport within the stack, i.e., internal transport to bring load units into the stack and to get them out of the stack again.

As will be discussed in Chapter 8 an internal stack at SDM model level consists of several stack blocks which are the actual locations at which load units are kept in storage. An example hereof is given in figure 7.11, showing at the left-hand side a plan of an internal stack consisting of eight stack blocks; between these stack blocks a number of paths for the movement of equipment is shown, along with a connection between these points and the gate of the internal stack.

The right-hand side of figure 7.11 shows the private network that is created automatically for the internal stack, along with the king-pin of the private network. Note that the connection between the network spanned between the stack blocks and the gate of the internal stack (i.e., the king-pin) is imaginary and does not need to be created by the user of the terminal design aiding tool, and therefore it is ambiguous which should be the transport distance to use during simulation. This problem is solved as follows, in which two different cases in which this transport distance is needed are discerned:

1. during actual simulation the transport distance is either determined (1) by a random distribution in case of simplified stacking (to be discussed in Chapter 8) or (2) by the rectangular distance from the gate of the internal stack to the most remote corner of the selected stack block in case of non-simplified stacking;
2. during equipment optimisation the transport distance is either determined (1) by the average of the just mentioned random distribution in case of simplified stacking, or (2) the average rectangular distance between the gate and all stack blocks in case of non-simplified stacking.

Various options to configure some details of the internal stack's private network are available in TORCMOD: Whether it should be "closed" at the top, bottom, left, and/or right side. Note that these details do not affect the functionality of the internal stack at SDM level at all; they are solely added to improve the visual resemblance of the modelled system with the real-life system.

Internal transport within the internal stack can be done by any equipment pool installed at its private network, but there are some specific rules according equipment pools that can transship but can not move. These equipment pools are considered as fixed at the stack blocks with a service area restricted to that stack block. Refer to table 7.4 for more details on this subject.

7.3.3. Stacking

Three types of locations can be used to model stacking of load units at SDM and DDM level of detail:
- bounding elements;
- train handle areas;
- internal stacks.

All three types of locations aim to model a different kind of stacking, either different in location or in detail. This section gives a brief description of the most important aspects on stacking, as far as needed for further discussion of the equipment optimisation model. A full overview of stacking can be found in Chapter 8.
Bounding elements
Stacking at bounding elements is called local stacking for reasons explained in Chapter 8 and models marine stacks at a coarse level of detail. Using this type of stacking is optional, but if it is used it must be specified whether the stacking capacity is restricted to a certain maximum capacity. This limited capacity will then be respected in the equipment optimisation model.

Train handle areas
Stacking at train handle areas is quite analogous to stacking at bounding elements: It too is called local stacking and is optional. The main difference is that this type of stacking is used to model stacking at handle bundles, next to the trains. The limited stacking capacity, if required, is expressed in a capacity per unit of length of all present handle bundles.

Internal stacks
Stacking at internal stacks is intended to be the most detailed type of stacking. For this purpose each stack can be divided into any number of stack blocks at which load units can be stored. To each of these stack blocks priority rules can be attached to control the assignment of individual load units to particular stack blocks. Refer to Chapter 8 for more details. In contrast to both previous types of stacking, stacking capacity at internal stacks is always unlimited. This facilitates to satisfy following condition.

Condition for stacking capacity

As mentioned before, limited stacking capacities are included in the equipment optimisation model. By consequence a situation as depicted in figure 7.12 could occur. This figure shows a closed task that runs over two bounding elements and one internal stack, each location either totally impeding stacking or offering solely a limited stacking capacity. If during simulation 750 load units were to be simultaneously on their way from the left-hand side bounding element to the right-hand side bounding element then an invalid situation would occur as the available stacking capacity is not sufficient to store all 750 load units at one time.

To prevent this type of invalid situation it is demanded that at each closed task at least one location with unlimited stacking capacity is present. In this way most erroneous situations due to a lack of stacking capacity can be avoided. This condition is called the limited stacking condition in the remainder of this thesis.

Normal use of stacking capacities would be to either set the stacking capacity of some bounding element to unlimited or to include an internal stack in a closed task, which always has an unlimited stacking capacity. Less frequently it will be decided to set the stacking capacity of a train handle area to unlimited.

7.3.4. Transhipment

Transhipment is included in the equipment optimisation model at those locations where two equipment pools meet with different capabilities: The first equipment pool must be able to tranship, while the second equipment pool may not. This will be clarified with an example.

---

\footnote{In reality the situation is even worse as this figure only shows flows over one closed task, while in reality several closed tasks may claim capacity at these stacking facilities.}
Consider figure 7.13 in which two sections on a particular equipment routing are depicted. At the leftmost section, which is named "section 1", two equipment pools with different capabilities are installed: One pool of MTS-trailers and one pool of straddle carriers.

At the rightmost section, two other equipment pools are installed which also have different capabilities: One pool of MTS-trailers and one pool of reach stackers. Between both sections load units can be transshipped according to the scheme depicted in figure 7.13, based on following rules:

1. Transshipment may only occur between (1) a transshipment pool like the reach stackers or straddle carriers and (2) a non-transshipment pool like either of both MTS-trailer pools.

2. Transshipment only occurs at locations that introduce an equipment break and by consequence only occurs at the transition from one section onto the other.

3. Transshipment may be done by any transshipment pool that is installed either of both sections. One exception to this latter rule is made for non-movable transshipment pools that are installed either at internal stacks or at handle bundles of train handle areas. This type of pools is considered to be the gantry cranes that are to be used exclusively for a transshipment into or out of the section they are installed at. For the moment this aspect will not be discussed further; it will be explained in more detail later on in this chapter and in Chapter 8.

Transshipment between two equipment pools that are both capable to transship is considered as invalid in the model and therefore is automatically ignored.

In case of figure 7.13 a special situation is modelled in which both transshipment pools (the straddle carriers and the reach stackers) can not only be used for transshipment from and onto the MTS-trailers, but can also be used for actual internal transport over either of both sections. This situation is shown in a schematic way in figure 7.14, and will be explained in more detail in section "Equipment sub-model" of this chapter.

In this way relatively complex assignment structures for equipment can be created in which parallel equipment pools "compete" with each other to transport and/or transship load units. The best solution for this problem is found by using equipment optimisation.

### 7.3.5. Closed- and broken internal transport

During installation of equipment pools some special configurations can be obtained that play an important role in the equipment optimisation model to discuss shortly. These special configurations will be explained by referring to figure 7.15 in which one particular equipment pool is used on two subsequent sections of the same equipment routing. Between both sections an exchange point is located and therefore introduces an equipment break between these sections which means that possibly two different equipment pools may be used to transport the same load units over either of both sections. At this equipment break load units may be transshipped from one equipment pool onto the other equipment pool. However,

---

6Different from the pool installed at the first section.
transhipping load units between two machines of the same equipment pool as shown in figure 7.15 seems to be illogical as it concerns a direct transhipment between both machines without intended stacking in-between.

This particular situation is recognised as broken internal transport and actually is considered as invalid. In other terms, in those cases where broken internal transport is detected the equipment break that caused broken internal transport is ignored\(^7\) and both sections are considered to be one. This will be discussed in the next paragraph in more detail. Broken internal transport may occur either at an exchange point as shown in figure 7.15 or at the connection between a private network and another network -- which is transition of local internal transport to normal internal transport.

Broken internal transport is mended by connecting both subsequent sections into one new fictitious section as shown in figure 7.16. In this way the function of an exchange point can be totally eliminated for the pool that causes closed internal transport! This will be given some more attention in the next paragraph. As indicated a little before, closed internal transport can not be used for internal stacks as they actually do decouple transport over both sections, and even add some additional, local sections that are used for transport inside the internal stack.

Some complications are introduced by closed internal transport as will be discussed by referring to figure 7.17. This figure depicts a situation in which two subsequent sections are discerned with different pools installed. At section 1 two equipment pools are installed: One pool of straddle carriers and one pool of trailers. At section 2 only this particular pool of trailers is installed.

Closed internal transport yields the situation that two different ways to transport load units over section 2 by means of the trailers are discerned: (1) transport by the trailers that are loaded at the edge of section 1 by the straddle carriers, and (2) transport by the trailers over both sections without transhipment in-between. This is the actual closed transport movement. In this example it is not specified how the trailers are loaded at the beginning of section 1; in the actual model this transhipment movement will probably be executed by the straddle carriers at the very beginning of section 1.

In the final equipment optimisation model these two different ways for transport will yield two different arcs for transport by the pool of trailers over section 2, each arc having different parameters assigned to it like the transport duration.

7.3.6. Equipment relocation

One important aspect that should be included in any model describing the processes at a large-scale terminal complex is the aspect of equipment relocation. For those controlling a real-life terminal it is the aspect that complicates planning of equipment movements over the terminal, especially in case of large driving distances. For those developing a computer model to optimise the assignment of equipment it is the aspect that increases the complexity of the model considerably. Nevertheless, equipment relocation is thus important that it must be included in a model to be valid.

\(^7\)For that particular pool only.
In short, equipment relocation is the unpleasant, but real-life effect that sometimes equipment needs to be moved from one location to another before it can be used for an actual internal transport movement. This relocation is always defined as empty transport and therefore is a waste of transport capacity. Therefore, equipment relocation should be avoided as much as possible and plays a major role in the planning of the use of equipment. An example of equipment relocation from A to B before the actual transport movement from B to C is given in figure 7.18.

In those cases where equipment relocation can be combined with transport of load units the relocation movement is not considered as a relocation movement anymore, but as a normal transport movement. In fact, such a combined transport is the result of good planning of equipment. An example can be seen from figure 7.18: If relocation from A to B can be avoided by combining equipment relocation with transport of load units from A to B then created two separate transport movements have been created: One from A to B and one from B to C. In that case there is no relocation movement.

7.4. Optimisation of equipment usage at SDM level

Now that it is known how to model almost any situation of internal transport at SDM level, there is a need to determine the optimal number of equipment to install at all equipment pools at the system. This is a balance between installing more equipment on one hand, thus complying better to the desired flow patterns, and inducing higher operational costs on the other hand, thus making the system possibly too expensive.

There are two major aspects complicating the problem. First, in the real-life world most load units remain definitely longer at the terminal complex than strictly needed to transport them from system entry to system exit. Normally, this can not be influenced by the terminal complex itself, but is imposed by the owners of the load units. Therefore, there is a great amount of liberty in planning when load units are moved by internal transport. Second, a great number of independent equipment pools that still have to co-operate are installed at most large-scale terminal complexes. As this optimisation problem is too complex to be answered by made calculations by hand, another approach is needed to solve it; there is a need to use a mathematical technique to find the optimal number of equipment to install at the various equipment pools in the system. In the remainder of this section the reader is assumed to be familiar with the contents of Appendix C on optimisation techniques.

♦ Use of equipment optimisation

Equipment optimisation is a method to find the best number of equipment to install at any pool at a complex internal transport system, taking into account the available liberty in scheduling transports, the preferred flow patterns, the preferred equipment assignment structure and the equipment costs.

♦ Application of equipment optimisation

Equipment optimisation is used at a regular basis during simulation of the model at SDM level. There is a fixed interval at which the long-term usage of equipment is planned and optimised; this is the planning interval. The results hereof are used to assign the actual transport jobs to individual machines. This is clarified in figures 7.19 and 7.20. Figure 7.19 shows that equipment optimisation and simulation, based on the results of this optimisation, are performed alternately.
Figure 7.20 shows the course of the simulation on two time axes; the lower axis represents the actual (calculation-) time and the upper axis the simulation time. In this figure two types of boxes are discerned, each one representing a different type of action. Dark grey boxes represent the time consumed by planning and optimising the usage of equipment, while light grey boxes represent the time consumed by simulating the transport movements based on the found optimal solution. Note that these two activities are executed alternately. All light grey boxes, representing simulation activities, have equal duration expressed in simulation time. This duration is the previously mentioned planning interval. All dark grey boxes, representing planning activities, have a duration of zero time units expressed in simulation time. So, every time the planning interval has elapsed in simulation time the future activities of all equipment are determined by optimising the long-term assignments. In this case "long-term" means: within about three full days from the current simulation time on.

Figure 7.20: Usage of equipment optimisation during simulation (II).

Load units

Figure 7.21\textsuperscript{8}: Example of peaked flows of load units crossing the system's boundary.
The basic idea about equipment optimisation was to shift flows of load units in time to reduce the number of required machines and in the mean time to meet the required flow patterns and the preferences of all equipment routings best possible. That such an optimisation model is to cover about three days is argued in figure 7.21 showing the total of all flows entering the terminal complex during four days, along with the sum of all flows leaving the system again. Note the rather peaked character of these flows. The aim of the equipment optimisation model is to flatten the use of equipment during these days as much as allowed thus reducing the number of equipment needed to install. This is shown in figure 7.22. Obviously, the working hours during the weekends play an important role in this special case, and therefore are included in the equipment optimisation model.

Machines used

Figure 7.22 Example of flattened equipment usage after equipment optimisation.

As equipment optimisation is to be performed at a regular basis during simulation, and the optimisation covers about three full days, it will be clear that there is no need to discern every single load unit (there may be several thousands of them) and every single machine (there may be about a hundred of them) in the optimisation model. Such a model would require too much time to solve, while there is no need at all to know the precise future activities in advance; some reasonable estimations will do as well. Instead, the optimisation model is based on

---

\textsuperscript{8}This type of flows is typical for continental transport, marine transport has a different character.

---

page 89
time windows, which are time intervals of a certain length, measured in simulation time. Each time window can cover a lot of transport jobs that can not be discerned individually. By consequence, the character of the optimisation method is rather continuous than discrete.

The major drawback of time windows is the difficulty of converting the results, which do not discern individual load units or machines, into the actual load units and machines. Even more, what to do if the optimisation method advises us to use 1.34 machines on a transport job? These problems should be given attention when using time windows. The precise definition of the time windows will be given later on.

**Method requirements**

A method for equipment optimisation should meet a number of requirements. These requirements are:

1. the method should be able to make a sensible optimisation of the number of equipment to use and therefore it should cover a sufficiently long period of time in order prevent peak loads by moving transports in time. Refer to figure 7.22. This aspect is elaborated when discussing the planning phases.
2. the method should take into account a number of real-life aspects to make the model valid. Some examples hereof are working hours and equipment relocation, i.e. empty transport.
3. the method should be fit for use during simulation at a regular time interval. In other terms, it should be (1) relatively simple, accepting simplifications where needed but including all relevant aspects as discussed before to be valid, and (2) be based on a technique that allows fast solving of the type of problem discussed here.

Following sections will show how these requirements are met by the optimisation model equipment hopping.

### 7.4.1. Basic definitions

Before the equipment optimisation model itself can be developed several aspects that play an important role in the optimisation model need to be discussed. These aspects are the knowledge horizon, the planning interval and the division into discrete time windows. Each of these parameters is discussed in this order.

These parameters highly influence the behaviour and the results of the equipment optimisation method along with the required computing time to solve the problem. Each model may therefore have its own setting of these parameters.

**Knowledge horizon**

One of the major differences between a computer model and the real-life system is that a computer model has exact information on future events at its disposition, while the real-life system is confronted with a number of uncertainties and lacks in the available information. In this case it means that the equipment optimisation model has at its disposition exact information on all future load units that will ever enter or leave the system, while the real-life system only has at its disposition some estimations on some future load units that will enter and leave the system.

This real-life aspect is modelled by introducing the knowledge horizon; this is the period of time that the computer model can look ahead from the current simulation time on. Any load unit entering the system between the current simulation time and the current simulation time plus the knowledge horizon is known to the computer model.

Compared to the real-life situation there are several differences. First, in the real-life system there is no such thing as a fixed knowledge horizon; on the contrary, some future load units are booked some days ahead, while others are booked just a few hours ahead. A single knowledge horizon as introduced here ignores this effect. Second, the computer model uses exact information on times, including disturbances in system entry or system exit time. In the real-life system these disturbances are not known. The effects of these discrepancies with the real-life system on the results of the model analysis are hard to predict, especially in those cases where accurate data on the real-life system lacks.

The user is free to chose the value of the knowledge horizon, just as long as he takes into consideration the comments given at the discussion of the planning phases are satisfied. In most cases the value of the knowledge horizon would be somewhere in the range from ten hours to five days. It may be a good idea to change the size...
of the knowledge horizon to study the effects of future electronic booking methods which may either result in extremely short-term booking or in (extremely) long-term booking of load units.

- **Planning horizon**
  
  Any load unit that is to arrive after the knowledge horizon is not known to the equipment optimisation model, while every load unit that is to arrive before the knowledge horizon is. As every load unit entering the system runs over an equipment routing with a known duration, it is also known at which time each load unit will leave the system again. By consequence, a planning horizon exists that equals at least the knowledge horizon plus the maximum duration of all equipment routings in the system; beyond that planning horizon there is no information on any flow of load units. Refer to figure 7.23. Note that in the time interval between the knowledge horizon and the planning horizon all information on load unit flows fades out; immediately after the knowledge horizon ends the model still has perfect information on all load unit flows that will leave the system again; these load unit flows have entered the system some time before the knowledge horizon. The further beyond the knowledge horizon the more information on load units leaving the system has been "used" without having been replenished with new information by new load units entering the system. Also note that the information that is available after the knowledge horizon is not complete anymore and therefore will cause the equipment optimisation model to plan less internal transport movements than eventually will be needed. By consequence, the knowledge horizon must be selected such that this incompleteness of information does not yield to invalid results of the equipment optimisation model. In other terms, it must be avoided that the incompleteness of information has any effect on the results of the equipment model. This is elaborated below at the planning phases.

- **Planning interval**
  
  As shown before, the planning interval is the interval in simulation time at which the equipment usage must be optimised. Refer to figure 7.20. There is a great freedom in choosing the value of the planning interval; for example it may be either fifteen minutes or eight hours of simulation time.

  Most important are the effects of a wrongly chosen planning interval. First, too small a planning interval will considerably increase the calculation time of the optimisation model, without any sensible pay-off. Second, too small a planning interval will introduce obvious rounding problems that are due to the differences in character between the optimisation model (which is a continuous model) and the actual simulation model (which is a discrete model) in which jobs are actually assigned to machines. By consequence, there will always be some rounding problems converting planned transport jobs into actual transport jobs. Third, too small a planning interval will enlarge the effects of some simplifications of the (continuous) optimisation model in the actual job assignment. Some examples hereoff are estimates of the ratio inter-/intra window transports, estimates of relocation times, estimates of transhipment and transport times. This is elaborated later on. Fourth, too great a planning interval will remove all opportunity for the optimisation problem to correct for unexpected problems that occurred during the actual execution of transport jobs, such as unexpected delays or alternative job assignments.

  Some demands are imposed on the possible values of the planning interval. For reasons of consistency of the optimisation model, the planning interval should be expressed in a natural number of micro windows; these micro windows are explained below. By consequence, the absolute lower limit of the planning interval is ¼ hour. A value between one and two hours seems reasonable, but in practice the actual value will depend on the computer on which the model is executed, on the available calculation time and on the aimed accuracy as well.

  A good expansion of the optimisation model would be to implement the planning interval as automatically dependent of the current intensity (i.e. at the current simulation time) of internal transport flows. The major advantage of this expansion would be a reduction of calculation time. Although this expansion is relatively simple to implement in the computer model, it is not used at the moment. This is due to a lack of experience.
with the optimisation mode, as it is not implemented yet at the time. By consequence, it is not yet possible to establish the effects of such a flexible planning interval on the results of the model.

**Section transport time**
A major problem of the equipment optimisation model is the wide variety in pure transport times. For instance, there may be internal transport jobs over extremely short distances (about 100 meters) as found within a small stack, along with jobs over extremely long distances (about 5 kilometres) as found at the Maasvlakte terminal complex.

Like there are differences in pure transport times, there are differences in pure transhipment times. Some types of equipment can tranship definitely faster than others. Any transport movement, except empty transport, corresponds to two transhipment movements: One transhipment movement before and one transhipment movement after the transport movement. The section transport time is defined as the sum of these three times per section, i.e. the sum of the pure transport time, of the transhipment time before and of the transhipment time after the actual transport movement9. By consequence, each equipment pool on each section may have a different section transport time as the various speeds differ between various equipment pools. The section transport time serves as an important element in the equipment optimisation model, facilitating to specify the duration of various transports later on in this thesis.

An example of the section transport time on two subsequent sections in a routing that are coupled through an exchange point is given in figure 7.24 in which equipment pool A is used for transport at the first section, while equipment pool B is used for transport at the second section. An overlap exists between the section transport times on these two sections during the transhipment between both equipment pools.

When establishing the correct value for each section transport time in the system some problems are encountered. First, we need to discern between the various equipment pools that can be used at each section, as just found. This, however, is not a major problem. Second, we are confronted with a wide variety of equipment types, some of which are not able to tranship while others are not even able to carry load units either; these latter types are pulling equipment types. To determine a valid estimate of the section transport time for these equipment types some rules need to be constructed. The first rule concerns purely the aspect of transhipment capacity. This capacity is estimated as the average value of the transhipment capacity of all equipment pools that are installed either (1) at any network covering the currently considered section or (2) at any network covering the previous or next section of the currently considered section--this latter choice depends on the transhipment movement that is considered: Loading or unloading. Equipment pools installed at the currently considered section may be used both for loading and for unloading. Referring to figure 7.24 it can be seen that transhipment onto equipment pool B can be done through equipment pool A only, the pool that is installed at the previous section of the currently considered section--i.e. the section of equipment pool B. It also can be seen that transhipment from equipment pool B can be done through some equipment pool that is installed at the section after the currently considered section--i.e. the section of equipment pool B-- which is not depicted in figure 7.24. Another aspect that is of importance at this point is the number of load units that is pulled per trailer by pulling equipment types. This problem is solved, as will be explained when discussing the networks for MLU-equipment10, by distinguishing explicitly between the various trailer-type pools that may be pulled. In this way the actual capacity of the right trailer-type equipment is used for each pulling equipment movement.

**Routing transport time**
While the section transport time was defined as the transport time on a single section including transhipment before and after the actual transport, the routing transport time is defined as the sum of all section transport times of all sections in a routing. As a matter of fact, the routing transport time is a rather pessimistic value: It is composed of the sum of all greatest section transport times of all sections in a routing. The maximum value of

---

9The time needed for equipment relocation has been omitted purposely from the section transport time to make it independent of the origin of the equipment. This relocation time, however, is included in the final equipment optimisation model.

10Abbreviation of: Multiple Load Unit-equipment, which is equipment that can carry several load units at one time.
all routing transport times in the system is used to determine the planning horizon as at least the sum of this maximum routing transport time and of the knowledge-horizon that was specified by the user. The value thus found for the planning horizon is rounded upwards to fit the third planning phase which is to be discussed shortly to a number of so-called micro-windows.

**Micro window**

A micro window is defined as the smallest time window that is discerned in the equipment optimisation model; it can not be divided into smaller time windows. Any time window in the optimisation model equals an integer number of micro windows. Therefore, the micro window serves as an atomic element of all time windows in the optimisation model. Micro windows are used to define all time windows in the optimisation model, and to define the planning interval, too.

The user is free to assign any value to the micro window, as long as it satisfies following conditions. First, the micro window should be at least as long as the smallest section transport time, i.e. at least as long as the mean duration of the fastest transport over any of the equipment sections including transhipment before and after the actual transport. This condition sets the maximal resolution of the equipment optimisation method, thus guaranteeing that the level of detail of the equipment optimisation method is not unnecessarily high. Second, the micro window should not be less than ¼ hour; otherwise the equipment optimisation model would be too detailed. Third, the micro window should not be greater than two hours; otherwise the results of the equipment optimisation method would be thus rough that they are no real basis for the actual transport job assignment. Fourth, the micro window should have a "neat" value, i.e. a value of 2^k hours. Combining with the last two conditions, n should be ∈ [-2, -1, 0, 1]. Note that the requirements stated here are quite subjective, but nevertheless are believed to be reasonable.

**Planning phases**

The equipment optimisation model discerns three different planning phases, each aiming a different level of detail and describing a different period of time. The general structure hereof is given in figure 7.25 showing the sequential character of these phases. Note that the first phase, which is closest to the current simulation time and actually starts at the current simulation time, consists of the smallest time windows of all phases. By consequence, the first phase reaches the highest level of detail of all; this is a necessity as the results of the first phase (and of the first phase only) form the basis for the actual transport job assignment.

While the first phase covers the "short-term" planning of all equipment usage, the second phase describes the "mid-term" planning, which can be a little more rough than the "short-term" planning. Finally, the third phase covers the "long-term" planning that lasts up to the planning horizon. Refer to figure 7.23.

Each planning phase has a different width of time windows and a different duration associated to it; these parameters influence the calculation time of the optimisation model and the accuracy of the results of the optimisation model. Although there is quite some freedom in choosing these parameters, they are subject to following rules: The time windows of the first phase are all to equal exactly one micro window as defined before; the first phase should cover a period of one to four hours. The time windows of the second phase should be at least two times as wide as those of the first phase, and lay in the range from two to eight hours; the second phase should cover a period of four to thirty-two hours. The time windows of the third phase should be at least two times as wide as those of the second phase, but at least eight hours and at most twenty-four hours long; the third phase starts directly after the second phase and ends at the planning horizon. Remember that this planning horizon was to equal at least the knowledge horizon plus the greatest routing transport time. Now we can derive what the precise ending time for the third phase is, or, which is equal, what the planning horizon is: It is the knowledge horizon plus the maximum routing transport time rounded upwards to match the ending time of a natural number of time windows of the third phase. Refer to figure 7.23 and table 7.1.
### Table 7.1: Parameters of all three planning phases.

| Phase length |  
|--------------|--------------------------------------------------|
| Micro windows | Hours | Hours | Smallest windows | Largest windows |
| First phase  | 1     | [¼..2 hrs] | [1..4 hrs] | [4..16] | [1..2] |
| Second phase | ≥ 2   | [2..8 hrs] | [4..32 hrs] | [2..16] | [1..4] |
| Third phase  | ≥ 4\(^\ddagger\) | [8..24 hrs] | [8..∞\(^\ddagger\) hrs] | [1..∞] | [1..∞] |

The knowledge horizon, in some cases, has a bad influence on the results of the first phase. Suppose, for example, that a situation as shown in figure 7.26 was modelled. In this situation the knowledge horizon is just a little greater than the length of the first phase. The upper part of this figure also shows one of the routing transport times in the system. The lower part shows the relative number of transport movements that is known to the equipment optimisation model to enter the system, relative to the total number of transport movements that eventually will enter the system. The hatched plane in figure 7.26 reads "falsely expected free capacity", which means that the equipment optimisation model "overlooks" that fraction of all transport movements; therefore is inclined to plan transport movements to start in the first window of the second phase, even when they should have started earlier for a good optimisation result. In other terms, it is inclined to postpone transport movement that were to start in the first phase due to a too small value for the knowledge horizon. This effect will be called the optimistic plan effect.

When the equipment optimisation model is invoked again, next time, it will have to pay for the fault shown in figure 7.26: The flows it has postponed to the first window of the second phase have not been planned to start in the first phase, before the planning interval would elapse, therefore have not been executed in the simulation between this and the previous optimisation, and therefore must be planned this time in the first phase in order to meet the required times of system exit of these flows. Apart from these postponed flows the newly known flows that start after the previous knowledge horizon are to be planned, too. Therefore, these latter flows will have a good chance to be postponed to the first window of the second phase, where the equipment optimisation model falsely expects a lot of free

\(^\ddagger\)At least two times as much micro windows as the second phase.

\(^\ddagger\ddagger\)At least the knowledge horizon plus the greatest routing transport time.
capacity... Now we have entered a loop of postponing flows due to invalid expectations on future free capacity that can not be broken. Refer to figure 7.27.

This situation can be avoided by setting the knowledge horizon to a value that equals the end of the first phase plus the maximum section transport time in the system, in the meantime ensuring that at least one full time window of the second phase is between the end of the first phase and the knowledge horizon. In that case the equipment optimisation model can not postpone flows that should have started in the first phase due to a falsely expected free capacity after that first phase. Note that, as we are studying systems at which each closed task consists of a number of partial transportation movements, it is even more safe to replace the maximum section transport time with the maximum routing transport time. In this way, it is avoided that the results of the first phase are influenced by the optimistic plan effect.

The planning interval may not be greater than the first phase is. Otherwise, the actual job assignment would partially be based on the results of the second phase which is too rough for this purpose. On the other hand, the planning interval may be smaller than the first phase is. This has the positive effect that the optimisation model can correct for unexpected problems that occurred during the actual execution of transport jobs, such as unexpected delays or alternative job assignments. Obviously, this increases the calculation time.

Note that the division in phases given here yields the simultaneous execution of all three phases, each time the optimisation model is executed. Although this may seem obvious, there is one drawback: Increased calculation time, compared to separate execution of all three phase. This concept (although not adapted in this thesis) is shown in figure 7.28, and uses the results of the optimisation of the third phase as input for the second phase. Likewise, it uses the results of the optimisation of the second phase as input for the first phase. Although this concept saves calculation time, it introduces major difficulties when translating results on individual load units and machines into input data for the other phases.

![Figure 7.28: Alternative planning phases (not adapted).](image)

7.4.2. Two optimisation models

Until now it has been spoken of the equipment optimisation model. In reality, in this thesis two different equipment optimisation models are discerned. The first model, called inter-window equipment hopping, is the most detailed one; this model tries to include transport across time windows as accurate as possible. In some cases, however, the assumptions that have been made in this model may turn out to be invalid. In these cases inter-window equipment hopping yields an infeasible problem that can not be solved, and by consequence some other method is needed to solve the problem. This method is called intra-window equipment hopping, and is a little less detailed than inter-window equipment hopping; this model does not include transport across time windows, thus avoiding infeasibility of the problem. Both models and their use will be discussed in detail.

For reasons of clarity, INTRA-window equipment hopping will be written in small caps as shown from now on; inter-window equipment hopping will be written in plain text.

**Two sub-models per optimisation model**

Both equipment optimisation models, i.e. inter- and INTRA-window equipment hopping, consist of two coupled sub-models: The first sub-model describes the flows of the load units through the system; this is the load unit sub-model. The second sub-model describes the assignment of equipment to the flows of load units through the system; this is the equipment sub-model. Both sub-models are coupled through side constraints which match the number of transported load units with the number of required machines. Refer to Appendix C for a description of these side constraints.
7.4.3. INTRA-window equipment hopping

This section describes the first of the two equipment optimisation models, named INTRA-window equipment hopping. This optimisation model consists, like the second optimisation model to be discussed later, of two separate sub-models: One sub-model describes the flows of load units through the system, and the other sub-model describes the flows of equipment through the system. Naturally, there is a strong relationship between these two sub-models, as flows load units can not exist without flows of equipment. Or, in other terms, load units require the presence of transporting equipment to move through the system.

First, the sub-model for load unit flows of INTRA-window equipment hopping will be described. Then the model for equipment flows of INTRA-window equipment hopping will be described.

7.4.3.1. Load unit sub-model

The purpose of the load unit sub-model, both for inter- and INTRA-window equipment hopping, is to describe the flow of load units through the system, taking into account for each single load unit various aspects like (1) the fixed time of system entry, (2) the fixed time of system exit, (3) the equipment routing, and (4) the preferred flow pattern of the equipment routing of the load unit.

♦ What is modelled?
The load unit sub-model is used to model all movements of load units through the system. Although you may have got the impression that the load unit sub-model is transparent with respect to the equipment pool that is used for each transport, this is not true. As a matter of fact, at a number of places the load unit sub-model is strongly related to the equipment that is installed at the various networks, and even clearly shows these relations. Equipment shows through in the load unit sub-model (1) as each flow of load units over a routing is related to exactly one equipment pool, and (2) at those points where two non-transshipping equipment pools meet. It will be shown shortly what the effect of these aspects on the load unit sub-model is. The load unit sub-model also includes stacking of load units at locations of various types. This does not only hold for INTRA-window equipment hopping, but also for inter-window equipment hopping.

♦ How are things modelled?
The load unit sub-model is entirely formulated as a generalised network problem, which means that the model can be presented in a graphical way. Generalised networks are described in Appendix C, along with a description of concepts like supply, demand, arc and node. A thorough understanding of these concepts is required for a good understanding of this section.

The load unit sub-model is an integer network which means that all flows that run over the network are expressed in natural numbers. Therefore, any flow over any arc of the load unit sub-model must be in the range [0,1,2,...] and can never have a real-type value. This choice facilitates conversion of the results of the equipment optimisation model into flows of discrete load units. Later on it will turn out that the equipment sub-model is not an integer network, and therefore may sometimes advise to use something like 1.56 machines for a transport job. It will be shown that this property of the equipment sub-model, although unpleasant, is necessary for a consistent optimisation model. From these differences in the nature of both sub-models you may have deduced that the load unit sub-model leads while the equipment sub-model follows the flows of the load unit sub-model.

♦ Getting started
First of all, all load units entering the system before the knowledge horizon are condensed into time windows; each single load unit is registered at two time windows: At one time window for system entry, and at another time window for system exit. In some cases these two time windows are the same. The selected time window is

---

13This choice is made as (1) generalised network problems (with or without side-constraints, refer to Appendix C) can generally be solved quicker than general LP-problems can, (2) the transportation problem (Appendix C) does not offer intermediate stacking, and (3) the transshipment problem (Appendix C) does not offer necessary features like (a) optimising two simultaneous networks that are interrelated but may not be mixed (i.e. load units and machines), (b) guarding a maximum stack capacity over several flows that may not be mixed, and (c) matching flows on two different sets of arcs (i.e. flows of trucks and flows of trailers). A study of available literature on this subject did not result in clues to solve these drawbacks, or in examples of other models that are based on the transshipment problem, allowing optimisation of a model alike the one described in this thesis consisting of several non-mixable, interrelated flows.
simply the one in which the time found lies. As stated before, the difference between system entry and system exit is determined by the equipment routing to which the load unit is assigned. As it is known from which load volume each single load unit stems, it is also known at what location the load unit enters the system and leaves the system again. The total numbers of load units thus calculated are used as demands and supplies (refer to Appendix C) per time window and per equipment routing, for each location at the system's boundary. By consequence, the optimisation method must find a solution such that these demands and supplies are satisfied; if no such solution can be found, then the model is infeasible.

After the load units have been condensed into time windows, the network for transport of load units must be created. This is a simple geometry/time network, connecting all locations in space (via equipment routings) and in time (via stacking).

There is a difference between the network for inter-window equipment hopping and INTRA-window equipment hopping: while INTRA-window equipment hopping allows only transport within a time window, inter-window hopping allows both transport within a time window and transport through time windows. This will be illustrated by elaborating an example given in Chapter 5, which is repeated in figure 7.29. It shows a single closed task running from "Central RSC" and "Nin-Jan" and vice versa. As indicated before, each closed task yields two separate equipment routings, so that's just what will be done here.

The network will contain only those locations that introduce an equipment break; there is no need to include additional locations in the optimisation model, as no special activities concerning equipment exist at those locations. This results in a very simple geometric network containing two equipment routings stemming from one closed task, connecting one train handle area, one internal stack and one bounding element through internal transport. Every arc in the network represents internal transport of load units on an equipment section. Information on the position of the various locations is not shown in figure 7.30. In fact, for both equipment routings a separate load unit network exists with its own nodes and its own arcs; there is no interference at all at this point between the two routings. This property holds for any pair of equipment routings.

Although figure 7.30 looks nice, it would be highly desirable to add some geographical information to it, just to get a better grasp on the model that is made. For this reason a set of three axes is defined according to figure 7.31: The X- and Y-axis are used to display information on the position of the locations, and the R-axis is used to display the various equipment routings in the network.

---

14Another option would be to include the duration of the transport of the load unit in the selection of the time window to prevent difficulties at the window's boundary. However, this option is not used due to the highly fragmented transports and great liberty in scheduling found at large-scale terminal complexes.
Chapter 7. Equipment

When using this axis definition for the example then figure 7.30 changes into figure 7.32. Especially for larger problems this axis definition gives a better insight in the actual flows of load units through the system.

- Equipment pools

As stated at the beginning of this section the structure of the equipment networks and of the equipment pools installed at these networks show through in the load unit sub-model. It especially shows through at those sections at which more than one equipment pool is installed, i.e. at which the equipment optimisation model must choose between to parallel, competing equipment pools.

Suppose, for instance, that we consider transport at a section at which three different equipment pools can be used. Later on it will be shown that an example of such a section is the so-called local section inside internal stack "Nin-Jan Stack". At this local section three pools named A, B and C can be used as shown in figure 7.33. Note that each arc of figure 7.33 represents the flow of load units that is transported by means of exactly one of these equipment pools over the considered equipment routing.

- Transhipment

As discussed earlier at section "Transhipment" of this section transhipment of load units is detected automatically based on the modelled structure of sections and of installed equipment pools. Transhipment is included at a number of points in the load unit sub-model. It is included at those points where load units are to be transhipped from one non-transhipping pool onto another non-transhipping equipment pool. In other words, at those points where a third equipment pool is needed for transhipment. At those points one or more arcs, depending on the number of equipment pools available for transhipment, connect both non-transhipping pools through one or more transhipping pools. This will be illustrated by a small example in which load units are to be transhipped from one MTS trailer onto another one (of another pool) through either a SC or a RS. Refer to figure 7.13.

Figure 7.34 shows some more details. The first detail appears to be that the left-hand side network is a public equipment network while the right-hand side network is a private network at an internal stack.

The second detail appears to be that there are a number of combinations of pools used for transport and transhipment. A total of three different cases can be found. (1) The most straightforward case is transport by both MTS pools and transhipment between these two pools by means of either the straddle carriers (installed at the public network) or by the reach stackers (installed at the private network). Although in most cases it would be more logical to tranship by using the reach stackers, the straddle carriers are available too. This possibly rare situation could have been avoided by not installing the reach stackers at the public network. (2) The second case is actual transport by means of the public straddle carriers and then transhipment --by the same straddle carriers, of course-- onto the private MTS pool. (3) The third case is actual transport by means of the public MTS pool and then transhipment and further transport by means of the reach stackers. Given this configuration of networks and pools, all combinations are

---

As the configuration shown in figure 7.29 does not contain a situation where two non-transhipping equipment pools meet as shown in figure 7.13, temporarily a new configuration will be adapted to clarify transhipment in the load unit sub-model. Refer to figure 7.13 for this configuration.

page 98
allowed to use for the equipment optimisation model. Nevertheless, the probability that certain combination will be used can be manipulated through the preferences specified at equipment routings.

The structure shown in figure 7.34 is the general structure that is encountered at almost any two subsequent sections at which two pools that can not tranship themselves meet. Two exceptions to this rule are sections that are separated (1) by the actual stacking location of an internal stack, or (2) by the local stacking nodes of a bounding element or of a train handle area. These two special types of locations and nodes will be discussed shortly.

- Closed internal transport

*Broken internal transport* is recognised automatically and is changed into closed internal transport. As mentioned before, broken internal transport can occur at the boundary of any private network and at any exchange point. Refer to section "Closed and broken internal transport" of this chapter for more details on this subject.

- Local internal transport and stacking

Several degrees of freedom to model local internal transport and local stacking at bounding elements and at train handle areas have been discussed before; these degrees of freedom will have an important impact on the structure of the load unit sub-model. Note that these degrees of freedom did not exist for internal stacks as both local internal transport and stacking were compulsory at internal stacks. Based on the settings of these options so-called partial networks can be constructed, which are just parts of the entire network that model either a bounding element or a train handle area. These partial networks will be given and explained, along with the partial network for an internal stack.

- Local internal transport and stacking - at bounding elements

At this point of modelling the main difference between bounding elements and train handle areas is, that at bounding elements both local internal transport and local stacking are optional, while at train handle areas only local stacking is optional; local internal transport, on the contrary, is compulsory.

All four possible combinations of the settings for a bounding element are given in figure 7.39, along with the resulting partial networks. Note that for train handle areas only the upper two combinations are available as local internal transport is compulsory for them.

<table>
<thead>
<tr>
<th>Local internal transport</th>
<th>Local stacking</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td><img src="#" alt="Diagram" /></td>
<td><img src="#" alt="Diagram" /></td>
<td><img src="#" alt="Diagram" /></td>
</tr>
<tr>
<td>No</td>
<td><img src="#" alt="Diagram" /></td>
<td><img src="#" alt="Diagram" /></td>
<td><img src="#" alt="Diagram" /></td>
</tr>
</tbody>
</table>

*Figure 7.39: Local internal transport and local stacking at bounding elements.*

Following notes should be taken into account to understand figure 7.39 correctly:

1. The triangle that appears at the left-hand side of each picture in figure 7.39 models the supply to the bounding element of load units entering the system on a certain closed task during a certain time window. For a description of these supplies refer to Appendix C.
2. The arc that connects this supply with the very first node of each picture does not require any form of internal transport and does not consume any time at all. As a matter of fact, this arc is only used to indicate which node is to be supplied and does not need to be included in the final model that is used for the actual calculations; it is just added for reasons of clarity.
3. Any node that connects with a vertical arc models local stacking. Two different arcs stem from such a node: One horizontal and one vertical arc. The horizontal arc models transport of load units that are possibly unstacked and are transported further into the system; the vertical arc on the contrary models stacking of load units through time. This arc will become important when more than one time window is shown.
4. The rightmost solid arc in both upper pictures models local internal transport. This local internal transport at bounding elements was modelled by specifying a fixed distance for internal transport, without the need to expand the private network of the bounding element. Local internal transport is modelled in a rather simple way -- without actually changing the location of equipment and/or load units.
In some cases the private network of the bounding element will be expanded to allow transport directly from the bounding element to another location, using the same equipment network and equipment pools that are used for the actual local internal transport. This option results in a structure that automatically yields closed internal transport. This option will mainly be used to model the situation of a bounding element that is supplied from a "private" internal stack -- i.e. the situation in which the stacking function of a bounding element needs to be modelled at the level of detail of an internal stack, rather than at the level of local stacking. In such a situation normally the same equipment used for local internal transport at the bounding element will be used for transport to its "private" internal stack.

The four possible combinations of local stacking and an expanded private network are given in figure 7.36. Note that if the private network is expanded, then the time needed for normal internal transport from the bounding element to the next location on a certain closed task over the private network (and only for those equipment pools that are installed at that private network) is increased by the standard duration of local internal transport. Equipment pools that are installed at another equipment network than the private network, but which are pinned to the bounding element, will always be modelled according to one of the upper two pictures in figure 7.36. The reason for this is, that they are not installed at the private network and therefore can not be used for local internal transport.

5. The dotted arc models normal internal transport, either over the location's expanded private equipment network, or over any public equipment network.

6. The partial networks shown in figure 7.39 only show the flows for load units entering the system. Obviously, such partial network also exist for load units leaving the system; the only difference is that the flow direction reversed with respect to that of figure 7.39.

**Local internal transport and stacking - at train handle areas**

Local stacking at train handle areas is modelled the same way it is at bounding element, with one difference: While the stacking node at bounding elements was located as remote from the system's boundary as possible, i.e. local internal transport was placed between the stacking node and the system's boundary, the stacking node at train handle areas is located as close to the system's boundary as possible.

The reason for this difference lies both in the aimed level of detail in the function of both types of location. At bounding elements, internal transport was meant to model transport over the sea-quay, and therefore local internal transport must be located between the system's boundary and the possibly requested local stacking facility. At train handle areas, internal transport was meant to model in all cases transhipment from and onto trains, after which load units might be stacked at the local stacking facility next to the rail tracks. Therefore, transhipment from and onto trains is located between the system's boundary and the local stacking (if requested), while transport either directly from the tracks or from the local stacking facility next to the tracks is located after the possibly present local stacking facility.

### Figure 7.36: Effect of an expanded private network on local internal transport at a bounding element (only valid for equipment pools installed at the private network).

<table>
<thead>
<tr>
<th>Private network</th>
<th>Standard</th>
<th>Expanded</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Diagram of local stacking with expanded private network" /></td>
<td><img src="image2" alt="Diagram of local stacking with standard private network" /></td>
<td><img src="image3" alt="Diagram of local stacking with expanded private network" /></td>
</tr>
</tbody>
</table>

### Figure 7.37: Local internal transport and local stacking at train handle areas.

<table>
<thead>
<tr>
<th>Local internal transport</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image4" alt="Diagram of local stacking at train handle areas" /></td>
<td><img src="image5" alt="Diagram of local stacking at train handle areas" /></td>
<td><img src="image6" alt="Diagram of local stacking at train handle areas" /></td>
</tr>
</tbody>
</table>

Invalid
This results in only two possible combinations for local internal transport and local stacking at train handle areas as shown in figure 7.37, as local internal transport is always obligatory at internal stacks. Note that, different from the previous figures, the left-hand side arcs in figure 7.37 model transhipment from and onto the trains, while the right-hand side arcs model local internal transport that possibly is preceded by another transhipment, depending on the type of equipment used for local internal transport.

The equipment optimisation model does not distinct between the various handle bundles at which load units could be stacked; all stacking capacity at all handle bundles is considered as one big stacking capacity. Local internal transport is obligatory at train handle areas. The transport distance, needed to model this local internal transport, is simplified to the average distance from the gate of the train handle area to the most remote king pin of each handle bundle associated to the train handle area.

♦ Local internal transport and stacking - at internal stacks

*Internal stacks* are handled somewhat different than bounding elements and train handle areas for following two reasons. *First*, local internal transport and stacking are both *compulsory* at internal stacks due to their aimed level of detail. *Second*, internal stacks do not reside on the system’s boundary as bounding elements and train handle areas do. By consequence, an internal stack *always* consists of three nodes and two arcs as shown in figure 7.38. The actual stacking of load units is modelled at the mid node. Note that *re-stacking* of load units within the internal stack is not modelled.

♦ Limited stack capacities

Limited stack capacities are not included in the load unit sub-model as stacking arcs with limited capacities, as one might think. The problem of such an approach is, that a limited capacity for any stacking facility holds for the total of load units stacked for all equipment routings. Therefore, limited stack capacities are modelled through side-constraints that ensure that the sum of the load units that are stacked at a stacking facility, summed over all equipment routings that run over that stacking facility, do not exceed its stacking capacity for any time window. Obviously, such side-constraints do not need to be included for those locations that have an unlimited stacking capacity.\(^{16}\)

Now internal transport and local stacking can be included at the various locations in the load unit sub-model. Concerning "Central RSC", suppose that local stacking is allowed and that the capacity is limited to 100 generalised load units; local internal transport will also be allowed as this is obligatory at all train handle areas. Concerning "Nin-Jan Stack" there is little choice: both (local) stacking and local internal transport are compulsory at an internal stack. Concerning "Nin-Jan", suppose that local stacking is not allowed (as there is a separate internal stack for "Nin-Jan"), but that local internal transport along the sea quay of "Nin-Jan" is modelled through a constant distance of 500 metres.

---

\(^{16}\) According to the *limited stacking condition* each closed task must include at least one stacking location with *unlimited* stacking capacity.
In that case figure 7.32 transforms into figure 7.39. This figure shows "Central RSC" as two, and "Nin-Jan Stack" as three separate circles, which is a consequence of the settings for local stacking and for local internal transport at these locations: Each time local internal transport is required in combination with local stacking a new arc for internal transport is required. Marine terminal "Nin-Jan" still consists of one circle only as the extra distance of 500 metres for local internal transport at "Nin-Jan" can easily be added to the standard distance between "Nin-Jan" and "Nin-Jan Stack" for normal internal transport, without the need to create an additional arc.

Equipment pools
The final step before adding time windows to the example is to include the various equipment pools that can be used for internal transport in the load unit sub-model. It is easy to determine which equipment pools can be used for (local) internal transport at the various section transports by following two rules. First, for each section it is known which equipment networks can be used through the pinned network pins. At each of these networks a number of equipment pools of various types will be installed. By consequence, it is known which equipment pools potentially could be used for (local) internal transport at each section. Second, whenever an equipment network is blocked for use at a specific section then none of the equipment pools that are installed at that blocked network is available for internal transport at the section. Third, each of the equipment pools that potentially could be used should be of an equipment type that meets some demands imposed by the load unit sub-model. These demands are that only those equipment pools are interesting that are capable of carrying load units and of moving; They do not necessarily need to be self-propelled nor do they need to be able to transship. These aspects are handled at the equipment sub-model. Following the rules for equipment type definition given at figure 7.1 there is no possibility that any of the equipment pools that are meant here can pull other equipment.

As mentioned before when discussing equipment routings, potentially more than one equipment network can be used for internal transport at a certain section; and potentially one and the same equipment pool can have been installed at these competing networks -- with varying preferences. This gives rise to the question which of these preferences should be adapted for such an equipment pool. This is solved by selecting the highest preference for the equipment pool at any of the competing networks.

So, when adapting the configuration of equipment pools on sections that was given in figures 7.6 and 7.7, and if showing only one equipment routing to keep the figure clear to understand, then a structure as shown in figure 7.40 is obtained. This figure shows for each section which equipment pools can be used for internal transport. Note that three of the five sections can be served by more than one equipment pool; this leaves the equipment optimisation model with the task to select one of them -- or use more than one of them. Also note, that equipment pool A can be used at the private networks of "Central RSC" and of "Nin-Jan" but can not be used on a section between these two locations. Therefore, whenever a machine from equipment pool A has been used at "Central RSC" and after that must be used at "Nin-Jan Stack", the machine must always be relocated without carrying load units. However, this aspect is modelled in the equipment sub-model.

One special case needs some more discussion: The case of pulling and trailer-type equipment. As shown before there is a special way in which trailer-type equipment is linked to pulling equipment. First of all, each trailer-type of equipment can be pulled by exactly one pulling-type of equipment; this pulling-type equipment was stored at the parameter et_ptr of the trailer-type pool. After that, a trailer-type pool may be pulled by any pulling pool of the required type that is installed at the same network as the trailer-type pool is. In this way it is
allowed to model a situation of two trailer- and pulling-type of pools that serve the same sections but may not be mixed mutually. Now such a combination of a pulling pool and a trailer-type pool that are installed at the same network and are related through the value of \( et_{ptr} \) are considered by the load unit sub-model as one, imaginary equipment pool and therefore an additional arc is created for that unique combination of pools.

So, when constructing a situation as shown in table 7.2 in which three trailer-type pools and two pulling pools are installed at two different, overlapping networks, then three different arcs are constructed as shown at the right-hand column of table 7.2. Each of these arcs models a unique combination of trailer-type- and pulling pool. In those cases where a certain combination can be found due to several, overlapping networks that serve the same section and at which the same pools are installed, then still only one arc is created. Obviously, the assumption in this example is that trailer-type pools 1 and 2 may be pulling by pulling pool A, and that trailer-type pool 3 may be pulled by pulling pool B. In the final, found combinations the networks are not discerned any more as such; these networks have only be used to determine the right unique combinations. This aspect will be discussed at the equipment sub-model, too.

We happen to be in the pleasant circumstances that in this configuration of equipment pools there is no need to create additional arcs for transhipment as there are no two subsequent sections\(^{17} \) at which two different trailer-type equipment pools meet. Nevertheless, one special case is encountered in this configuration: Equipment pool D can be used both for (1) local transport at "Nin-Jan Stack" followed by transhipment on the AVGs, or (2) by fully self-supported transport from "Nin-Jan Stack" to "Nin-Jan". Which of the two transport combinations is preferred by the equipment optimisation model depends on various aspects like the preferences specified at the equipment routings, the number of available AVGs, the number of machines in pool D, the number of load units that is to be transported between "Central RSC" and "Nin-Jan Stack" and must be transhipped by pool D due to a lack of machines in pools A and C, ... and so on.

Note that figure 7.40 only displays just one time window. In reality the equipment optimisation method covers a number of time windows. So, the network should not just include internal transport within a single time window (as shown in figure 7.40) but also internal transport across time windows and stacking of load units. At this point inter- and INTRA-window equipment hopping start to differ. In this section, however, only INTRA-window equipment hopping is discussed.

\(^{17}\)Direct transhipment between two sections as meant here is not allowed when the sections are separated (1) by the actual stacking location of an internal stack, or (2) by the local stacking nodes of a bounding element or of a train handle area.
The multiple-time window network, based on solely two time windows, is presented in figure 7.41. These time windows are indicated by "w=1" and "w=2" for the first and second time window respectively. The first time window starts at the current simulation time. The second time window might end at the current simulation time plus the planning horizon, provided that the whole optimisation model covers solely two time windows\textsuperscript{18}. Normally, however, the second time window will end long before the planning horizon has been reached and there will be several more time windows after it.

Figure 7.41 also shows the data on incoming and outgoing flows of load units. These flows, or demands and supplies according to Appendix C, specify the number of load units crossing the system's boundary in a specific time window, over a specific equipment routing. These demands and supplies are determined by the various load schedules, contra schedules and flow patterns. Note that figure 7.41 shows some more "time windows" than the two mentioned before. These "time windows" are the initial window and the final window and play a crucial role in the equipment optimisation model.

The initial window specifies the number of load units that are present at the various locations at the current simulation time. For the way initial stack sizes are determined refer to Chapter 8. The final window specifies the number of load units that have to remain in the system after the planning horizon ends, and therefore is the sink of the load unit network. The number of load units that is demanded at the sink equals the sum of all supplies in the system including those of the initial window minus the sum of all demands in the system. Note that there is only one sink in the system: There is no need to create different sinks for different routings, as flows towards the sink are purely needed to balance the load unit network.

Figure 7.42: Application of closed internal transport for equipment pool D.
Now the aspect of closed internal transport can be included in the model as it appears to apply for equipment pool D; this particular pool can be used both for transport local internal transport at "Nin-Jan Stack" and for

\textsuperscript{18}In reality this could never be the case as the equipment optimisation model consists of three phases, each covering at least one time window.
normal internal transport between "Nin-Jan Stack" and "Nin-Jan" itself. The original and the corrected network are shown in Figure 7.42.

Figure 7.43 shows the incorporation of closed internal transport in the load unit sub-model. Although the difference with Figure 7.41 is slight, it nevertheless is a very important difference. Note that equipment pool D can still serve to get load units out of stack "Nin-Jan Stack" and transship them onto pool B.

Analogous to overlapping equipment pools, closed internal transport leaves the equipment optimisation model with the problem to choose which preference for use of the equipment pool D should be adapted. Unlike at the overlapping equipment pools, there is no way to find an obvious solution from the real-life system. Therefore, it is selected to adapt the preference of the normal internal transport (i.e., not of the local internal transport) as the most likely preference. The reason for this is that in most cases the distance covered by normal internal transport is much greater than that covered by local internal transport, and therefore has the greatest impact on the overall preference.

Although Figure 7.37 only shows closed internal transport, theoretically it still is possible to transport load units by broken internal transport by transshipping them from equipment pool D onto equipment pool D, according to Figure 7.42. In reality closed internal transport will always be less expensive than broken internal transport as it requires less transshipping, and therefore the equipment optimisation model will always prefer closed transport above broken internal transport. Also note that the load unit sub-model does not specify at all how load units are transshipped and what requirements should be met; this will be discussed at the equipment sub-model.

**Stacking costs**
The only costs that are to be included in the load unit sub-model are the costs of stacking load units either at internal stacks either at bounding elements or train handle areas (i.e., through local stacking). Costs of internal transport are included in the equipment sub-model.

The stacking costs are derived from the flow patterns. Although one might expect that the stacking costs per load unit at each location at each equipment routing equal the REMBRANDT value as specified at the corresponding flow pattern, they do not. The reason for this is the logarithmic character of these REMBRANDT values which is fine for multi-criteria analysis but not for an optimisation model as described here. A particular benefit of the logarithmic character of the REMBRANDT method was its natural pay-off behaviour, which is already implicitly present in the equipment optimisation model as network models are always solved in a kind of hierarchic way: First all resources yielding the cheapest solution are used and only then the more expensive resources are used, even if the costs do not differ much. As this already results a pay-off behaviour there is no need to make the differences in costs extremely large as the logarithmic REMBRANDT method does. Therefore, we can simply use linear values as stacking costs according to following table:

---

19 Refer to Appendix A for a description of the REMBRANDT method.
Two special aspects when converting REMBRANDT semantic preference judgements into stacking costs can be seen from Table 7.3. First, the cost sequence is inverse of the preference judgement sequence. In other terms, a high preference judgement yields a low stacking cost, and a low preference judgement yields a high stacking cost. Second, stacking costs equal zero if the preference judgement equals "Indifference". In this way the stacking costs are most analogous to their human interpretation.

Stacking costs are totally independent of the duration of the time window in which they are stacked. This is a necessity, needed to guarantee equal preference of the equipment optimisation model to stack load units at a particular location, for all time windows. This aspect is also found when discussing costs in the equipment sub-model, showing through more clearly than here in the load unit sub-model.

<table>
<thead>
<tr>
<th>semantic preference judgement</th>
<th>stacking costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Very strong preference</td>
<td>-4</td>
</tr>
<tr>
<td>2. Strong preference</td>
<td>-3</td>
</tr>
<tr>
<td>3. Average preference</td>
<td>-2</td>
</tr>
<tr>
<td>4. Weak preference</td>
<td>-1</td>
</tr>
<tr>
<td>5. Indifference</td>
<td>0</td>
</tr>
<tr>
<td>6. Weak negative</td>
<td>1</td>
</tr>
<tr>
<td>7. Average negative</td>
<td>2</td>
</tr>
<tr>
<td>8. Strong negative</td>
<td>3</td>
</tr>
<tr>
<td>9. Very strong negative</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 7.3: Stacking costs corresponding to REMBRANDT semantic preference judgements.

- **Limited stack capacities**

A little while back it was found that limited stack capacities were to be modelled by means of side-constraints rather than by means of limited arc capacities. This was due to the multitude of equipment routings over which the stacking capacity must be divided.

So, when storing the number of load units that is stacked at a location at the end of a time window \( w \) for an equipment routing \( r \) in the location's parameter \( l_e_{\text{sn}_w,r} \), then the limited stack capacity is guaranteed not to be exceeded if the side-constraint:

\[
l_e_{\text{lsco}} \geq \sum_{r} l_e_{\text{sn}_w,r}
\]  

(7.1)

is satisfied. Obviously, this side-constraint only needs to be constructed for those locations that have a limited stack capacity, that is, if the option \( l_e_{\text{lsco}} \) is used. It may be ambiguous which time window is meant by index \( w \), as any flow modelling stacking of load units is related to two time windows: The time window in which the load units are stacked and the successive time window.

The time window referred to by the index \( w \) is the time window in which the stacking activity starts. This rule will also be applied for flows of equipment to be discussed in the equipment sub-model: The index \( w \) will indicate the time window from which the flows of machines originate rather than in which they end. This especially will become important when discussing flows that skip a time window.

- **Meaning of flows**

Flows in the load unit sub-model represent numbers of load units that are transported in time window \( w \) over section \( s \) of a certain equipment routing, by means of machines from equipment pool \( p \). These flows are all stored in variables named \( l_u_{\text{tn}_{w,p,s}} \). Unlike the stacking flows \( l_e_{\text{sn}} \) there is no discussion what the meaning of index \( w \) is for these transport flows \( l_u_{\text{tn}} \) as these latter do not cross a time window and therefore only flow through one time window. Therefore, index \( w \) indicates the time window in which the transport flows start and end as well.

- **Initial flows**

The initial flows of the load unit sub-model specify the number of load units that are present at all locations at the current simulation time, i.e. at the simulation time at which the equipment optimisation model is executed. For those load units that are stored at some location there is no difficulty in finding the initial flows: These just equal the number of load units stored at each location. These numbers of load units can be used as the first part of the supply values of the load unit sub-model in the very first time window, along with the numbers of load units that are to enter the system at that very location during that first time window.
The second part of the supply values of the load unit sub-model is determined by the numbers of load units that are being transshipped or transported at the current simulation time, and need to be registered at some location in some time window, too. Therefore the location and time window to supply the load units to should be determined. This problem is solved by using the expected time at which the currently executed transport or transhipment will end. This time is estimated as the starting time of the transport- or transhipment movement plus the so-called section transport time, which will be discussed at the equipment sub-model, of that transport- or transhipment movement. The thus found estimate of the ending time of the transport- or transhipment movement determines the time window in which the load unit is registered as a supply. An example hereof is given in figure 7.44 showing an arc that ends in time window "w=1" and therefore results in the assignment of the transported or transshipped load units at that same time window "w=1".

Note in this case the remaining transport time of the load unit during the time window "w=1" is ignored by the equipment optimisation model. Although a threshold could be introduced to minimise the amount of time available in the time window for transport after the actual, currently being executed, transport has ended, this has not been chosen for. The reasons for this choice are (1) the ambiguity of the amount of time to select as a threshold, and (2) the danger of blocking transport of the load unit on a subsequent, short section, thus possibly introducing an infeasible solution of the optimisation problem without any reason.

In case of pure transhipment the duration of the arc in the equipment sub-model in general will be thus short measured in time that it will point to the very first time window of the equipment optimisation model. In case of actual transport however, possibly combined with transhipment by the transporting equipment, there is a greater chance to enter the second or third time window.

7.4.3.2. Equipment sub-model

While the previous section described the flows of load units through the system for INTRA-window equipment hopping, this section describes the flows of equipment through the system. Although strongly related to each other (load units can not move without the help of equipment) there are some differences between these flows. Actually, these differences make the equipment sub-model a little more complex than the load unit sub-model.

♦ What is modelled?
The equipment sub-model covers some details that have not been found in the load unit sub-model. These details can be divided in (1) details on physical abilities of equipment like transport speeds, acceleration times, transhipment times and carrying capacities, (2) organisational aspects like preferences for pools, working schedules, blocked networks, overlapping networks, and equipment insertion points, and (3) geometric data on the lay-out of the modelled system, which is important to model distances for transport and relocation. Stacking of load units, on the other hand, is not included at all in the equipment sub-model and therefore is purely a concern of the load unit sub-model.

Obviously, the equipment sub-model also covers costs, like the load unit sub-model did. The main difference is, that the equipment sub-model covers both operational (variable) and investment (fixed) costs, while the load unit sub-model only covered variable costs.

♦ Pools for internal transport
The structure of equipment pools available for various transports in the equipment sub-model is equal to the structure described at the load unit sub-model. So, wherever an arc for transhipment or transport of load units by some pool on some equipment routing appeared in the load unit sub-model, a copy of the arc appears in the equipment sub-model. At least, seems to appear, as in reality each arc in the load unit sub-model will yield more
than one arc in the equipment sub-model, needed to include the aspect of equipment relocation sufficiently detailed in the equipment optimisation model.

Special concern is given to equipment types that can carry more than one load unit, the so-called MLU-types. They will be related to so-called dummy-load units rather than to actual load units. Special concern is also given to pulling equipment and trailers, resulting in a remarkable structure in which flows of (dummy-) load units are assigned to the pulling equipment rather than to the trailers. A special role will be given to non-movable equipment types that can tranship and are installed at the private networks of train handle areas and of internal stacks. In the remainder of this section these aspects will be discussed in detail.

♦ **Getting started**

One major aspect to bear in mind when creating the equipment sub-model of INTRA-window equipment hopping is, that all flows through this sub-model represent the average number of equipment (machines) to be used during the time window at a certain section of a certain equipment routing, of a certain equipment pool. The equipment sub-model does not specify when these machines exactly should be used. Instead, it only specifies in which time window they should be used. Furthermore, unlike the load unit sub-model, the equipment sub-model is a real-type network, which means that flows on the network do not necessarily need to be integers. This is a necessity in order to match the use of equipment, which is subject to varying working hours, various speeds, and non-single carrying capacities, with the flows of load units which are (and must be) integer-type flows; as each load unit is traced as detailed as possible through its network. By consequence, the load unit sub-model is the leader while the equipment sub-model is the slave, i.e. the equipment sub-model has to match to the load unit sub-model -- and not vice versa.

♦ **Transhipment**

Analogous to the load unit sub-model, transhipment of load units is included in the equipment sub-model, too. Nevertheless, as this subject is more strongly related to the equipment sub-model discussed here than to the load unit sub-model, it will be discussed in more detail over here. It will be discussed what equipment pools can be selected for transhipment, what rules apply for modelling gantries at internal stacks and train handle areas, and what effect closed internal transport, local internal transport and local stacking have on the equipment sub-model.

♦ **Transhipment - at bounding elements**

First of all, transhipment at bounding elements will be discussed. As mentioned earlier two major parameters determine the character of a bonding element in the equipment optimisation model: The presence of local stacking and the presence of local internal transport. Whether or not these two parameters are used is entirely up to the user of the terminal design aiding tool, and depends on the aimed level of detail. Creating closed internal transport (by installing the same pools on neighbouring networks) yields the same configuration of equipment assignment, and therefore the same structure of the load unit- and the equipment sub-model as an expanded private network of a location would have.
Depending on the settings for local internal transport, local stacking and closed internal transport five different configurations for internal transport and transhipment at bounding elements are found (compare with figure 7.39). All of these configurations are shown in figures 7.45 and 7.46 in which transhipment is shown as "\( \Rightarrow \).

![Figure 7.45: Effect of various settings on the transhipment of load units at a bounding element (transhipment is shown as \( \Rightarrow \).)

Three details can be seen from figures 7.45 and 7.46. (1) First, in configuration B load units are transhipped directly from the private network onto the next public network. As a matter of fact, this transhipment is equal to the one at an exchange point. This type of construction is used when a strong relation between quay-side internal transport (which is fed by load schedules and contra schedules) with the remaining internal transport system is expected to be important in the modelled terminal system. (2) Second, in configuration C the internal transport distance is simply incorporated in the actual internal transport distance over the connecting public network. In other terms, the transport distance over the connecting public network is simply increased by the amount of the local internal transport distance. Local internal transport at bounding elements does not actually change the position of any load unit or machine but merely is an abstract extra transport distance. (3) Third, in case of local internal transport the transhipment at the sea-quay is also included in the model.

For each of these transhipment movements different rules apply to select the equipment pools that may be used to those movements. This will be explained by referring to figures 7.45 and 7.46. A distinction between three types of transhipment, shown as 1, 2 and 3 in figures 7.45 and 7.46 must be made: Transhipment 1 is always located at the sea-quay and always uses equipment installed at the bounding element's private network. There is, however, one special case for this transhipment. Whenever an equipment pool has been installed at the bounding element's private network that can tranship but can not move, then that equipment pool is taken as the single pool of quay-crane(s) of the single installed non-movable transhipment pool; it is not allowed to install more than one pool of this type at the bounding element's own private network. Transhipment 2 always uses equipment pool installed at the bounding element's private network, while transhipment 3 may use pools installed either at the bounding element's private network or at the neighbouring network. Whenever no pools that can tranship but can not move are installed at the bounding element's private network then there is no difference between the pools that can be used for both transhipments 1 and 2.

---

20 For an expanded private network, which is equal to closed internal transport.
21 The vertical arc representing stacking is only shown for reasons of clarity; it is not included in the actual equipment sub-model.
22 For an expanded private network, which is equal to closed internal transport.
23 This case will be elaborated in more detail for internal stacks.
24 This type of structure will be found at internal stacks and at train handle areas, too.
Chapter 7. Equipment

Note that each transhipment shown in figures 7.45 and 7.46 may result in any of both structures as shown in figure 7.47. In the upper structure transhipment is done by a pool of straddle carriers named SC onto a pool of trailers which take care of the actual transport movement. In the lower structure both transhipment and the actual transport movement are done by the straddle carriers.

Also note that configuration B could be modelled according to any flow shown in figure 7.46, based on figure 7.17. In this situation a pool of straddle carriers named "SC" and a pool of trailers named "Trailer 1" is installed at the bounding element's private network, while a pool of reach stackers named "RS" and a pool of trailers named "Trailer 2" is installed at the connecting network. This structure yields four different flows of load units, corresponding with four different assignment sequences of equipment.

\* Transhipment - at train handle areas

Transhipment at train handle areas is analogous to transhipment at bounding elements with one major difference: The relative position of local internal transport and local stacking. While local stacking at bounding elements was separated from the system's boundary through local internal transport, local stacking at train handle areas is always located at the system's boundary. As mentioned before this is a consequence of the difference in meaning of local stacking at both types of locations.

Depending on the settings for local stacking and closed internal transport four different configurations for transhipment can be found as shown in figures 7.49 and 7.50 which also show that local internal transport is obligatory at train handle areas. Transhipment \( \oplus \) in figures 7.49 and 7.50 represents transhipment onto and from trains and therefore is always included in a modelled train handle area. Analogous to transhipment \( \oplus \) at bounding elements, transhipment \( \oplus \) at train handle areas is a little different from the other transhipment movements concerning pool selection: Whenever a pool of gantries is installed at the train handle area's private network, and one is allowed at most, then this equipment pool is selected as the only pool that may be used for transhipment \( \oplus \). Gantries are automatically recognised through their modelled capabilities: They are able to tranship but are not able to move. Transhipment movement \( \ominus \) can be done by any transshipping pool installed at the train handle area's private network. So, it is possible to model a situation in which movement \( \ominus \) is done by gantries and movement \( \ominus \) is done either by straddle carriers or by the same

\[25\] or an expanded private network, which is equal to closed internal transport.
gantries. In that case two details need to be considered. First, it must be safeguarded that equipment like straddle carriers can enter the local stacking area with the gantries without disturbing these gantries, which in practice only occurs when gantries are equipped with an outreach allowing the straddle carriers to pick up the load units. Second, if the only two equipment pools installed at the train handle area’s private network are these straddle carriers and gantries, then transshipment ③ will only use the straddle carriers as the gantries require the presence of non-transshipping movable equipment between transshipment ② and ③. In other terms, gantries can not be used for transshipment ② unless some trailer-type equipment is installed at the train handle area’s private network.

**Transshipment - at exchange points**
Transshipment at exchange points is completely analogous to transshipment movement ③ in configuration D of train handle areas, and represents direct exchange of load units between two equipment pools. As exchange points do not own a private network, and therefore do not have the possibility to install equipment, any transshipment at any exchange point must be done by means of an equipment pool that is installed at a network connected to the exchange point. In case of closed internal transport at an exchange point, the function of that exchange point is eliminated completely as can be seen from the left-hand side of figure 7.51.

Any transshipment structure as shown in figure 7.34 can be used for transshipment ③ shown in figure 7.51, so transshipment between two trailer-type pools through any movable transshipping pool from one of the connected networks, and transshipment between one trailer-type pool and a movable transshipping pool.

**Transshipment - at internal stacks**
Transshipment at internal stacks is the most complicated of all types of locations due to the relative large number of transshipment movements: At most four different transshipment movements can be found at an internal stack. This is caused by the situation of the internal stacks off the system’s boundary resulting in an additional transshipment with respect to locations on the system’s boundary for load units entering the stack.

This is shown in figure 7.52. As stacking at internal stacks is compulsory, like local internal transport is, transshipment movements ② and ③ are always included in the model of an internal stack. The presence of transshipment movements ② and ③ depends on the presence of closed internal transport, either through identical installed equipment pools at neighbouring networks or through an expanded private network of the internal stack. Any combination of closed internal transport when entering or leaving the internal stack is allowed, as can be seen from figure 7.52.

Analogous to bounding elements and train handle areas some special rules apply for pool selection at internal stacks. As the situation at internal stacks is the most complicated of all, these rules will be elaborated now in more detail for internal stacks. Whenever it is spoken of “stack blocks” in the remainder of this section, it may also be replaced by “handle bundles” for train handle areas, as the general structure is almost the same for both types of locations with two major differences: (1) At internal stacks load units are stacked at stack blocks, while at train handle areas load units are stacked at handle bundles, and (2) at internal stacks it is not possible to assign
specific equipment pools to specific stack blocks, while at train handle areas it is allowed to assign a different equipment pool to each handle bundle.

As can be seen from figure 7.52 (slightly modified, partially repeated in figure 7.53) load units are transshipped at most at four places at an internal stack. But, as transshipment movements ✈ and ◐ can be considered as the reversed of transshipment movements ✈ and ✐, the same rules apply for both pairs of transshipment movements; only the latter pair of movements is discussed. This is also shown in figure 7.53 in which transshipment movements I and II (shown in thick lines) are the reverse of transshipment movements shown in dotted lines. Whenever closed internal transport is applied, the leftmost and the rightmost transshipment movement of figure 7.53 do not exist.

The main difference between transshipment movements I and II are the equipment pools that are available: For transshipment movement II only those pools installed at the internal stack's private network are available, while for transshipment movement I both the internal stack's private network and any network connecting to that private network are available26. These networks connecting to the internal stack's private network have been used for transport of load units towards the internal stack27.

Only two cases can occur for the organisation of all internal transport shown in thick lines in figure 7.53. Both cases refer to broken internal transport. The main difference between these two cases is, whether the equipment pool moving the load units from transshipment I to transshipment II is able to transship itself. If it is then this (local) equipment pool is to take care of all internal transport shown in thick lines in figure 7.53. This case is shown in the upper part of figure 7.54. If it is not then this (local) equipment pool is to take care solely of the actual transport movement between both transshipment movements; transshipment movement I is taken care of either by a pool from some network previous to the internal stack, or another pool from the private network of the internal stack. Obviously, such a pool must be able to transship. Transshipment II is taken care of by an equipment pool from the private network of the internal stack that can transship. This case is shown in the lower part of figure 7.54.

Although both for transshipment movement I and II in the lower part of figure 7.54 an equipment pool of the internal stack's private network can be used, following difference between both transshipment movements exists: The private equipment pool for transshipment movement I must be able to move, while the private equipment pool for movement II needs not to be able to move. This is clarified in table 7.4, showing all allowed combinations.

<table>
<thead>
<tr>
<th></th>
<th>Can transship</th>
<th>Can transship and can move</th>
</tr>
</thead>
<tbody>
<tr>
<td>Movement I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement II</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 7.4:** Requirements for equipment shown in lower part of figure 7.54.

---

26 That is, for as far as these connecting networks may be used for transport over the equipment routing that is considered, by means of blocking some of these networks may be excluded for a particular equipment routing.

27 By consequence, if closed internal transport is used, then the internal stack's own private network may be used for this transport and transshipment I is not defined.
Transhipment - at internal stacks - fixed gantries at stack blocks

The structure from table 7.4 is used to model a situation is shown in figure 7.55 in which transhipment I is done by a straddle carrier, transport between transhipment I and II by a trailer system, and transhipment II by a fixed gantry crane. This latter gantry crane is modelled as an equipment type that can tranship but can not move. This immediately gives rise to the question how this gantry crane can store a container in a stack block, or even retrieve a container from a stack block as it can not move. The explanation lies in the fact that non-movable transhipment equipment that is installed at the private network of an internal stack is to be interpreted differently than expected: They are to be interpreted as gantries that are installed at the stack blocks, can not leave the stack blocks, but still can drive along these stack blocks. So, installing non-movable transhipment equipment at an internal stack's private network is a "trick" that is used to model gantries installed at stack blocks.

![Figure 7.55: Example of broken internal transport using non-movable equipment for transhipment movement II at an internal stack.](image)

Note that figure 7.55 shows a situation that will seldomly occur in practice: It will rarely happen that a private pool of trailers will be used for transport inside the internal stack combined with another pool of trailers that is used for transport outside of the internal stack. As can be seen from figure 7.55 this induces an additional transhipment movement by equipment like a straddle carrier, and therefore increases costs.

A more realistic situation can be obtained by using closed internal transport which is shown in a schematic way in figure 7.56. After closed internal transport has been used the same equipment pool can be used for transport inside and for transport towards the internal stack. This can be obtained by installing the same equipment pool at the internal stack's private network and at one of the connecting networks, connecting to the internal stack.

![Figure 7.56: Closed internal transport at an internal stack.](image)

The result of using closed internal transport is shown in figure 7.57. It shows how a trailer system can be used both for transport towards and inside the internal stack, where it is handled by a pool of straddle carriers. In

---

28 Or at the handle private network of some handle bundle belonging to a train handle area.
29 This pool must be installed at the internal stack's private network.
Chapter 7. Equipment

despite this case for gantry cranes, modelled as non-movable transhipment equipment, are installed at the internal stack's private network.

![Diagram of internal stack transport](image)

**Figure 7.58:** Example of broken internal transport at an internal stack.

If the trailer system of figure 7.57 was excluded from the internal stack's private network then a situation as depicted in figure 7.58 would have been created in which the trailer system is not allowed to transport load units within the internal stack. Instead, the straddle carriers are the only machines available for local transport within the stack.

![Diagram of internal stack transport](image)

**Figure 7.59:** Example of closed internal transport using non-movable equipment for transhipment movement II at an internal stack.

Another situation, which is a mix of the situations shown in figure 7.55 and 7.57, is given in figure 7.59. It shows a situation of closed internal transport in which the only transhipment equipment installed at the internal stack is a non-movable transhipment type, modelling a gantry crane. In this situation there is no need to install any other movable transhipment pool at the internal stack's private network as the gantries take care of all transhipment at the internal stack.

Returning to the non-movable transport equipment for movement II of figures 7.53 and 7.54, some other detail needs to be discussed: The relation between non-movable transhipment equipment and stack blocks; refer to Chapter 8 for a description of these stack blocks. This relation is different for both stacking modes that are available according to Chapter 8: Detailed stacking and simplified stacking. Both stacking modes will be discussed in brief.
In case of detailed stacking an internal stack is divided into a number of stack blocks. To each of these stack blocks different, individual gantry cranes can be assigned which all stem from the same non-movable transhipment pool. From the real-life system this can be seen as a gantry that is installed at a certain stack block and can only be used to transship load units into and out of that single stack block.

<table>
<thead>
<tr>
<th>First internal stack</th>
<th>Installed “fixed” gantries</th>
<th>Stack blocks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="image1" alt="Diagram" /></td>
<td><img src="image2" alt="Diagram" /></td>
</tr>
</tbody>
</table>

**Figure 7.60:** Example of fixed transhipment pools installed at stack blocks of two different internal stacks, in case of a pool capacity of two.

From this concept two consequences can be derived: (1) Each internal stack may have only one non-movable transhipment pool associated to it; otherwise it would be ambiguous which type of machine is to be assigned to each individual stack block. (2) The capacity of the non-movable pool represents the number of machines installed per stack block rather than the total number of machines in that pool. Based on this consequence, one further consequence can be drawn: Any non-movable equipment pool that is installed at more than one internal stack will have the same number of machines installed at each stack block of each internal stack. This is clarified by figure 7.60 showing a situation in which the capacity of a pool of “fixed” gantries equals two, and which pool is installed at two different stacks; the first internal stack consisting of one stack block, and the second internal stack consisting of three stack blocks. As the individual stack blocks are not known to the equipment optimisation model, the capacity of a non-movable transhipment pool at an internal stack gets known to the equipment optimisation model simply as a single capacity which equals the user-specified pool capacity times the number of stack block of the internal stack.

<table>
<thead>
<tr>
<th>First internal stack</th>
<th>Installed “fixed” gantries</th>
<th>Stack blocks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="image3" alt="Diagram" /></td>
<td><img src="image4" alt="Diagram" /></td>
</tr>
</tbody>
</table>

**Figure 7.61:** Example of fixed transhipment pools installed at two different internal stacks, in case of a pool capacity of two and simplified stacking.

In case of simplified stacking the entire internal stack is considered as one single stack block, and by consequence the capacity of the non-movable transhipment pool specified by the user will be applied for the internal stack as a whole. This is clarified by figure 7.61 showing a situation in which the capacity of a pool of “fixed” gantries equals two, and which pool is installed at two different stacks; both internal stacks are modelled through simplified stacking. By consequence, the “stack blocks” at these internal stacks in reality do not exist, and therefore are shown in dotted lines in figure 7.61. Any mix of figures 7.60 and 7.61 is valid, i.e. installing a non-movable transhipment pool at any mix of detailed or simplified internal stacks.

---

30: This in contrast to gantries at handle bundles at train handle areas, which all may have a different pool of gantries or other machines assigned to them.
Chapter 7. Equipment

If the capacity of a non-movable transhipment pool is set to "unlimited" then the total capacity per internal stack will, of course, be unlimited.

In case a non-movable equipment pool is installed at an internal stack (and at most one is allowed!) then transhipment movement II in the lower part of figure 7.54 can only be done by this equipment pool; there is no way that another, possibly movable, equipment pool installed at that internal stack can take care of this transhipment, too. This is in coherence with normal real-life operations, where transhipment at any stack block is either done by a gantry installed at that stack block, or by a movable machine like a straddle carrier that can freely move around at the internal stack.

♦ Transhipment - at internal stacks - pool selection summary
In short, the rules for the selection of equipment pools in case of broken internal transport for use on the various internal transport movements shown in thick lines in figure 7.54 are:
1. if local internal transport is done by an equipment pool that can carry, move and tranship, then this pool takes care of all movements, as shown in the upper part of figure 7.54;
2. otherwise, transhipment movement I as shown in the lower part of figure 7.54 is taken care of:
   • either by a movable transhipment equipment pool installed at the private network of the internal stack
   • or by a movable transhipment equipment pool installed at the network connecting to the private network of the internal stack;
   the actual local internal transport within the internal stack is taken care of by an equipment pool installed at the private network of the internal stack that can carry and move, but can not tranship.
   transhipment movement I as shown in the lower part of figure 7.54 is taken care of:
   • either by a movable transhipment equipment pool installed at the private network of the internal stack provided that there is no non-movable transhipment pool installed at that network.
   • or by a non-movable transhipment equipment pool installed at the private network of the internal stack.
   At most one non-movable transhipment pool may be installed at any internal stack.

In case of closed internal transport the rules for the selection of equipment pools for use on the various internal transport movements shown in thick lines in figure 7.56 are:
1. local internal transport is always (by definition of closed internal transport) done by an equipment pool that can carry and move, but can not tranship, and that is installed both at the private network of the internal stack and on the network connecting to that network.
2. transhipment movement II as shown in figure 7.56 is taken care of:
   • either by a movable transhipment equipment pool installed at the private network of the internal stack provided that there is no non-movable transhipment pool installed at that network.
   • or by a non-movable transhipment equipment pool installed at the private network of the internal stack.
   At most one non-movable transhipment pool may be installed at any internal stack. In that case any movable transhipment pool installed at the private network of the internal stack will not be used for closed transport.

Another detail of non-movable transport equipment for movement II of figures 7.53 and 7.54 is the time needed to place a load unit in the stack (i.e. to "move" the transhipment and the load unit in the stack block, although the equipment officially is not able to move) to the proper position in the stack block in the time to tranship. This means that the time to move in the stack should be included either in et_at, et_ut, et_dt or et_rt of the equipment's type. Obviously, this time is a rather rough estimate of the actual value, even more as the precise stack block is not known to the equipment optimisation model.

♦ Transhipment - at internal stacks - constructing partial networks
Based on the just discussed configurations for internal transport at internal stacks the networks modelling those internal stacks at INTRA-window equipment hopping can be found. These (partial) networks are shown in figure 7.62.
Figure 7.62: Scheme to construct the equipment sub-model for an internal stack (excluding equipment relocation; only incoming transport).

Equipment relocation

Until now it has not been spoken of equipment relocation yet, an aspect that will turn out to have a major impact on the appearance of the equipment sub-model: It will increase the complexity of the equipment sub-model considerably. Nevertheless, it is required to create a model sufficiently accurate, especially for the first phase of the equipment optimisation model.

The need for inclusion of equipment relocation will be illustrated with two examples. (1) First, consider a large-scale terminal complex. Suppose that during a certain time window at a certain location a lot of equipment is required, but that only 80% of this requirement can be satisfied by using equipment that is already present at that particular location. Also suppose that some more machines that can be used at this location can be found at another location in the terminal complex, which is about 5 kilometres away. Now the equipment optimisation model should decide whether to drag these machines for 5 kilometres to come and help, and thereafter to drag these machines possibly back again over 5 kilometres, or to mud around with only 80% of the required machines, leaving some less important load unit flows for what they are during that time window. This shows that large relocation distances can have a major impact on the flows through terminal complexes. (2) Second, consider the situation that a small transport cycle at a bounding element of about 500 metres models transport between the sea-quay and the boundary of the bounding element. Suppose that a large sea-vessel --the only one at that bounding element-- is being unloaded. By consequence, al machines serving that cycle will have empty return trips. Omitting the aspect of relocation will yield that the equipment optimisation model assigns (in the worst case) only the half of the machines that actually is required. This shows that empty return trips, which are modelled through equipment relocation, are important to establish the actually needed number of machines. Equipment relocation is less important in those cases where only very little machines are assigned to internal transport during a certain time window, relative to the time window's length.

Equipment relocation arcs model the relocation of machines of one equipment pool. In other terms, it is the redistribution of the machines of an equipment pool over the network, taking into account travelling distances. Therfore, it will have the character of a set of arcs related to one equipment pool, connecting the various
locations between two subsequent time windows at which the equipment pool can be used. In case of INTRA-equipment hopping this is the relocation between time windows.

So, for example a small network as shown in figure 7.63 displays the relocation of equipment that is used for transport between two locations, A and B, on two equipment routings, R1 and R2, in two time windows, "w=1" and "w=2". Each location in each time window (except the last time window) yields one extra node which is displayed slightly greyed; this node is used to gather all machines of a certain pool that have been used during one time window at that location, on all equipment routings running over that location. Therefore, this additional node is called the collection node of an equipment pool at a location, at one time window. All machines that have been gathered at the collection node in a certain time window can be re-distributed over other nodes that are served by the equipment under consideration for use in the next time window.

Figure 7.63 gives a small pre-view of coupling of the load unit sub-model and the equipment sub-model that is still to be discussed: All horizontal lines belong to the load unit sub-model, and thus show the flow of load units. The vertical and slant flows belong to the equipment sub-model and thus show the flow of equipment through the model.

Figure 7.63 should be read as follows: Each arc that represents equipment relocation also represents internal transport of load units. This gives rise to the question from which location these load units are to be transported, as the network only shows from which location the equipment stems and to which location it drives. It is simple to derive the locations between which the load units are transported, as each equipment relocation is related to one single equipment routing, like all flows in the equipment sub-model are. By consequence, if it is known to which location a machines drives, and if it is also known over which equipment routing it drives, then it becomes immediately known which location is located before the machine's destination, and therefore which location is the origin of the load units. This is also clarified in figure 7.64 showing how equipment relocation from location A to location B over routing R1 actually consist of following two parts: (1) Empty driving of the machine to location C from which the load units must be transported, and (2) full driving from location C to location B, carrying the load units to their destination.
Empty transport is only included in the model as part of equipment relocation; empty transport is not included as such, without a strong relation to a future actual transport movement. This is selection is considered logical as there is no point in empty transport without the perspective of transporting load units immediately after that empty transport.

One special type of empty transport is found at the so-called equipment supply nodes: For each equipment pool in the modelled system one such an equipment supply node is automatically "created" and used during optimisation. All newly created machines of an equipment pool use that equipment supply node as their equipment entry point. As a matter of fact this equipment supply node represents the number of machines that is added to the equipment pool during equipment optimisation in case of a lack of transport capacity, and therefore can be regarded as an additional supply of equipment capacity. Transport from this equipment supply node to the first actual transport of a newly created machine is always empty transport. These equipment supply nodes will be given more attention later on.

Each collection node yields a number of incoming arcs that equals the number of routings running over the associated location, and a number of outgoing arcs that equals the number of locations at which the equipment pool may be used times the number of equipment routings the equipment pool may serve. This is clarified in figure 7.66, showing a situation in which equipment of one equipment pool that has been used at location A for two different equipment routings, is re-distributed over three locations (including A) for three equipment routings. Although this equipment pool can not be used for equipment routing R3 at location A (either as this routing is not connected to A, or this equipment is not allowed to serve R3 at location A), it can be re-distributed for use at routing R3 at locations B and C.

The number of locations meant right here may be a little greater than the actual number of UCM-locations because of additional "locations" that are introduced by local internal transport. Whenever local internal transport is modelled, then equipment can move within a location, and by consequence can relocate, too. In other terms, one location that causes internal transport will yield at least two nodes as shown in figure 7.66, which are all distinct geographical positions. By consequence, relocation between these nodes should be included in the model.

\[^{31}\text{i.e. those UCM-locations that introduce an equipment break.}\]
Consider, for instance, an internal stack. As found in figure 7.52, this internal stack could be modelled through a number of additional nodes, depending on the network structure and the equipment pools installed at these networks. Consider the "worst case" situation, in which internal transport at the internal stack uses an MTS trailer system, and two pools of straddle carriers. Suppose that we consider the situation that this holds both for incoming and outgoing internal transport. In that case the situation would be as shown in figure 7.67 modelling the internal stack as five nodes and six arcs (including the arcs modelling the AGVs that transport load units to and from the internal stack, but excluding the stacking arc). Note the parallel arcs for the two straddle carrier pools which compete with each other. Although figure 7.67 shows the situation for one equipment routing only, namely R1, the situation for the reversed routing, R2, is completely mirrored with respect to figure 7.67.

The network for the AGVs would be as shown in the lower part of figure 7.68, showing not only the relocation for the internal stack (in thin lines) but also for two other locations B and C (in dotted lines) that are not shown in figure 7.67. Both incoming and outgoing flows of AGVs refer to the same collection node of the internal stack, which in fact is quite logical as geographically they do refer to the same location.

When taking a close look at figure 7.68 it may be found that the complexity of this figure is rather high for such a relatively simple case. On the other hand, if equipment relocation, which is very important for the accuracy of the results of the equipment optimisation model, is to be included in the model, than such a complex structure is inevitable.
The network for MTSs would be as shown in the lower part of figure 7.69, provided that they are not used outside the internal stack. Unlike the network of the AGVs, the network of the MTSs contains two collection nodes for the same internal stack; this is the aspect of local internal transport at the internal stack, i.e. the different locations at the boundary of the internal stack and the actual stacking location.

The network for the SCs is (for each of the two SC pools) equal to the network of the MTSs, with one difference: while the MTS network modelled transport (influenced by transhipment), the SC network purely models transhipment and relocation.

These rules do not apply to internal stacks only, but also to bounding elements at which internal transport is allowed (excluding closed internal transport). So, when giving the networks for all equipment pools as installed at the various networks of figure 7.6 for the two equipment routings as shown in figure 7.4, then figures 7.93 through 7.73 are obtained. Unfortunately, these figures would become too complex to read when all relocation relations would be included in the model. Therefore only the relocation relations stemming from the boundary of "Central RSC", the boundary of "Nin-Jan Stack" and of "Nin-Jan" itself are fully elaborated.

---

32 This figure purposely does not show (1) transhipment at the actual stacking location and (2) local internal transport by the straddle carriers only. Compare with figure 7.34.
Figure 7.70:  Equipment network for AGV equipment pool B from figure 7.7 (incomplete)

Figure 7.71:  Equipment network for SC 1-over-2 equipment pool A from figure 7.7 (incomplete).
Figure 7.72: Equipment network for SC 1-over-4 equipment pool C from figure 7.7 (incomplete).

Figure 7.73: Equipment network for SC equipment pool D from figure 7.7 (incomplete).
The way relocation is modelled quantitatively is discussed at the section dealing with the coupled load unit- and equipment sub-model. The way equipment relocation is modelled requires to discuss one special case that is to occur especially in the first phase of the equipment optimisation model in combination with a large terminal complex. This is the case that the time needed to relocate equipment from one location to another equals or exceeds the duration of the time window. In that case something odd happens: an entire time window is skipped (or potentially over two or more time windows) to model equipment relocation properly. This is explained through an example in which the duration of a time window is 15 minutes, and the relocation time equals 15 minutes, too. In that case the machines must relocate as is shown in figure 7.74, skipping exactly one time window during relocation. This means that whenever a machine comes available after time window 1 (i.e. may have been used during time window 1) then it will not be available for transport starting at the location just before location B on routing R1, and ending at location itself B on routing R1 in time window 2; but it will be available in time window 3.

Figure 7.74: Skipping one time window during relocation.

7.4.3.2.1. Multiple load unit equipment types

The equipment network that has been given until now is apt for all equipment types that carry precisely one load unit at the time. Normally, these are the types of equipment that will be used for internal transport at most terminals. Nevertheless, in some cases special equipment is used that is able to transport more than one load unit at the time. This type of equipment is called multiple load unit equipment, abbreviated to MLU-types. Some good examples hereof are the Multi Trailer System (MTS) and double stacking trailers.

Multiple load unit equipment types introduce some specific problems. As mentioned earlier, the flows of equipment through the equipment sub-model are based on the flows of load units in the load unit sub-model, where the first is to follow the latter. When considering single load unit equipment types only, then each machine that is planned for internal transport can be related to exactly one load unit in the load unit sub-model. When considering multiple load unit equipment types on the other hand, then it is found that in some cases these equipment types will transport their full capacity of load units, and in other cases they will transport less than their full capacity. Without any further measures we could easily end up in a situation in which multiple load unit equipment is split into two or more parts, thus yielding

---

33 Actually, a time window is skipped if the transport-time, which is to be discussed at the coupled models, exceeds available working hours available in that time window. This will be discussed in more detail later on.
an invalid model. An example hereof is given in figure 7.75 showing an MTS trailer system consisting of three trailers, which in the real life system always remain coupled and therefore need to be considered as an inseparable unit in the model. However, if no special measures are taken then the three trailers may be split up by the equipment optimisation model into separate trailers, as shown in the right-hand part of figure 7.75. In this example 2/3 of the original trailer system is used for one transport movement and the remainder 1/3 of the original trailer system is used for another transport movement. Obviously, this is in conflict with the real-life system and therefore needs to be avoided. One important conclusion from this example is, that the problem of invalid splitting of multiple load unit equipment only occurs in those situations where the multiple load unit equipment transports less than the equipment's capacity. Or, in other terms, this invalid situation only occurs if the assigned flow of load units is less than et_ce of the considered equipment pool.

Fortunately, a solution exists to prevent this invalid splitting of multiple load unit equipment into several individual pieces of equipment. This solution is to assign so-called dummy load units to the multiple load unit equipment, such that each transport movement with any such equipment type is a one hundred percent full transport movement. In other terms, the transport capacity that is not used to transport the actual load units from the load unit sub-model is always used to transport dummy-load units over the same section as the actual load units are. These dummy-load units may not be mixed with the actual load units as they do not exist in the real-life system; they are solely introduced for a "trick" to prevent splitting of multiple load unit equipment. Therefore, they have no physical meaning at all.

Suppose that we consider a situation as depicted in figure 7.76 in which an MTS trailer system can be used to transport load units from location A to location B over two equipment routings, namely R1 and R2. Figure 7.76 also shows that these MTS trailers can be relocated from locations A through C before they can be used at location A. Remember that such a situation would be modelled in the equipment sub-model as a set of 6 arcs, connecting locations A through C for transport of load units from A to B for both equipment routings. Nevertheless, these six arcs are not shown properly in figure 7.76 to make it more easy to understand, but they are shown in figure 7.77.

**Figure 7.76:** Example of a multi-load unit equipment type on two equipment routings.
Figure 7.77 gives a more abstract representation of figure 7.76. It does not only show the six arcs from the equipment sub-model (shown in thin lines), but also the two arcs from the load unit sub-model (shown in thick lines) that represent the total number of load units that is transported over an equipment routing by this specific equipment pool. Furthermore, it shows two dotted arcs: These arcs represent the integer-type flow of dummy-load units, each arc being associated with exactly one arc from the load unit sub-model. As can be seen from both equations shown in figure 7.77 these flows of dummy-load units are forced to equal the modulus-value of the flows of actual load units. So, if et_{ce} of the considered equipment pool equals 3 as is in this example, then the flow of dummy load units would be as shown in table 7.5 for some values of the flows of actual load units.

The number of equipment needed for transport of the actual load units is based on the number of dummy-load units rather than on the number of actual load units. This property will be used later on when coupling both the load unit sub-model and the equipment sub-model.

As these dummy-load units have no physical meaning there is no need to have separate networks for dummy-load units for all equipment pools in the system; all equipment pools may use the same network for dummy-load units. Nevertheless, there is some distinction between the various equipment pools as each equipment pool has its own arcs between the locations that are served by these equipment pools, as shown in figure 7.77. In this way the flows of dummy-load units are sorted per equipment pool and per section.

<table>
<thead>
<tr>
<th>Load units</th>
<th>No. of machines</th>
<th>Dummy load units</th>
</tr>
</thead>
<tbody>
<tr>
<td>y1</td>
<td>x1+x2+x3</td>
<td>y2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 7.5: Conversion of actual load units into dummy-load units (et_{ce}=3).

7.4.3.2.2. Pulling equipment types

Until now the impact of pulling equipment on the equipment sub-model has not been discussed. It will turn out that this impact is great, even requiring a simplification of the model in some cases to prevent the equipment sub-model to explode in terms of the number of arcs and nodes. This is caused by the aspect of double relocation.

Consider the situation depicted in figure 7.76 in which load units are to be transported from A to B by an MTS trailer system. Suppose that the MTS truck is always uncoupled after a trip with a fixed number of MTS trailers.
In that case the trailers can be relocated either from A, B or C to location A where they are loaded. This trick has been used before. But, unfortunately, these trailers are not in the position to move by their own; they need a truck to pick them up either at A, B or C. These trucks in their turn may need to relocate from locations A, B or C before they can pick up the trucks.

For the trucks, this yields a set of nine different double relocation arcs that are shown somewhat simplified in figure 7.78 for two different equipment routings. In reality these nine double relocation arcs would run directly from the leftmost set of nodes to the rightmost shown location A.

For the trailers, this yields a set of three different single relocation arcs that are shown in thick lines in figure 7.78 for the same two equipment routings. Although it might seem logical to construct nine relocation arcs for the trailers instead of only three, this approach has not been used for following reason: As each different double relocation movement induces a different ratio of time spent on relocation and spent on actual internal transport (A→B), each different double relocation movement requires a different number of trailers to transport the same amount of load units. Also the flow of trailers must be according reality and therefore an arc must originate at each location A through C for each double relocation movement. Fortunately, there is no need to construct such an immense network both for trailers and for trucks; at least the network for the trailers can be kept a little more moderate. This is the result of a little "trick" that will be played with trailer-type equipment.

The "trick" is that flows of load units are not assigned to flows of trailers but rather to flows of trucks. The advantage of this structure is, that the effect of double relocation can be incorporated without any problems in the side-constraints that are to be discussed later on, along with aspects like the working hours of trucks and their speeds. The result of this "trick" is that the problem of double relocation is kept away from the trailers; we only need to match the number of trucks that double-relocate over a certain location with the number of trailers that originate from that location. Another consequence is, that we need to make a distinction between the use of trucks on various equipment routings.

**Figure 7.78:** Example of exploding equipment sub-model due to double relocation\(^{34}\).

Note that in some cases there is no need to relocate at all: If the load units are to be picked up at A, and both the MTS-trailer and the MTS-truck are present at A then the distance for relocation will be zero. On the other hand,

\(^{34}\)For this situation the equipment model contains nine different arcs for the trucks, and another three for the trailers. So, for this single transport of load units a total of twelve arcs is generated for the equipment sub-model! Therefore, care should be taken when deciding for non-simplified double relocation.
an MTS-truck that already is present at A may need to drive to B first to pick up an MTS-trailer, then drive back to A with the empty trailer and finally drive from A to B with the loaded trailer.

It will be clear that, especially in case of a system consisting of a lot of locations and equipment routings that are served by such pulling equipment types in combination with trailer-type equipment types, the equipment sub-model may explode rapidly in this way. Although the situation of figure 7.78 is the most accurate way to model double relocation, in many cases it is desirable to use a simplification of this aspect to keep the size of the equipment sub-model in control. Note that this simplification is optional and should be avoided when both computational speed and computer memory are sufficient to obtain an acceptable solving time of the equipment optimisation model.

![Figure 7.79: Simplification of the double relocation shown in figure 7.78.](image)

The simplification meant here is shown in figure 7.79 where the dotted lines mean to indicate that the aspect of double relocation of pulling equipment to the trailers and then from there to the starting point of the load unit flow is included in the single relocation time of the pulling equipment itself to the starting point of the load unit flow. The main problem of this simplification is to determine a reasonable estimate of the double relocation times, as the number of trailers that are to be picked up from the various locations are not known on beforehand. Therefore, there is a need to make an estimation about these numbers. The basic idea about this assumption is, that the chance for presence of a trailer-type machine is the same for all locations at which these trailers possibly could be present. In other terms, at all locations to which the networks over which the trailers drive are pinned the same number of trailers is supposed to be present and available. Suppose that the number of locations equals \( n \) and that at each of these \( n \) locations exactly \( m \) trailers are available.

Furthermore, it is assumed that there is equal chance that any number out of all these available trailers is needed for internal transport during any time window. In other terms, there is a uniform distribution ranging from 1 to \( n \cdot m \) that specifies the number of required trailers to start for transport of load units on an equipment routing.

Finally, it is assumed that the trailers that require the least amount of relocation of pulling equipment are used first and, when they have been used, the trailers that require the one-but-least amount of relocation are used, and so on. Therefore, if somewhere between 1 and \( m \) trailers are required, they are all taken from the location that requires the least relocation. If somewhere between \((m + 1)\) and \((2m)\) trailers are required, the first \( m \) trailers are taken from the location that requires the least relocation, and the next \( m \) trailers are taken from the location that requires the one-but-least relocation, and so on.

<table>
<thead>
<tr>
<th>Double relocation scheme</th>
<th>Distance [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-A-A</td>
<td>( \Delta x_1 = 0 )</td>
</tr>
<tr>
<td>A-B-A</td>
<td>( \Delta x_2 = 800 )</td>
</tr>
<tr>
<td>A-C-A</td>
<td>( \Delta x_3 = 2300 )</td>
</tr>
</tbody>
</table>

Table 7.6: Example of distances for double relocation of pulling equipment.
Suppose that we consider relocation of pulling equipment from location A to location A (seems ambiguous, but is not), and therefore need to consider double relocation according to following three schemes: (A-A-A), (A-B-A) and (A-C-A). Suppose that the relocation distances $\Delta x_i$ associated with these schemes are as shown in table 7.6. Also suppose that the value of $m$ equals 5. Then it is found that the relocation distance per trailer for the various numbers of trailers possibly required during a time window is as given in figure 7.80. Note that exactly $n \cdot m$ entries are shown in figure 7.80, each value having equal chance of being selected. Now the average of these $n \cdot m$ values must be found and used as the estimated relocation distance for pulling equipment from location A to location A. This value equals 368$^{35}$ metres. In the same way the estimated relocation distance for pulling equipment over locations B and C to location A must be determined.

This average relocation distance, which is called $dx_{reloc\_{avg}}$ can be formulated in a more abstract way. This will be shown by writing down the first six terms of the formula used to determine the above result of 368 metres:

$$dx_{reloc\_{avg}} = \frac{1}{n \cdot m} \left( \frac{1 \cdot \Delta x_1}{1} + \frac{2 \cdot \Delta x_1}{2} + \frac{3 \cdot \Delta x_1}{3} \right) + \frac{1}{n \cdot m} \left( \frac{3 \cdot \Delta x_1 + 1 \cdot \Delta x_2}{4} + \frac{3 \cdot \Delta x_1 + 2 \cdot \Delta x_2}{5} + \frac{3 \cdot \Delta x_1 + 3 \cdot \Delta x_2}{6} \right) + \ldots$$

When writing this in a more formal way then we obtain:

$$dx_{reloc\_{avg}} = \frac{1}{n \cdot m} \left( \Delta x_1 \left( m + \sum_{i=m+1}^{m+m} \right) + \Delta x_2 \frac{2m}{i=m+1} + \sum_{i=2m+1}^{m+m} \right) + \Delta x_3 \left( \sum_{i=2m+1}^{m+m} + \sum_{i=m+1}^{m+m} \right) + \ldots$$

which can be reduced to:

$$dx_{reloc\_{avg}} = \frac{1}{n \cdot m} \left( \sum_{j=1}^{n} \Delta x_j \left( \sum_{i=(j-1)m+1}^{i=m} \frac{i-(j-1)m}{i} + \sum_{i=m+1}^{m+m} \right) \right) \quad (7.2)$$

This formula will not only be used to calculate the average relocation distance for pulling equipment in case of simplified double relocation, but also to calculate the estimated time needed for coupling and loading trailers in case of simplified double relocation. This will be explained when discussing the side-constraints for the coupled load unit- and equipment sub-model.

---

$^{35}$This value depends on the choice for $m$ and therefore needs to be used with care.
Chapter 7. Equipment

Figure 7.81 shows the sensitivity of \(dx_{\text{reloc_\text{avg}}}\) according to formula (7.2) for the number of trailers \(m\) that is assumed to be present at each of the \(n\) locations. This figure shows that the influence of \(m\) is marginal for values greater than 5.

There are two parallel networks that need to be balanced: One network modelling the flow of the trailer-type machines and one network modelling the flow of the pulling machines. This balancing is done through some additional side constraints as will be discussed at the next section. These side constraints will demand that the total number of trailers used for an equipment routing equals the total number of pulling machines used for that very equipment routing, regardless of the origin of these machines. We will also see, that the transport times for trailer-type equipment is derived from the associated pulling equipment pool. That's one of the reasons why to each trailer-type pool exactly one pulling equipment pool must be associated. Otherwise, it would be ambiguous which equipment pool's driving speed should be adapted.

- **Meaning of flows**
  As mentioned before, the flows in the **load unit sub-model** are quite straightforward: A flow of size \(f\) over a section of an equipment routing in a time window means that during that time window \(f\) load units are transported over that section of that equipment routing, ending either in that same time window (in case of internal transport) or in the next time window (in case of stacking).

  The flows in the **equipment sub-model** are a little different: Instead of the total number of machines that is used during a certain time window, they specify the *average* flow of machines during that time window. There is a good reason for this choice. Suppose that we consider two time windows of different duration; the first one lasting 2 hours and the second one lasting 24 hours. Suppose that the number of load units transported per time unit on one specific section of an equipment routing is equal in both time windows, and therefore equals 30 in the first and 360 in the second time window. Also suppose that we would take the *total* number of machines required during each time window as the flow in the equipment sub-model. In that case 12 times as much machines would be needed in the second time window as in the first time window, for the same transport load *per time unit*. Therefore, the equipment sub-model would suppose that in the second time window 12 times as much machines are to be installed as in the first one. Obviously, this is not what was intended, and therefore the *average* flow of equipment should be considered instead of the total flow\(^{36}\).

\(^{36}\)If we would take the *total* number of machines used during a time window as unit of the flows in the equipment sub-model, then the flows in the equipment sub-model would be equal to those in the load unit sub-model, provided that we neglect MLU-equipment and the different arcs for equipment relocation.
So, when speaking of a flow of "15" in the equipment sub-model this means that at any intersection of the arc over which the flow runs, 15 machines will pass during a time span that equals the time unit chosen to express the flow in. So even over the collection nodes connecting two time windows of totally different lengths the flow through the collection node equals the flow through the smallest time window and of the greatest time window. Working hours are not included in this flow as each collection node is related to only one equipment pool; regardless of the time window that is considered, each arc stemming from that collection node is confronted with exactly the same working hours, as all these arcs are related to the same pool in the same window and therefore to the same working hours. Therefore, no need to include the working hours in the flows through the equipment sub-model. Compare with figure 7.82 and find that the limited working hours that apply for transport in time window "w=3" do not influence the flow of machines at all.

Obviously, the effect of limited working hours during a time window is that the actual flow of machines during the simulation will at some times be higher than the average flow, spread over the whole time window, as found by the equipment optimisation model. However, although this peaked character is not included directly in the size of the flows running through the equipment sub-model, it is included in the costs of the equipment sub-model. And therefore it is included in the flows of the equipment sub-model in an indirect way.

In some cases the duration of the time windows in the second and third phase could be thus wide that an unpleasant situation might occur: A situation in which two equipment pools that are supposed to transport load units on subsequent sections through direct interference, and therefore are mutually dependent, are confronted with mutually excluding working hours as shown in figure 7.83. The equipment optimisation model won’t detect this discrepancy in working hours until the first planning phase where it will try to correct as much as possible for the invalid situation. Normally, this situation would yield an infeasible solution of the equipment optimisation model. It will be clear that such a situation is merely the result of an incorrectly modelled situation than of an obvious omission of the equipment optimisation model.

**Working hours**

Working hours are not directly included in the flows of machines, but they are used at some other points: They are used (1) to guarantee that any arc in the equipment sub-model represents a valid transport movement that can be done within the time window(s) that the arc covers. In other terms, they are used to guarantee that arcs in the equipment sub-model are only created within time window(s) if the available working time during that or those time window(s) exceeds the so-called transport time, which is discussed shortly at the costs. They are also used (2) to model differences in the costs between various equipment pools that can be used on the same section. Again, this aspect will be discussed shortly at the costs.

**Helpful hint**

See Appendix D for a description of the variables used here and the way they are coded.

---

37 The same holds for all arcs ending in that collection node.
Costs

Costs in the equipment sub-model are somewhat different from the costs in the load units sub-model: While the load unit sub-model only included costs per load unit induced by stacking, the equipment sub-model includes costs per flow of machines per time unit rather than of individual machines, induced by actual transport, equipment relocation and transhipment. It could also be chosen for to express costs per machine instead of costs per flow unit, approach has not been selected to be in correspondence with the unit of flow through the equipment sub-model which is numbers of machines per time unit.

To aspects are included in the costs: Working hours and transport times. These transport times contain information on travelling speeds, transhipment speeds, transport distances, relocation distances and carrying capacities as will be discussed shortly.

Costs - working hours

First of all, there is a need to discuss the way working hours of equipment pools are included in the equipment optimisation model. Actually, they are included in a rather rough way, as a constant factor that increases the costs of the use of a certain equipment pool in a certain time window. That same factor is also used to determine whether any arc in the equipment sub-model must skip a time window or not, to be valid.

Suppose that we consider the use of an equipment pool \( p \) of straddle carriers during a certain time window \( w \). Suppose that we find that they are available for 6 out of the 8 hours that this time window lasts, according to their time schedules. In this case the factor modelling the working hours in the side constraints would be, for that equipment pool in that time window (Refer to Appendix D for details on the coding system of parameters):

\[
\text{ep\_wtf}_{w,p} = \frac{8}{6} = 1.33,
\]

or, in a more formal way:

\[
\text{ep\_wtf}_{w,p} = \frac{w\_dt_w}{\text{ep\_wt}_{w,p}} \tag{7.3}
\]

in which:

\[
\begin{align*}
\text{ep\_wtf}_{w,p} &= \text{working time factor modelling limited working hours for pool } p \text{ in time window } w \\
\ w\_dt_w &= \text{duration of time window } w \\
\text{ep\_wt}_{w,p} &= \text{working time of pool } p \text{ in time window } w.
\end{align*}
\]

The values of \( \text{ep\_wtf}_{w,p} \) can be derived easily from the working schedules of the equipment pool, while the duration of the time window is known through the planning phase it belongs to.

In some cases there is a need to determine the working time factor over various time windows, needed for those arcs in the equipment sub-model that span more than one time window. In that case formula (7.3) transforms into:

\[
\text{ep\_wtf}_{w\_start,p} = \frac{\sum_{w=w\_start}^{w\_end} w\_dt_w}{\sum_{w=w\_start}^{w\_end} \text{ep\_wt}_{w,p}} \tag{7.4}
\]

in which:

\[
\begin{align*}
w\_start &= \text{time window from which the multiple-window arc stems} \\
w\_end &= \text{time window in which the multiple-window arc end.}
\end{align*}
\]

This aspect is also discussed at page 148.
Working hours of trailer-type equipment are neglected automatically through the chosen structure in which flows of load units are related to pulling equipment rather than to the carrying trailer-type equipment. In this way working hours of trailer-type equipment have no meaning at all, which is in correspondence with reality.

- Costs - transport times
Second, there is need to discuss the duration of each transport over each section, which depends on various aspects like the capabilities of the equipment pool. These transport times can be constructed of at most following components: (1) relocation, (2) coupling of trailers, (3) loading of load units, (4) actual transport, (5) unloading of load units and (6) uncoupling of trailers. In case of pulling equipment relocation must be replaced by double relocation, and the sequence of coupling and loading may be inverted, like the sequence of uncoupling and unloading may. To help understanding the phenomenon of transport times figure 7.64 is repeated in figure 7.84.

The values for all aspects included in the transport time for an equipment type that is able to carry, to move by itself and to transship is given in table 7.7. An example of such an equipment type is a straddle carrier.

<table>
<thead>
<tr>
<th>Ability</th>
<th>Aspect</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can carry</td>
<td>Carrying capacity</td>
<td>own ( \text{et}_cc )</td>
</tr>
<tr>
<td>Self-propelled</td>
<td>Travelling speeds</td>
<td>own ( \text{et}<em>{vat}, \text{et}</em>{lat}, \text{et}<em>{udt}, \text{et}</em>{lft}, \text{et}<em>{uv} ) and ( \text{et}</em>{lv} )</td>
</tr>
<tr>
<td>Can tranship</td>
<td>Transshipment speeds</td>
<td>own ( \text{et}<em>{at}, \text{et}</em>{ut}, \text{et}<em>{dt} ) and ( \text{et}</em>{rt} )</td>
</tr>
</tbody>
</table>

Table 7.7: Parameters for transport time for fully self-supported equipment like straddle carriers.

The next type of equipment to discuss is a little bit simpler than the type discussed in table 7.7 in that this type is not able the transship by itself. Therefore, it depends on the transshipment capacity of equipment pools that are installed either at the same sections as the pool of this type, or at the neighbouring section. Refer to page 92 for another discussion of this aspect.

For each transshipment onto or from equipment that can not transship itself the average transshipment speed of all equipment pools that can be used for that particular transshipments at the same section as the load units are transport is used as the estimated transport speed. The various points at which transshipment may occur can be found in figures 7.45 through 7.52 along with a description of the pools that may be used for each transshipment movement. Note that, if some equipment pools are available for transshipment at the same section as the pool of this type, then their transshipment speeds will be used both for loading and for unloading. This is summarised in table 7.8 giving an example of AGVs, and shown in figure 7.85 where the AGV can be loaded either by a pool of RMGs at a neighbouring section or a pool of SCs at its own section, and can be unloaded either by another pool of RMGs at another neighbouring section or the same pool of SCs at its own section. At this point it does not matter at what equipment network these equipment pools are installed, just as long as their equipment network may serve one or more of the sections meant here.
### Chapter 7. Equipment

<table>
<thead>
<tr>
<th>Ability</th>
<th>Aspect</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can carry</td>
<td>Yes</td>
<td>Carrying capacity</td>
</tr>
<tr>
<td>Self-propelled</td>
<td>Yes</td>
<td>Travelling speeds</td>
</tr>
<tr>
<td>Can transship</td>
<td>No</td>
<td>Transshipment speeds</td>
</tr>
</tbody>
</table>

#### Table 7.8: Parameters for transport time for non-transhipping equipment like AGVs.

The third type of equipment to discuss can neither transship nor move by itself. This is the normal trailer-type of equipment which is only capable of carrying load units. As any equipment type that is not self-propelled and can move must be associated with another equipment type that is self-propelled and can pull, the travelling speeds can simply be adapted from that associated pulling equipment type. This is done in a rather indirect way, as will be explained in more detail when discussing the total side-constraints. Namely, flows of (dummy)-load units are not related to flows of trailer-type equipment, as one might expect, but rather to flows of pulling equipment that drag the trailer-type equipment around the terminal yard. This remarkable structure has some advantages that will become clear when discussing the relocation time for pulling equipment.

Therefore, table 7.9 states that various aspects are not applicable (N/A) for trailer-type equipment although in reality these aspects are included in the model, in an indirect way.

Figure 7.86 shows the relation between trailer-type equipment and pulling equipment in a somewhat different way: The shown trailer-type of equipment requires the presence of any pulling pool of the type specified at the trailer-type's parameter et_ptr to be present at any network at which the trailer-type equipment is installed. This figure also shows that, after the right pulling pool has been installed, the pulling pool merely refers to the trailer-type pool than vice versa. This structure, which has been discussed before, tries to express that flows of trailer-type equipment are matched with flows of pulling equipment that are used for that particular trailer-type pool, rather than vice versa. In other terms, rather than that flows of pulling pools are matched with flows of trailer-type equipment. This structure is used both for INTRA- and inter-window equipment hopping.

<table>
<thead>
<tr>
<th>Ability</th>
<th>Aspect</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can carry</td>
<td>Yes</td>
<td>Carrying capacity</td>
</tr>
<tr>
<td>Self-propelled</td>
<td>No</td>
<td>Travelling speeds</td>
</tr>
<tr>
<td>Can transship</td>
<td>No</td>
<td>Transshipment speeds</td>
</tr>
<tr>
<td>Can couple</td>
<td>No</td>
<td>Couple speeds</td>
</tr>
</tbody>
</table>

#### Table 7.9: Parameters for transport time for trailer-type equipment like MTS-trailers.

The fourth type of equipment is the pulling type of equipment. As shown in the last paragraph trailer-type equipment is related to exactly one pulling-type equipment. However, this does not mean that a pulling-type of equipment can serve only one type of trailer-type equipment. On the contrary, a score of totally different trailer-type equipment types could be associated to one pulling-type of equipment.

It has been shown before that in this model any pulling-type of equipment is not able to transship nor to carry load units. In this way the model can be kept relatively simple without offending most real-life situations. Table 7.11 shows the way the various parameters are derived for pulling-type equipment. One special parameter is the moment of coupling, which is said to be optional and specified in the pulling pool's parameter ep_ebo. If this option is used then the times for transhipment onto and from the trailer-type equipment are to be included both in the transport time. Otherwise, they are to be left out from the transport time.
There is another funny thing about the transport time when ep_cbt0 is not used, i.e. when the pulling equipment is to couple to the trailer after the trailer has been loaded and to uncouple before the trailer is being unloaded. This is strongly related to the position from which the trailer that is to be loaded has to be picked up by the pulling equipment: (1) Whenever the trailer is present at the starting location of the actual transport movement, i.e. at the right collection node, then there is no need to relocate the trailer before loading (at least, that is the assumption in this model) and therefore the pulling equipment may couple after the trailer has been loaded. No problem at all. (2) Otherwise, the trailer is to be relocated before it can be loaded, and therefore needs to be pulled by the pulling equipment to the starting location of the actual transport movement. Obviously, this relocation requires the pulling equipment to couple first, and therefore the setting of ep_cbt0 is to be ignored for the relocation of the trailer; so the coupling time before loading is to be included in the transport time anyhow, regardless of ep_cbt0. The uncoupling time after the actual transport movement, on the contrary, is not to be omitted from the transport time as this aspect is not affected by the relocation of the trailer. This is shown in figure 7.87, showing the effect of ep_cbt0 and trailer relocation on the transport time of the pulling equipment. Note that this figure does not show the relocation of the pulling equipment itself.

So, if ep_cbt0 is not used in combination of simplified double relocation then the value of the transhipment onto the pulled trailers is to be included in the total transport time for those cases in which the trailers are to be relocated before the actual transport. There is a great resemblance with the average relocation distance dx_reloc_avg as discussed before at figure 7.80 and at table 7.6.

A slight error is introduced here in the model due to the chosen structure in which flows of load units are related to pulling equipment rather than to trailer-type equipment. In those cases where the time needed to tranship onto or from trailers is omitted from the transport time of pulling equipment, it automatically is omitted for the transport time of trailer-type equipment, too! This is caused by matching the number of pulling equipment exactly with the number of trailer-type equipment. This error has noticeable consequences when (1) ep_cbt0 is not used, (2) the actual transport time is short relative to the transhipment time, and (3) the carrying capacity et_ce of the trailers is definitely greater than 1. The best way to avoid these problems is to model trailer and pulling equipment together as one, self-propelled equipment type, thus avoiding coupling and uncoupling of trailers. Obviously, it depends on the situation being modelled whether this simplification is allowed or not.

Analogous to table 7.6, table 7.10 can be constructed in which it is indicated for which double relocation schemes the transhipment time dt_tranship onto the trailers is to be included in the relocation time. One major difference with table 7.6 is that at this point we are not interested at all in the origin of the pulling equipment; that is the reason for showing this origin as '??' in table 7.10. The only factor of interest is the origin of the trailers as this is the factor that determines whether the transhipment time onto the trailers is to be included in the total transport time.

<table>
<thead>
<tr>
<th>Double relocation scheme</th>
<th>Transhipment time [hr]</th>
</tr>
</thead>
<tbody>
<tr>
<td>?-A-A</td>
<td>$\Delta T_1 = 0$</td>
</tr>
<tr>
<td>?-B-A</td>
<td>$\Delta T_2 = dt_{tranship}$</td>
</tr>
<tr>
<td>?-C-A</td>
<td>$\Delta T_3 = dt_{tranship}$</td>
</tr>
</tbody>
</table>

Table 7.10: Inclusion of transhipment onto trailers for several double relocation-origins. (Only valid if ep_cbt0 is used in combination with simplified double relocation.)
Table 7.10 shows the additional time $\Delta t_i$ that is to be included in the transport time for each origin of the trailer, but we need to estimate the average value of these $\Delta t_i$'s for all possible double relocation schemes. To make this estimate, formula (7.2) is used with following adaptations: First, the values of $\Delta x_i$ are replaced by the values of $\Delta t_i$. Second, the result is called $dt_{\text{transh avg}}$ instead of $dx_{\text{reloc avg}}$. Now the average transhipment time can be determined as:

$$dt_{\text{transh avg}} = \frac{1}{n \cdot m} \left( \sum_{j=1}^{n} \Delta t_j \cdot \left( \sum_{i=1}^{m} \frac{i \cdot (i-1) m + i}{i = (i-1) m + 1} \right) + \sum_{j=1}^{m} \right),$$

(7.5)

in which $\Delta t_i$ is derived from table 7.10 and equals either zero or $dt_{\text{tranship}}$. Analogous to the average relocation distance, the resulting average transhipment time is dependent of the selected value of $m$. If this parameter is set to 5 then the resulting average transhipment time is 0.32 $dt_{\text{tranship}}$. Some other values are shown in figure 7.88 giving an indication of the sensitivity of the average transhipment time of $m$ and of $n$; note that in case of only one location the average transhipment time equals zero for any $m$. Nevertheless, this case is purely theoretical as in practice at least two different relocation nodes will exist for each pool.

As pulling equipment could potentially be associated to a score of totally different trailer-type equipment pools, it might be ambiguous which of these carrying capacity is meant in table 7.11. The answer lies in the demand discussed before that a pulling pool can only pull trailer-type pools of the right type(s) that were installed at the same equipment network as the pulling pool; overlapping networks can not overrule this fixed assignment. Therefore, for each network at which a pulling pool is installed (for each equipment routing) a score of arcs is created, each arc being associated with exactly one trailer-type pool. As seen in table 7.2 when discussing the load unit sub-model, each unique combination of a trailer-type pool with a pulling pool is considered as a new, fictitious equipment pool. This new, fictitious equipment pool does not exist as such nor is created automatically somehow by the model, but the load unit sub-model does contain a separate arc modelling the flows of load units transported by that combination of trailer-type and pulling pool.

**Figure 7.88:** Sensitivity of $dt_{\text{transh avg}}$ for $m$ and $n$.

As pulling equipment could potentially be associated to a score of totally different trailer-type equipment pools, it might be ambiguous which of these carrying capacity is meant in table 7.11. The answer lies in the demand discussed before that a pulling pool can only pull trailer-type pools of the right type(s) that were installed at the same equipment network as the pulling pool; overlapping networks can not overrule this fixed assignment. Therefore, for each network at which a pulling pool is installed (for each equipment routing) a score of arcs is created, each arc being associated with exactly one trailer-type pool. As seen in table 7.2 when discussing the load unit sub-model, each unique combination of a trailer-type pool with a pulling pool is considered as a new, fictitious equipment pool. This new, fictitious equipment pool does not exist as such nor is created automatically somehow by the model, but the load unit sub-model does contain a separate arc modelling the flows of load units transported by that combination of trailer-type and pulling pool.

**Figure 7.89:** Networks showing the allowed combinations of pulling and trailer-type equipment.
Chapter 7. Equipment

Analogous, in the equipment sub-model some separate arcs are created for this new, fictitious equipment pool: Both for the trailer-type pool (which has a separate flow-network) and for the pulling pool (which has a separate flow-network, too) a number of arcs is created, each arc being related to only one, unique combination of trailer-type and pulling pool at each section. This is clarified in figure 7.89. This figure is based on the situation described at table 7.2 in which two trailer-type pools may be pulled by one pulling pool named A.

Figure 7.89 shows how for each relocation arc over which it may (double) relocate to pick up trailers the pulling pool refers to only one trailer-type pool. Analogous, for each relocation arc over which the trailers may relocate a separate arc per allowed pulling pool is created. Fortunately, in this situation each trailer-type pool may be pulled by only one pulling pool, that is pool A. Note that this figure shows the situation for one equipment routing only, and therefore for one section only.

Through this structure of imagining a new, fictitious equipment pool, it is always known which carrying capacity should be adapted for that pool: The carrying capacity of the trailer-type that makes part of the fictitious equipment pool. Also, refer to figure 7.86.

<table>
<thead>
<tr>
<th>Network</th>
<th>Trailers</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>trail_X</td>
</tr>
<tr>
<td></td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>trailer relocation</td>
</tr>
<tr>
<td></td>
<td>actual transport</td>
</tr>
<tr>
<td></td>
<td>pick up trailers</td>
</tr>
<tr>
<td></td>
<td>load trailers</td>
</tr>
<tr>
<td>II</td>
<td>trail_Y</td>
</tr>
<tr>
<td></td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>od</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 7.90:** *Arcs for pulling pool pull_A per network and per trailer-type pool.*
This will be illustrated by a small example in which one pulling pool pull_A is installed at two different networks (I and II) and is to drag two different trailer-type pools: Pool trail_X at network I and pool trail_Y at networks I and II. Figure 7.90 shows the two networks that are constructed for pulling pool pull_A. The thick lines represent the arcs for equipment relocation, which could be either simplified or non-simplified double relocation.

One remarkable conclusion that can be drawn from figure 7.90 is that there is no difference between the two networks concerning equipment pool trail_Y. This is caused by the double installation of pulling pool pull_A at both networks; should some other pulling pool pull_B be installed at network II instead of pool pull_A then with three different networks were needed: Two for pool pull_A at network I and one for pool pull_B at network II.

Note that it would not make any difference if pool pull_B would have been installed at network II next to pool pull_A and not instead of pool pull_A. In both cases pool pull_A could pull pool trail_Y, irrespective of the network that is considered.
Table 7.11: Parameters for transport time for pulling-type equipment like MTS-trucks.

- Costs - transport times – distances – general
  The distances for relocation and for normal transport are derived according to following rules. If the equipment routing to which the relocation movement is related connects both the origin and the destination of the relocation movement then the travelling distance over the physical paths between this origin and destination is used. Otherwise, the rectangular distance between the origin and destination is used. Suppose that the relocation distance found thus is named **dx_reloc** and that the actual transport distance is named **dx_transp**. Note that this distance may include, or even solely consist of, the local transport distance at a bounding element.

- Costs - transport times – distances – pulling equipment
  As you may have deduced from the oncoming coupling of trailers to pulling equipment, and of the pulling equipment to load units, transportation times are only modelled for the pulling pools; the trailer-type pools simply follow the flow of pulling pools and therefore only indirectly influence the total flow of load units through the system. So, the subject "transportation times at trailer-type equipment" can be skipped.

Leaves us with the transport times of pulling equipment: We need to determine the relocation distance and the actual transport distance. The latter is determined completely analogous to the actual transport distance **dx_transp** described before for normal equipment types, although we should bear in mind that actually we are looking at the transport distance of the trailer-type equipment rather than of the pulling equipment. The relocation distance is somewhat more complicated, and depends on the setting for simplified double relocation.

If simplified double relocation is used then the relocation distance **dx_reloc** will be estimated as shown in figure 7.80. It will be found later on that the way flows of pulling equipment are related to flows of trailer-type equipment is quite "loose" compared to non-simplified double relocation with respect to the locations from which both pools relocate.

If simplified double relocation is not used then the relocation distance **dx_reloc** can be determined accurately for each (o-od, od-s) pair as shown in figure 7.90 as the sum of both partial relocation distances at both paths o-od and od-s. Each of these partial relocation distances is determined as described above for all normal equipment types.

One consequence of the way double relocation is modelled in this thesis is that it induces a small simplification of the model: The model assumes that during the total relocation time of pulling equipment, including the time for double relocation, the trailer-type equipment is "reserved" and can not be used for other purposes. In reality in some cases the pulling equipment may have started for the first step of double relocation while the trailer that is to be picked up is still in use for another transport movement. As this type of planning is too complex for normal terminals it will hardly occur in practice (except by coincidence) and does not need to be included in the optimisation model.

- Costs - transport times – duration
  The times for relocation and for normal transport are derived from **dx_reloc** and **dx_transp**, and of the travelling speeds that apply. Suppose that the speeds found for loaded and unloaded travelling are stored in **v_l** and in **v_u** respectively, and that the times for loaded and unloaded acceleration and deceleration are stored in **ta_l, ta_u, td_l** and **td_u** respectively. In most cases these values just correspond with the pool's own parameters **et_uv**, **et_lv**, **et_uat** and **et_lat** respectively. When determining the time for relocation and for normal transport over a distance of, say, **dx** we need to distinct between the situation in which full speed is reached and that in which full speed is not reached. The threshold distance **dx_thres** that is sufficient to let the equipment reach full speed for a split second only and let it decelerate immediately after is crucial at this point. This threshold distance for the parameters:
- speed **v',** for which either **et_uv** or **et_lv** can be used;

```plaintext
<table>
<thead>
<tr>
<th>Ability</th>
<th>Aspect</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can carry</td>
<td>No</td>
<td>Carrying capacity</td>
</tr>
<tr>
<td>Self-propelled</td>
<td>Yes</td>
<td>Travelling speeds</td>
</tr>
<tr>
<td>Can tranship</td>
<td>No</td>
<td>Transhipment speeds</td>
</tr>
<tr>
<td>Can couple</td>
<td>Yes</td>
<td>Couple speeds</td>
</tr>
</tbody>
</table>
```

<table>
<thead>
<tr>
<th>Ability</th>
<th>Aspect</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can carry</td>
<td>No</td>
<td>Carrying capacity</td>
</tr>
<tr>
<td>Self-propelled</td>
<td>Yes</td>
<td>Travelling speeds</td>
</tr>
<tr>
<td>Can tranship</td>
<td>No</td>
<td>Transhipment speeds</td>
</tr>
<tr>
<td>Can couple</td>
<td>Yes</td>
<td>Couple speeds</td>
</tr>
</tbody>
</table>
```
• acceleration time $T_a$, for which either $ta_u$ or $ta_l$ can be used;
• deceleration time $T_d$, for which either $td_u$ or $td_l$ can be used;
is determined as follows:

$$dx_{thres} = dx_{accel} + dx_{decel} = \int_{t=0}^{T_a} \left( \frac{V}{T_a} \right) \cdot t \cdot dt + \int_{t=0}^{T_d} \left( \frac{V}{T_d} \right) \cdot t \cdot dt$$

which results in:

$$dx_{thres} = \frac{V}{2} \cdot (T_a + T_d)$$  \hspace{1cm} (7.6)

Pulling equipment is not that different from normal equipment types concerning the duration of relocation and of actual transport. The only difference is that both steps in double relocation are considered empty transport. Therefore, even driving with empty trailers while relocating them is considered empty transport. The transport times involved with transport of trailer-type equipment are taken into by at the appropriate pulling pools only; they are not taken into account by the trailer-type pools.

♦ Costs - transport times - duration - [large distances]
Whenever either the transport distance or the relocation distance equals or exceeds the value of $dx_{thres}$ then the time for transport or relocation $\Delta T$ over distance $\Delta X$ is determined as follows. First, the time needed for acceleration and for deceleration is determined as $(T_a + T_d)$ during which time the distance $dx_{thres}$ is covered. During the remaining part of $\Delta X$, which equals $(\Delta X - dx_{thres})$, the equipment drives at its nominal speed $V$. Therefore, the total time $\Delta T$ over distance $\Delta X$ equals:

$$\Delta T = (T_a + T_d) + \frac{\Delta X - dx_{thres}}{V} =$$

$$= (T_a + T_d) + \frac{V \cdot T_a - V \cdot T_d}{2} =$$

$$= \frac{T_a + T_d}{2} + \frac{\Delta X}{V}$$  \hspace{1cm} (7.7)

♦ Costs - transport times - duration - [small distances]
Whenever either the transport distance or the relocation distance equals or is less than the value of $dx_{thres}$ then the time for transport or relocation $\Delta T$ over distance $\Delta X$ is determined as follows. First, we need to realise that the time $t_a$ that the equipment can accelerate is related to the time $t_d$ that the equipment must decelerate according to:

$$t_a \cdot \frac{V}{T_a} = t_d \cdot \frac{V}{T_d}$$  \hspace{1cm} (7.8)

as the speed at the end of accelerating equals the speed at the start of decelerating. Furthermore, the distance $dx_a$ covered during acceleration and the distance $dx_d$ covered during deceleration are given by:

$$dx_a = \int_{t=0}^{t_a} \left( \frac{V}{T_a} \right) \cdot t \cdot dt = \frac{V \cdot t_a}{2}$$

and:

$$dx_d = \int_{t=0}^{t_d} \left( \frac{V}{T_d} \right) \cdot t \cdot dt = \frac{V \cdot t_d}{2}$$

respectively. Combining these results with formula (7.8), and realising that $(dx_d + dx_d)$ must equal $\Delta X$ the values for $\Delta X$, $t_a$ and $\Delta T$ are found as:
\[ \Delta X = dx_{-a} + dx_{-d} = \frac{V \cdot t_{-a}}{2} \left(1 + \frac{T_{-d}}{T_{-a}}\right) \]

\[ \Rightarrow t_{-a} = \frac{2 \cdot \Delta X \cdot T_{-a}}{V \cdot (T_{-a} + T_{-d})} \]

\[ \Rightarrow \Delta T = t_{-a} + t_{-d} = t_{-a} \left(1 + \frac{T_{-d}}{T_{-a}}\right) = \frac{2 \cdot \Delta X \cdot T_{-a}}{V \cdot (T_{-a} + T_{-d})} \frac{T_{-a} + T_{-d}}{T_{-a}} = \frac{2 \cdot \Delta X}{V} \quad (7.9) \]

**Costs - transport times - duration - [synopsis]**
Both formulas (7.7) and (7.9) are used to determine the time needed for equipment relocation and for the actual transport movement. In other terms, they are used to convert \( dx_{\text{reloc}} \) and \( dx_{\text{transp}} \) into \( dt_{\text{reloc}} \) and \( dt_{\text{transp}} \) respectively.

**Costs - transport times - pulling and trailer-type equipment**
The values of \( dt_{\text{reloc}} \) and \( dt_{\text{transp}} \) for pulling equipment can be calculated with formulas (7.7) and (7.9), too. Only bear in mind that in case of simplified double relocation the relocation distance \( dx_{\text{reloc}} \) is to be calculated according to formula (7.2). Whenever simplified double relocation is not used then the value of \( dx_{\text{reloc}} \) can be determined accurately for each double relocation scheme of which some examples have been given in table 7.6.

**Costs - transport times - [synopsis]**
Suppose that the results of the above calculations on transport times are named \( dt_{\text{reloc}} \) and \( dt_{\text{transp}} \) respectively. Also suppose that the time needed for transshipment \( dt_{\text{transh}} \) has been determined as the carrying capacity times the transshipment speed for the equipment pool, according to the pool's characteristics shown in tables 7.7 through 7.11. In case of pulling equipment pools the value of \( dt_{\text{transh}} \) may be determined according to formula (7.5) as an estimate of the average transshipment time of all double relocation schemes, used in case of simplified double relocation and when the option \( \text{ep\_cbto} \) is used. If \( \text{ep\_cbto} \) is not used then the transshipment onto the trailers is always included in the transport time. If simplified double relocation is not used then the transshipment time can be determined accurately for each double relocation scheme according to figure 7.87. Finally, suppose that the time needed to couple trailer-type equipment \( dt_{\text{couple}} \) has been determined according to tables 7.9 through 7.11 for trailer-type equipment and for pulling equipment.

Then the total transport time can be determined simply as the sum of the previous times according to:

\[ \text{ep\_tt}_{p,s,o} = dt_{\text{reloc}}_{p,s,o} + dt_{\text{transp}}_{p,s} + dt_{\text{transh}}_{p,s} + dt_{\text{couple}}_p \quad (7.10) \]

in which:

\[ \text{ep\_tt}_{p,s,o} = \text{transport time involved with the used of equipment pool } p \text{ for transport on section } s, \text{ when relocating from origin } o. \]

In case of pulling equipment that does not use double relocation the origin \( \text{od} \) over which the pulling equipment is to double-relocate to pick up the trailers must be included in the transport time, and therefore:

\[ \text{ep\_tt}_{p,\text{pull},s,o,\text{od},p\_trail} = dt_{\text{reloc}}_{p,s,o,\text{od}} + dt_{\text{transp}}_{p\_\text{pull},s,p\_\text{trail}} + dt_{\text{transh}}_{p\_\text{pull},s,p\_\text{trail}} + dt_{\text{couple}}_{p\_\text{pull}} \quad (7.11) \]

in which:

\[ p\_\text{trail} = \text{trailer-type equipment pool} \]

\[ p\_\text{pull} = \text{pulling equipment pool}. \]

The double index of these parameters, \( p\_\text{pull},...,p\_\text{trail} \) is used to express that they hold for the unique combination that the pulling pool is pulling the trailer-type pool for transport over one section; for any other combination of pulling pool and trailer-type pool other parameters may hold, either due to different speeds of the pulling pool or due to a different carrying capacity of the trailer-type pool.

In case of pulling equipment that does use double relocation there is no need to include the origin \( \text{od} \) in the index of the transport time, and therefore formula (7.11) reduces to:
\[ ep_{-}tt_{p\_pull,s,o,p\_trail} = dt_{-}reloc_{p,s,0} + dt_{-}transp_{p\_pull,s,p\_trail} + dt_{-}transh_{p\_pull,s,p\_trail} + dt_{-}couple_{p\_pull} \]  

(7.12)

Remember that for equipment that is purely used to transship load units between to non-transhipping pools (like a SC transshipping between two pools of MTS trailers) the transport time \( dt_{-}transp \) will equal zero, as these transhipping pools do not drive with the load units. You will find, however, no example of this type of transshipment in figure 7.41.

\* Costs - [synopsis]

Now that we have found the factor modelling the working hours and the transport times as well, we finally can determine the costs that are to be included in the equipment sub-model. These costs are based on the preferences as specified at the equipment routings. In some cases these preferences may be stored at fixed preferences and in other cases these preferences may be stored at the equipment routings, different for different sections and networks.

The costs need to express three elements: (1) They need to express the operational costs involved with the use of equipment for certain transports. These costs include the transport times, needed to find the nearest or fastest equipment, and the preferences from the equipment routings, needed to choose between alternative pools. (2) They need to express the "peakness" costs of flows of equipment involved with limited working hours. These costs only control selection between various pools and not selection of machines of one and the same pool between various collection nodes. In other terms, these costs do not control the relative preference for various collection nodes for a particular equipment pool, but they do control the relative preferences for various equipment pools for a particular section. (3) The need to express a neutral attitude of the equipment optimisation model towards transport by any equipment pool with respect of the various time windows in which load units could be transported. In other terms, the differences in duration of different time windows may not show through in any way in the results of the equipment optimisation model.

(1). To include operational costs of the use of equipment we include the preference to use equipment at certain sections as stored at the equipment routings in the costs, multiplied by the previously determined transport time on that section. In this way it is modelled that machines that require little relocation are preferred over other machines (of the same pool) that require more relocation.

(2). To include the "peakness" of the use of equipment in time windows with limited working hours we include the previously defined working time factor \( ep_{-}wtf_{w,p} \) in the costs. In this way we express that, although the average flow, average over the full time window(s) over which the arc in the equipment sub-model runs, models the number of equipment that is needed at average during the time window(s), at simulation the flow of machines will be concentrated during the working hours only; and therefore will be \( ep_{-}wtf_{w,p} \) times the average flow that was returned by the equipment optimisation model. But there is no problem at all with this simplification, as it holds for all flows of machines of any particular equipment pool in that time window. Therefore, at simulation the flows of machines will be peaked in the same way for all arcs at which that particular equipment pool is used. The number of machines needed, both at the equipment optimisation model and at simulation still equals the found flow times the full duration of the time window(s). Also, refer to figure 7.82. In case of arcs covering more than one time window, the total

(3). To include a neutral attitude of the equipment optimisation model towards time windows of different length it might seem as if we need to include the duration of each time window in the costs. However, this is not the case. As it requires some more knowledge on the side-constraints matching the flows of load units to flows of machines this topic is postponed until page 154

So, the total costs involved with the assignment of one unit of flow of equipment to transport over section \( s \), starting is time window \( w \) equal:

\[ ep_{-}ac_{w,p,s,0} = ep_{-}tt_{p,s,0} \cdot ep_{-}sp_{p,s} \cdot ep_{-}wtf_{w,p} \]  

(7.13)

in which:

\[ ep_{-}sp_{p,s} = \text{preference to use equipment pool } p \text{ on section } s \text{ as stored at the equipment routings} \]
Chapter 7. Equipment

\[ \text{ep}_{-ac_{w,p,s,o}} = \text{costs involved with the use of equipment pool p, relocating from origin o, on section s per unit of flow in time window w}. \]

Again, for pulling equipment that does not use simplified double relocation we need to include the second origin od over which the pulling equipment must relocate to pick up the trailers. So, for pulling equipment we find, instead of formula (7.13):

\[ \text{ep}_{-ac_{w,p,pull,s,o,od,p_trail}} = \text{ep}_{-tt_{p,pull,s,o,od,p_trail}} \cdot \text{ep}_{-sp_{p,pull,s}} \cdot \text{ep}_{-wtf_{w,p,pull}} \] (7.14)

in which:

- \( p_{\text{trail}} \) = trailer-type equipment pool
- \( p_{\text{pull}} \) = pulling equipment pool.

And for pulling equipment that does use double relocation there is no need to include the origin od in the index of the assignment costs, and therefore formula (7.14) reduces to:

\[ \text{ep}_{-ac_{w,p,pull,s,od,p_trail}} = \text{ep}_{-tt_{p,pull,s,od,p_trail}} \cdot \text{ep}_{-sp_{p,pull,s}} \cdot \text{ep}_{-wtf_{w,p,pull}} \] (7.15)

**Initial Flows**

Analogous to the load unit sub-model the equipment sub-model incorporates initial flows which specify the flows of equipment at the current simulation time, i.e. the moment in simulation time at which the equipment optimisation starts. As found at the load unit sub-model, the initial flows can be split into two parts: (1) Flows of equipment that are present at the collection nodes and currently not in use, and (2) flows of equipment that are currently in use either for a transport movement or for a transshipment movement. Both flows will be handled somewhat different, as discussed at the load unit sub-model.

(1). Flows of equipment that are present at the collection nodes and not in use at the current simulation time at which the equipment optimisation model starts are converted into initial flows by dividing that amount of equipment by the duration of the time window. So, supposing that this amount of machines of equipment pool \( p \) that present at location \( l \) equals \( n_{-l} \), we find the initial flow \( \text{flow}_{-l_p} \) to be:

\[ \text{flow}_{-l_p} = \frac{n_{-l_p}}{w_{-dt_l}} \] (7.16)

This amount is the first part of the flow assigned to the equipment supply node of location \( l \). In case of unlimited capacity of an equipment pool the size of the initial flow at the equipment insertion point of the equipment pool will be "infinite" to allow newly created equipment enter the system at a pre-defined location. In this model a value of 1,000 machines stored at the equipment insertion point of the pool will be sufficient to model an "infinite" flow in all cases.

(2). Flows of equipment that are currently in use at the current simulation time, either for transport or for transshipment, are to be included in the initial flows, too, though possibly not in the initial flows of the first time window as the first part did. Completely analogous to the problem of initial flows of load units in the load unit sub-model, we are confronted with the problem to determine the time window in which currently used equipment should be inserted. The solution to this problem is the same as found at the load unit sub-model and illustrated in figure 7.44: The time that determines the time window to register the currently used equipment as supply flows is found using the expected time at which the currently executed transport or transshipment will end. This time is estimated as the starting time of the transport- or transshipment movement plus the so-called section transport time, which will be discussed at the equipment sub-model, of that transport- or transshipment movement. As described at the load unit sub-model this approach does induce some simplifications, but these are accepted.

So, the additional initial flow for time window \( w \), if the number of equipment finishing at that time window at location \( f \) would be \( n_{-enter_{-f}} \), we find the initial flow \( \text{flow}_{-enter_{-l_p}} \) to be:

\[ \text{flow}_{-enter_{-l_w,p}} = \frac{n_{-enter_{-l,w,p}}}{w_{-dt_w}} \] (7.17)

For the first time window the initial flows of the collection nodes would equal the sum of (7.16) and (7.17), so:
\[
flow_{-1,p} = \frac{\text{n\_enter\_l}_{1,p}}{\text{w\_dt}_1},
\]

but for the remaining time windows the initial flows would be equal only the value of (7.17), so:

\[
flow_{-w,p} = \frac{\text{n\_enter\_l}_{w,p}}{\text{w\_dt}_w} \quad \forall \ w \neq 1
\]

Obviously, the collection nodes referred to by formula (7.19) also receive flows of equipment that have been assigned to transport movements currently determined by the equipment optimisation model, but these movements all start after the current simulation time and are not yet actually executed. These movements are modelled as flows over relocation arcs in the equipment sub-model rather than as flows entering the equipment sub-model for the first time.

- **Equipment supply nodes**
  As discussed before, all newly created machines are inserted at the equipment supply nodes which can be regarded as abstract "security supplies" of machines that are available for use in case of a lack of transport capacity. These new machines can only be inserted at one fixedly specified location\(^{38}\) and the very beginning of the first time window of the first planning phase. In this way any newly created machine is forced to be present in the system during the entire planning horizon, thus inducing costs for the entire time span covered by optimisation. This prevents the optimisation model from inserting a new machine halfway the planning horizon, thus skipping the costs of presence for that machine during the first half of the planning horizon -- which would yield an invalid result of the optimisation model.

### 7.4.3.3. Coupled load unit sub-model and equipment sub-model

The load unit sub-model and the equipment sub-model need to be coupled; in other terms, it must be guaranteed that the flows of load units in the load unit model are coupled to the flows of equipment in the equipment sub-model. This will be established through so-called side-constraints which are available in some modern network computer packages. Refer to Appendix C for a description of these side-constraints. Some other side-constraints than meant here, concerning limited stacking capacities in the system, are defined at the load unit sub-model. Refer to formula (7.1).

- **Helpful hint**
  See Appendix D for a description of the variables used here and the way they are coded.

- **Side-constraints**
  Now we are in the position to construct the total side-constraints. In short, a total side-constraint can be said to match one flow of load units over one section of one equipment routing by one equipment pool with the sum of the various flows of machines of that pool related to that section over the various relocation arcs. So, if the number of load units that is transported by equipment pool \( p \) over section \( s \) over an equipment routing equals \( \text{lu\_tn}_{w,p,s} \)\(^{39}\) and if the number of machines of that equipment pool that is assigned to that flow and is to relocate from origin \( o \) first equals \( \text{epa\_n}_{w,p,s,o} \), then the total side-constraint for that flow of load units is:

\[
\text{lu\_tn}_{w,p,s} = \text{w\_dt}_w \cdot \sum_o \text{epa\_n}_{w,p,s,o} \quad \quad (7.20)
\]

- **Total side-constraint - self-propelled MLU-equipment**
  The side-constraints for self-propelled MLU-equipment are somewhat different from the one shown in formula (7.20). In contrast to normal side-constraints as shown in formula (7.20) they are related to flows of dummy-load units instead of flows of actual load units. These numbers of dummy-load units in their turn are the rounded-up values of the flows of actual load units, that flow through a separate network. Refer to figure 7.77

\( ^{38} \)Which location must be served by that equipment pool.

\( ^{39} \)These are the flows through the load unit sub-model. At this point the load unit- and equipment sub-model are actually coupled.

\( ^{40} \)Next section explains how in some cases these side-constraint may over more than one time window.
for a better understanding of the side-constraints that convert the actual load units into dummy-load units shown in formula (7.21).

\[
lu_{-dlutn_w,p,s} \geq et_{-ccp} \cdot lu_{-tn_w,p,s}
\]  
(7.21)

Analogous to formula (7.20) for single-load unit equipment types, the side-constraint for self-propelled MLU-equipment becomes:

\[
lu_{-dlutn_w,p,s} = w_{-dt_w} \cdot \sum_{o_e} epa_{-n_w,p,s,o}
\]  
(7.22)

* **Total side-constraint - pulling and trailer-type equipment**

In case of pulling equipment we have to deal with two different types of side-constraints rather than with one. The first type of side-constraint matches the flows of pulling machines with the flow of (dummy-)load units, while the second type of side-constraint matches the number of trailers with the number of pulling machines. Both types of side-constraints are affected by the use of simplified double relocation; the first type of side-constraint is also affected by the carrying capacity of the trailers, depending on their MLU-character.

At figure 7.90 we have seen that at each section any combination of a pulling pool with a trailer-type pool is unique: Each combination is modelled at most once at each section, even if pools are double installed at several networks covering that section. We also have seen that for each such unique combination a separate arc in the load unit sub-model is created, expressing that such a combination is considered as a single equipment pool in the load unit sub-model. The same holds for the equipment sub-model: Each combination of a pulling pool with a trailer-type pool is considered at some points as one equipment pool. To stress this aspect, at some points this unique combination of a pulling pool with a trailer-type pool at a section is indicated as a single pool named \texttt{p_pull\_trail}, which in fact is a fictitious equipment pool as shown in table 7.2, consisting of trailers stemming from equipment pool \texttt{p\_trail} and of pulling equipment stemming of equipment pool \texttt{p\_pull}. These various names for equipment pools will be used in the remainder of this section without further explanation. In some cases the index of variables will be \texttt{p\_pull\_s, p\_trail} which indicates that the variable refers to a flow of the pulling equipment pool in service of the trailer-type pool.

As for each unique combination of pulling pool and trailer-type pool a separate arc is created in the load units sub-model, the single index \texttt{p\_pull\_trail} will be used when referring to flows of (dummy-) load units and to side-constraint factors, while the composed index \texttt{p\_pull\_s, p\_trail} will be used when referring to flows of pulling equipment in service of the trailer-type pool.

There is a difference in the side-constraints for pulling equipment that uses simplified double relocation and for pulling equipment that does not. If simplified double relocation is not used, then we have found that for each collection node from which trailers could be retrieved one arc for relocation of trailers, and a number of arcs for double relocation of pulling equipment is generated. In case of three collection nodes the trailers can be relocated over three arcs, while the pulling equipment can be double-relocated over nine arcs. Refer to figure 7.78. Also, we have seen that the network for the trailers contains precisely the same collection nodes as the network for the pulling equipment, as both pools are forced to be installed at the same equipment networks. One important aspect to bear in mind is that for each network at which pulling and trailer-type equipment is installed a separate set of arcs can be created. Refer to figure 7.90 for more information on this subject.

So, in total four different sets of side-constraints can be given, each for another combination of simplified double relocation and of MLU-type trailers. First all these side-constraints are discussed for each of these sets, after which the final results are shown in table 7.12.
Case I
In the first case to discuss simplified double relocation is not used and the pulled trailers are no MLU-trailers. In that case we need only two side-constraints: One to match the load units with the pulling equipment and one to match the pulling equipment with the trailers. This yields following two side-constraints:

\[
lu_{\text{tn}_{w_{,}}_{p_{-}}_{\text{pull}_{-}}_{\text{trails}}} = w_{-} dt_{w} \cdot \sum_{o_{,} od} \text{epa}_{_{n_{w_{,}}_{p_{-}}_{\text{pull}_{-}}_{s_{,}}_{o_{,}}_{od_{,}}_{p_{-}}_{\text{trail}}} (7.23)
\]

\[
\text{epa}_{_{n_{w_{,}}_{p_{-}}_{\text{trails}_{,}}_{o_{-}}_{\text{trail}_{,}}_{p_{-}}_{\text{pull}}} = \sum_{od=0_{-}\text{trail}} \text{epa}_{_{n_{w_{,}}_{p_{-}}_{\text{pull}_{-}}_{s_{,}}_{o_{,}}_{od_{,}}_{p_{-}}_{\text{trail}}} (7.24)
\]

in which:

- \(o_{-}\text{trail}\) = origin collection node of trailer-type equipment pool
- \(p_{-}\text{pull}_{-}\text{trail}\) = fictitious, combined equipment pool of trailer-type pool and pulling pool
- \(p_{-}\text{trail}\) = trailer-type equipment pool
- \(p_{-}\text{pull}\) = pulling equipment pool.

The double index of these parameters, \(p_{-}\text{pull}_{-}\text{,...}_{p_{-}}_{\text{trail}}\) is used to express that they hold for the unique combination that the pulling pool is pulling the trailer-type pool for transport over one section; for any other combination of pulling pool and trailer-type pool other parameters may hold, either due to different speeds of the pulling pool or due to a different carrying capacity of the trailer-type pool. Each unique combination of \(p_{-}\text{pull}_{-}\text{,...}_{p_{-}}_{\text{trail}}\) refers to a different arc both for the trailer-type pool and for the pulling pool. In the remainder four cases this property is considered as known without further explanation.

Note that in this formula three origins, \(o_{,} od\) and \(o_{-}\text{trail}\) respectively, are included. The first one is the origin of the pulling equipment pool, the second one is the origin of the trailer-type equipment pool, and the third one is the collection node over which the pulling equipment double relocates, analogous to figure 7.90. Only those arcs for which the values of \(od\) and \(o_{-}\text{trail}\) are equal can be matched. In this formula each combination of a trailer-type pool with a pulling pool in unique for each section, even if this combination is double installed at the equipment networks serving that section. Refer to figure 7.90 for more details on this subject.

Case II
In the second case to discuss simplified double relocation is used, but the trailers are still no MLU-trailers. Even in that case we need only following two side-constraints:

\[
lu_{\text{tn}_{w_{,}}_{p_{-}}_{\text{pull}_{-}}_{\text{trails}}} = w_{-} dt_{w} \cdot \sum_{0} \text{epa}_{_{n_{w_{,}}_{p_{-}}_{\text{pull}_{-}}_{s_{,}}_{o_{,}}_{p_{-}}_{\text{trail}}} (7.25)
\]

\[
\sum_{o_{-}\text{trail}} \text{epa}_{_{n_{w_{,}}_{p_{-}}_{\text{trails}_{,}}_{o_{-}}_{\text{trail}_{,}}_{p_{-}}_{\text{pull}}} = \sum_{o_{-}\text{pull}} \text{epa}_{_{n_{w_{,}}_{p_{-}}_{\text{pull}_{-}}_{s_{,}}_{o_{,}}_{p_{-}}_{\text{pull}}_{,}}_{p_{-}}_{\text{pull}}} (7.26
\]

in which:

- \(o_{-}\text{trail}\) = origin collection node of trailer-type equipment pool
- \(o_{-}\text{pull}\) = origin collection node of pulling equipment pool

The rather strict coupling of trailer-type equipment with pulling equipment in side-constraint (7.24) can not be guaranteed in side-constraint (7.24); this is the foretold rather "loose" coupling between these two pools due to simplified double relocation.

Case III
In the third case to discuss simplified relocation is not used analogous to Case I, but the trailers are MLU-type trailers, able to carry more than one load unit at the time. Therefore, an additional side-constraint is needed compared to Case I to convert the flows of actual load units into flows of dummy-load units. Therefore, the side-constraints for this case become:

\[
l u_{-}\text{dlut}_{w_{,}}_{p_{-}}_{\text{pull}_{-}}_{\text{trails}} \geq et_{-}\text{ccp}_{_{p_{-}}_{\text{trail}}_{,}} lu_{-}\text{tn}_{w_{,}}_{p_{-}}_{\text{pull}_{-}}_{\text{trails}}, (7.27)
\]
\[ lu_{\text{dltn}}_{w,p\_\text{pull}\_\text{trails}} = w_{\text{d}t_w} \sum_{o,od} epa_{n_{w,p\_\text{pull}s,o,od,p\_\text{trail}}} \]  
(7.28)

\[ epa_{n_{w,p\_\text{trails},o\_\text{trail}}p\_\text{pull}} = \sum_{od=o\_\text{trail}} epa_{n_{w,p\_\text{pull}s,o,od,p\_\text{trail}}} \]  
(7.29)

in which:
- \( o\_\text{trail} \): origin collection node of trailer-type equipment pool
- \( o\_\text{pull} \): origin collection node of pulling equipment pool.
- \( p\_\text{pull}\_\text{trail} \): fictitious, combined equipment pool of trailer-type pool and pulling pool.

In these formulas the composed index \( p\_\text{pull}\_\text{trail} \), which was found at the load unit sub-model to express that the flow of load units was transported by a fictitious pool that is the unique combination of one particular trailer-type pool with one particular pulling pool, installed at the same network, is also used for the flows in the dummy-network. Quite logical, actually, as the flows in the dummy-load unit network were related to flows of actual load units in the load unit sub-model through the pool that transported those (dummy-) load units. So, if the equipment pool that transports the actual load units is a fictitious pool, the equipment pool that transports the corresponding dummy-load units must be that same, fictitious, pool too.

♦ Case IV

In the fourth case to discuss simplified relocation is used and trailers are MLU-type trailers. So, in this case we need an additional side-constraint to match the flows of actual load units to the flows of dummy-load units. The side-constraints for this case are:

\[ lu_{\text{dltn}}_{w,p\_\text{pull}\_\text{trails}} \geq et_{ccp\_\text{trail}} lu_{\text{fn}}_{w,p\_\text{pull}\_\text{trails}}. \]  
(7.30)

\[ lu_{\text{dltn}}_{w,p\_\text{pull}\_\text{trails}} = w_{\text{d}t_w} \sum_{o} epa_{n_{w,p\_\text{pull}s,o,p\_\text{trail}}} \]  
(7.31)

\[ \sum_{o\_\text{trail}} epa_{n_{w,p\_\text{trails},o\_\text{trail}}p\_\text{pull}} = \sum_{o\_\text{pull}} epa_{n_{w,p\_\text{pull}s,o\_\text{pull}}p\_\text{trail}} \]  
(7.32)

in which:
- \( o\_\text{trail} \): origin collection node of trailer-type equipment pool
- \( o\_\text{pull} \): origin collection node of pulling equipment pool.
- \( p\_\text{pull}\_\text{trail} \): fictitious, combined equipment pool of trailer-type pool and pulling pool.

<table>
<thead>
<tr>
<th>Simplified double relocation</th>
<th>YES</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLU-equipment</td>
<td>(7.30), (7.31) and (7.32)</td>
<td>(7.27), (7.28) and (7.29)</td>
</tr>
<tr>
<td>(7.25) and (7.26)</td>
<td>(7.23) and (7.24)</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.12: *Side-constraints for various settings for pulling- and trailer-type equipment.*

Table 7.12 shows a summary of these side-constraints for the various combinations of settings for pulling and trailer-type equipment.

One aspect influences the construction of side-constraints for pulling- and trailer-type equipment: As trailer-type equipment pools are only allowed to be pulled by pulling pools of the proper type that are installed at the same equipment network, an ambiguous situation may occur in case of overlapping equipment networks. This will be clarified with following example. Suppose that at one network named I one pulling pool named \( p\_\text{A} \) and two trailer-type pools named \( \text{trail}_X \) and \( \text{trail}_Y \) are installed. Also suppose that at another network named II the same pulling pool \( p\_\text{A} \) is installed together with pulling pool \( p\_\text{B} \), along with trailer-type pool \( \text{trail}_Y \). Refer to the alternative situations discussed at figure 7.90. In that case, side-constraint (7.26) results in following three separate side-constraints:
\begin{align*}
\sum_{\text{trail}} \text{epa}_{n,w,\text{trail},X,s,o,\text{trail pull}_A} &= \sum_{\text{trail}} \text{epa}_{n,w,\text{pull}_A,s,o,\text{trail}_X} \\
\sum_{\text{trail}} \text{epa}_{n,w,\text{trail},Y,s,o,\text{trail pull}_A} &= \sum_{\text{trail}} \text{epa}_{n,w,\text{pull}_A,s,o,\text{trail}_Y} \\
\sum_{\text{trail}} \text{epa}_{n,w,\text{trail},Y,s,o,\text{trail pull}_B} &= \sum_{\text{trail}} \text{epa}_{n,w,\text{pull}_B,s,o,\text{trail}_Y} 
\end{align*}

(7.33)
(7.34)
(7.35)

These three formulas show that overlapping networks do not affect the guarantee that for each unique combination of pulling pool and trailer-type pool only one separate arc is created\textsuperscript{41}. But, as the number of arcs and side-constraints may increase rapidly by installing different combinations of equipment pools at overlapping equipment networks, these overlapping equipment networks should be avoided as much as possible in order to keep the optimisation model from exploding.

The most complicated structure of side-constraints is found in cases III and IV in which MLU-trailer-type equipment is modelled. Referring to table 7.12 we need three different side-constraints to model this situation.

The various flows needed to model this situation are depicted in figure 7.92, showing four different flows and three different relations marked 1, 2, and 3. Relation 1 models the conversion of flows of actual load units into the rounded-up value to match the carrying capacity et_cc of the considered trailer-type equipment. This relation uses side-constraint (7.27) or (7.30). Relation 2 models the conversion of the flows of dummy load units into pulling equipment that may double-relocate from various collection nodes. This relation uses side-constraint (7.28) or (7.31). Relation 3 models matching of pulling equipment with trailer-type equipment. This relation uses side-constraint (7.29) or (7.32).

\textbf{Multiple time windows}

It has been shown before that in some cases the transport time for some section may require more time than the time window load units are transported offers. This may be caused either by long (double) relocation times, by long transport or transhipment times, or by limited working hours, relative to the time window's length. This yields two problems. (1) Flows of equipment may cover more than one time window, and (2) it is ambiguous in which of these multiple time windows the load units are to be considered as actually transported in the load unit sub-model. These two problems will be tackled sequentially.

41]

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure7.92}
\caption{Relation of flows for MLU-trailer-type equipment.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure7.93}
\caption{Transport time exceeding time window.}
\end{figure}

(1). The first problem, how to model multiple window transports in the equipment sub-model, is easily solved by creating arcs that cover more than one time window. These arcs must cover just enough time windows to cover a time span that exceeds\textsuperscript{42} the transport time multiplied by the factors ep_wtf modelling the working hours in each time window. So, if we consider an arc that starts in time window \texttt{w_start} then we find the corresponding window \texttt{w_end} to end the arc in as the first window that matches the inequality:

\[ \texttt{w}_\text{end} \geq \texttt{w}_\text{start} + \sum_{\text{time window}} \text{ep}_\text{wtf} \times \text{transport time} \]

\textsuperscript{42}This time may not equal the sum of the duration of each time window, as such an arc would be valid for one transport movement only, starting at the very beginning of the first time window.
\[
\sum_{w = \text{w}_{\text{start}}}^{\text{w}_{\text{end}}} \text{ep}_{\text{wt}, w, p} > \text{ep}_{\text{tt}, p, s, o, \text{start}}
\]

(7.36)

or, for pulling equipment using non-simplified double relocation:

\[
\sum_{w = \text{w}_{\text{start}}}^{\text{w}_{\text{end}}} \text{ep}_{\text{wt}, w, p} > \text{ep}_{\text{tt}, p, \text{pull_trails}, o, \text{od}}
\]

(7.37)

The resulting arc is shown in figure 7.93. It shows how the transport time exceeds the duration of the first time window, thus forcing the transport to end in the second time window. In this example the factors modelling the working hours are considered equal in both time windows, so that we can focus on the transport time only. This problem could not be solved before the transport time \( \text{ep}_{\text{tt}} \) was defined and determined. Therefore, the equipment model that was shown before might need some adaptations for some of the smaller time windows as some of the arcs might turn out to be multiple-time window arcs. In case a multiple-window arc covers two different planning phases, then several collection nodes of the latter planning phase might be fed by several (multi-window) arcs that stem from the one and the same location, but from different time windows. This is shown in figure 7.94 in which two different arcs stem from different collection nodes at the same location A, but from different time windows. All these same (collection) node at location B in the second planning phase. This effect is caused by the differences in time window length for different planning phases, making inequality (7.36) or (7.37) point to the same ending window \( \text{w}_{\text{end}} \) for different starting windows \( \text{w}_{\text{start}} \).

(2). The second problem, which arc in the load unit sub-model should be matched with the multiple-time window arc in the equipment sub-model, is solved by considering all possibilities and their consequences. First we need to realise that arcs modelling flows of load units do not skip time windows in the load unit model. Therefore, we must select either the first or the last time window of the multiple-window equipment arc as the arc to which the load unit flow must be matched. Suppose that the first time window was selected. In that case the load units were considered to be transported sooner than they actually are, thus yielding a too optimistic solution. Furthermore, the time needed for (double) relocation would be counted as actual transport time, too. This situation obviously is not desirable, and therefore the last window over which the equipment arc runs is selected to match the flows of load units and of equipment. So, in all the side-constraints matching flows of load units to flows of equipment, the index \( w \) of the flow of equipment must be replaced by \( \text{w}_{\text{start}} \), such that the ending window \( \text{w}_{\text{end}} \) of each arc equals \( w \).

A ambiguous situation might occur if for some multiple-window equipment arc the starting window \( \text{w}_{\text{start}} \) is the last time window that is defined; in that case it is not possible to find a valid ending time window \( \text{w}_{\text{end}} \), there is no "next" time window. Fortunately this problem will never occur as it would imply an error in the modelled system: It would imply a transport time that exceeds 8 hours (refer to 7.1, planning phase 3) which can be regarded as invalid for any terminal complex that is intended to be modelled with the terminal design aiding tool described in this thesis.
Balanced flows

![Diagram of balanced flows]

**Figure 7.95:** Example of balancing flows of equipment.

Some errors are introduced with respect to "return trips" on a section of an equipment routing. This will be clarified by referring to figure 7.95 in which 50\(^{43}\) load units are transported from A to B, and 50 other load units vice versa, using the same equipment pool. In that case the optimisation model partially ignore the effect of re-use of equipment that has been used for transport from A to B, for reverse transport from B to A. But fortunately these errors are slight as will be argued in following paragraphs.

First of all, we need to distinguish between two different situations: (1) a "steady-state" situation and (2) a "transient" situation.

(1) The "steady-state" situation is defined as a situation in which the demands for equipment do not differ much between two time windows. In other terms, the flows of load units from A to B and from B to A are more or less the same in two subsequent time windows\(^{44}\). In that case the flows of equipment could be as shown in figure 7.96 in which equipment that has been used in a time window previous to the one shown for transport from B to A, and therefore is collected at the collection node at A, is re-used in the next window, which is shown, for return-transport from A to B. So, in a steady-state situation of subsequent time windows the aspect of combined "return-trips" is fully satisfied.

(2) The "transient" situation is defined as any situation not equal to the "steady-state" situation, and therefore requires equipment relocation to meet the demands of 50 machines per hour in both directions. This is shown in figure 7.97 in which a flow of only 7 machines per hour is available at the collection node at A. The remaining (25\( - 7\))=18 machines per hour must be relocated from another collection node, which possibly could even be location B, as shown in figure 7.97. The result of this unbalance between location A and B is therefore an additional relocation of equipment, and therefore higher total costs.

The simplification that is included in the equipment optimisation model is found when the additional 18 machines required at location A could be retrieved from location B through equipment relocation. In that case, in reality, the machines that were used for transport from B to A would be re-used for transport from A to B again, provided that the transport time from B to A and backwards does not exceed the time window's length.

However, the equipment optimisation model is not in the position to re-use equipment within one time window,

\(^{43}\)This number of load units is determined dynamically by the equipment optimisation model itself, and therefore is the result of a complex equilibrium with the equipment sub-model.

\(^{44}\)A "steady-state" situation as meant here could also be reached through equipment relocation from other nodes than A and B to B and A respectively.
and therefore has to relocate equipment from B to A, thus inducing extra empty travelling distances. So, in that case the total costs of the equipment optimisation model's solution would be a little higher than absolutely necessary as the costs for relocation would be included in the total costs, while in reality there would not have been any need for relocation.

♦ Preference for wide time windows
Any optimisation model that includes time windows of different lengths could have an invalid preference for wide time windows, or, on the contrary, for narrow time windows. It will be shown here that the equipment optimisation model described in this thesis is not sensitive for differences in duration of different time windows in two steps.

First, at the load unit sub-model we found that the stacking costs are purposely independent of the time window in which the load units are stacked. Therefore, differences in stacking costs can only be found at different stacking locations, rather than at different time windows for the same stacking location.

Second, at the equipment sub-model we found a somewhat more complex structure of costs and of side-constraints matching flows of load units to flows of equipment. Consider a situation in which a constant flow of load units runs over a section of one closed task through a number of time windows of different lengths. A fair comparison between two time windows of different lengths can only be made when this flow of load units is constant per time unit. Therefore, consider the situation in which a constant flow of load units that equals 15 load units per hour, flows through two different time windows, lasting 2 and 6 hours respectively. Therefore, the total numbers of load units transported in both time windows equal (2*15)=30 and (6*15)=90 respectively. As we have found at the equipment sub-model, these numbers of load units are converted into flows of machines to be assigned to these load units through side-constraints like the one shown in formula (7.25). These side-constraints specify that the flow of machines assigned to a number of load units is found by dividing this number of load units by the duration of the time window. So, for both time windows we would find a flow of machines\(^\text{45}\) that equals (30/2)=15 and (90/6)=15 for both time windows respectively: There is no difference in the flow of equipment assigned in both time windows. The costs involved with any assignment of equipment to load units are related to the unit of flow of machines in the equipment model, and therefore are related to the just found values of (30/2)=15 and (90/6)=15 machines per hour. By consequence, as the flows of machines found in both time windows are equal, and the costs per unit of flow of machines are equal in all time windows\(^\text{46}\), the total costs of assigning equipment to load units is equal in both time windows. In this way it has been shown that the equipment optimisation model is not sensitive for varying lengths of time windows.

♦ Simplification by multiple transports through one time window
An apparent drawback of INTRA-window equipment hopping is the effect that one single load unit can be transported over a number of subsequent sections during one time window. In those cases where the total transport time over all these sections comes near to, or even exceeds the duration of the time window, an erroneous result is produced by the optimisation model. This drawback is perceived and will be solved by introducing inter-window equipment hopping in the next section. This section also reformulates the just described drawback in a more extensive way.

7.4.4. Inter-window equipment hopping
Finding an optimal flow of load units and equipment on the long term as described at INTRA-window equipment hopping at an acceptable speed of calculation is only possible when all those flows are condensed into a limited number of discrete time windows as previously described: Considering all transports up to three days ahead individually would yield unmanageable calculation times of a few hours, while resulting in too detailed output data. However, using discrete time windows also has some disadvantages as they introduce some simplifications and by consequence some errors.

\(^45\)Ignoring the effect of (double) relocation at this point.
\(^46\)Ignoring the effect of various working hours.
An example of such a simplification is shown in figure 7.98, showing two different transports of load units on the same section by the same equipment pool. By consequence, both transports last equally long as is shown in the figure. But although both transports start in the same time window, they do not end in the same time window; INTRA-window equipment hopping ignores this effect by presuming that all transports start at the very beginning of the first time window -- and therefore all end in the second time window.

Another example of such a simplification is shown in figure 7.99, showing two subsequent transports on an equipment routing that are separated by an exchange point. Both transports individually last less than the shown time window does; but summed up they exceed that time window's duration. Therefore, a single load unit that is transported from A to B during the shown time window could not be transported from B to C in that same time window; this is physically impossible. However, although physically impossible, INTRA-window equipment hopping does allow such an assignment\(^7\) which is especially severe for the first planning phase in which the time windows are small and a situation as shown in figure 7.99 is likely to occur.

To fix these errors the concept of inter-window equipment hopping is introduced. The basic idea about this concept is shown in figure 7.100 by the arcs marked with the sign "\(\oplus\)". Inter-window equipment hopping does not only allow internal transport within a time window but also across a time window. By consequence, in case of inter-window equipment hopping flows of load units do not necessarily need to cross the border of a transport window by stacking only; they also can cross the border of a time window by internal transport.

\(\oplus\) = arc specific for inter-window equipment hopping

---

\(^7\)So, assignment both of transport A-B and of transport B-C of one single load unit to the same time window.
Inter-window equipment hopping imposes some restrictions on the numbers of load units that are transported over on the inter- and on the intra-window arcs (and therefore on the number of equipment that is used). Inter-window equipment hopping forces the flows over both arcs to have a fixed ratio which is determined by:

1. the length of the starting time window,
2. the duration of the internal transport movement, and
3. the working hours in the starting time window.

In the example shown in figure 7.101 a ratio of 60/40% is used, which means that the number of load units transported within the first time window equals (60/40)=1.5 times the number of load units transported between the first and second time window.

It will be shown that this fixed ratio is imposed by means of a side-constraint for flows in the load unit sub-model only; it is not imposed for flows in the equipment sub-model. Nevertheless, the structure shown in figure 7.101 of inter- and intra-window arcs is found both in the load unit- and in the equipment sub-model.

Inter-window equipment hopping will be able to fix the previously signalled disadvantages of intra-window equipment hopping, but in the meantime introduces another (possible) error. It introduces a fixed ratio between inter- and intra-window flows of load units which in the actual situation does not necessarily need to be valid, thus possibly yielding an infeasible model. This drawback and the solution to it will be discussed later on in more detail.

7.4.4.1. Load unit sub-model

First of all, the effects of inter-window equipment hopping on the load unit sub-model will be shown. These effects are twofold: (1) First of all, the number of arcs is increased analogous to the structure shown in figure 7.101; for each arc modelling internal transport within a time window in the intra-window equipment hopping load unit sub-model a new, slant, arc is created that models internal transport between two time windows. (2) Second, an additional side-constraint is added to impose a fixed ratio on the flows over these two intra- and inter-window arcs.

Finding the ratio between the inter- and intra-window arc

The ratio between each pair of intra- and inter-window arcs that model one internal transport movement is determined based on figure 7.102, and will include:

1. the length of the starting time window,
2. the duration of the internal transport movement, and
3. the working hours in the starting time window.

Figure 7.102: Finding the ratio between the inter- and intra-window arc.

Figure 7.102 shows three time windows (the first of which is split into two parts $t_2$ and $t_3$) and one arc modelling the transport time of some internal transport movement\(^{49}\). It is this particular internal transport movement for which the ratio of the inter- and intra-arc as shown in figure 7.101 is to be determined. This will be done through both parameters $t_2$ and $t_3$.

---

\(^{48}\) Not for arcs modelling stacking.

\(^{49}\) Which could either be solely a transhipment movement or an actual internal transport movement.
Chapter 7. Equipment

Consider the internal transport movement starting at the beginning of the first time window of Figure 7.102, which is the upper time window of Figure 7.101. Adding the internal transport movement's transport time to the start time of this very first time window yields the ending time of this particular internal transport movement; the time window in which this ending time lies is called the second time window, which is the upper time window of Figure 7.101, too. (Note in this particular case the second time window is equal to the first time window; this is the case whenever the internal transport movement's transport time is less than the first time window's duration).

As the second time window is known, its end time is known too, which also is the start time of the third time window, which is the lower time window of Figure 7.101. Subtracting the internal transport movement's transport time from the end time of the second time window as shown in Figure 7.102 yields the subdivision of the first time window into $t_2$ and $t_3$. The meaning of these two times $t_2$ and $t_3$ is as follows: Any internal transport movement starting in the first time window before time $t_2$ will end in the second time window; and any transport movement starting in the first time window at or after time $t_2$ will end in the third time window.\(^{50}\)

Now the ratio between the flows over the inter- and intra-window arc is determined as the ratio of the working hours within both parts of the first time window: $t_2$ and $t_3$, rather than as the ratio of the entire duration of both parts.

An example hereof is shown in Table 7.13 in which $t_2$ and $t_3$ last 1.3 and 0.7 hours respectively. As these two times sum up to 2.0 hours the first time window of both Figures 7.101 and 7.102 apparently equals 2.0 hours in this example. Table 7.13 shows that the working hours in both periods are totally different; period $t_3$ fully concerns working hours, while period $t_2$ only partially concerns working hours. By consequence, the flow of load units that is to flow to the third time window is $(77.8/22.2)=3.5$ times the flow of load units that is to flow to the second time window.

<table>
<thead>
<tr>
<th></th>
<th>Duration [hrs]</th>
<th>Of which: working hours [hrs]</th>
<th>Percentage [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_2$</td>
<td>1.3</td>
<td>0.2</td>
<td>22.2%</td>
</tr>
<tr>
<td>$t_3$</td>
<td>0.7</td>
<td>0.7</td>
<td>77.8%</td>
</tr>
</tbody>
</table>

Table 7.13: Finding the ratio between the inter- and intra-window arc.

Based on these data the side-constraint modelling the fixed ration between the inter- and intra-window arc can be found as:

$$L_{u \cdot t_{n_{w_3, p, s}}} = \left( \frac{E_{p \cdot w_{t_{w_3, p}}}}{E_{p \cdot w_{t_2, p}}} \right) \cdot L_{u \cdot t_{n_{w_2, p, s}}}$$  \(7.38\)

in which:

- $L_{u \cdot t_{n_{w_3, p, s}}}$ = number of load units starting to be transported at time window $w$, being transported by equipment pool $p$ over the intra-window arc of section $s$, and ending in the second time window of figure 7.102.
- $L_{u \cdot t_{n_{w_2, p, s}}}$ = number of load units starting to be transported at time window $w$, being transported by equipment pool $p$ over the inter-window arc of section $s$, and ending in the third time window of figure 7.102.
- $E_{p \cdot w_{t_{w_3, p}}}$ = working hours of equipment pool $p$ during the part $t_2$ of time window $w$.
- $E_{p \cdot w_{t_{w_3, p}}}$ = working hours of equipment pool $p$ during the part $t_3$ of time window $w$.

In those cases where both $t_2$ and $t_3$ equal zero (that is, in those cases where there are no working hours at all present within the first time window) not a single arc is created stemming from the first time window. This is in correspondence with the load unit sub-model of INTRA-window equipment hopping.

\* Minimum-ratio threshold

In some cases the ratio between the inter- and intra window arc is relatively small (5% or less for one arc). Including an additional side-constraint for such relatively small flows is considered not worthwhile the additional calculation effort. Therefore, a minimum-ratio threshold is introduced that prevents the creation of

\(^{50}\)This also is the reason why $t_2$ and $t_3$ are called that way: Transports starting during $t_2$ end in window 2, and transports starting in $t_3$ end in window 3.
both an inter- and an intra-window arc, and the creation of a side-constraint matching these two arcs; in any case where either:
1. \( \text{ep}_{\text{wp2,p}} \) is less than this threshold times the sum of \( \text{ep}_{\text{wp2,wp3,p}} \) and \( \text{ep}_{\text{wp3,p}} \), or where
2. \( \text{ep}_{\text{wp3,p}} \) is less than this threshold times the sum of \( \text{ep}_{\text{wp2,wp3,p}} \) and \( \text{ep}_{\text{wp3,p}} \).
the arc corresponding with the smallest flow is omitted from the model, along with the related side-constraint.

The value of this minimum-ratio threshold may be specified by the user, but is advised to be about 5%. In some cases it may be decided to increase this minimum-ratio threshold up to 10%. Especially in cases where pure transhipment movements are modelled, which tend to last relatively short compared to the time window's length, this minimum-ratio threshold is likely to be useful.

**Simplifications in inter-window equipment hopping**
From the above discussion it can be deduced that inter-window equipment hopping does take into account the fact that several transport movements starting from the same time window may end in different time windows, but also silently makes the assumption that all transport movements start equally divided over the available working hours within the starting time window. In practice, this may not be the case. In practice it may even turn out that, in the example at table 7.13, all transports start during time period t2 and none starts during time period t3, irrespective the major differences in available working hours between these two periods. However, although this assumption may have an important impact on the results of the optimisation, it is considered as a reasonable assumption.

**7.4.4.2. Equipment sub-model**

When discussing INTRA-window equipment hopping it turned out that the load unit sub-model was to be considered as the master sub-model, and that the equipment sub-model was to be considered as the slave sub-model. This property still hold for inter-window equipment hopping, and even shows through in a more apparent way, both in the model's structure of arcs and nodes and in the presence of side-constraints.

**No additional side-constraints**
As shown at the load unit sub-model, the side-constraints that match the flows over the inter-window arcs with the flows over the corresponding intra-window arcs are included in the load unit sub-model. By consequence, there is no need to include that type of side-constraints in the equipment sub-model. An advantage of this choice is that the number of side-constraints remains limited, as the load unit sub-model does not include various arcs for equipment relocation\textsuperscript{51}.

**Same structure**
The structure of the equipment sub-model for inter-window equipment hopping can be considered as a mix between the structure of:
1. the equipment sub-model of INTRA-window equipment hopping, and of
2. the load unit sub-model of inter-window equipment hopping.
This means that for each arc that exists in the previously described equipment sub-model of INTRA-window equipment hopping two different arcs may be created for inter-window equipment hopping: One for the intra-window flow and one for the inter-window flow.

Note that the ratio of flows of over an intra- and its corresponding inter-window arc for load units is automatically found in the equipment sub-model, too. (That is, summed over all equipment pools that can be used for the same section, and ignoring the aspect of multiple-load unit equipment). By consequence, transport of machines through time windows is modelled by inter-window equipment hopping, too.

In case of pulling equipment types a number of additional arcs is created, compared to INTRA-window equipment hopping: First of all the number arcs for pulling equipment are doubled, and second the number of arcs for trailers are doubled, as to each flow of pulling equipment a flow of trailers is related. Also, in case of

\textsuperscript{51}In this way the duration of equipment relocation is left out of the fixed ratio between inter- and intra-window flows. This is obvious, as this ratio was intended to control the flow of load units in a more accurate way, and not the flow of equipment. Equipment merely has to follow these flows of load units, and relocate where needed.
MLU-equipment the arcs for dummy-load units are doubled too, to match up with the total arcs for MLU-equipment.

Costs
There is very little difference between the costs of flows of equipment in inter-window equipment hopping from those in INTRA-window equipment hopping. In both models they are used to include (1) operational costs of equipment and (2) "peakedness" of the use of equipment in the optimisation model. In both models these costs consist of the elements:
1. preference\(^{52}\);
2. transportation time;
3. working hours,
As discussed at INTRA-window equipment hopping. Only the latter cost element changes somewhat for inter-window equipment hopping: This cost element is calculated separately for the inter- and for the intra-window arc of figures 7.100 and 7.101 to model differences in working hours and in "peakedness" for both periods \(t2\) and \(t3\) from figure 7.102.

By consequence, the working time factor \(\text{ep}_\text{wtf}_{wp}\) will be different for both the inter- and for the intra-window arc that stem from the same time window, and will be determined in precisely the same way as previously discussed at INTRA-window equipment hopping; only the duration of the time window is limited to either period \(t2\) or period \(t3\) rather than to the entire time window. This will be shown with the example shown in table 7.13.

In case of the example of table 7.13 the working time factor is determined according to table 7.14, both for the inter- and the intra-window arc. Note that there may be a major difference between these two working time factors, thus yielding a clear distinction between arcs that yield "smooth" and arcs that yield "peaked" usage of equipment of one pool.

Because of the additional side-constraint (7.38) for inter-window equipment hopping, the flow over both arcs -- with different costs -- are forced to have a fixed ratio. Therefore, if equipment is assigned to the relative "cheap" inter-window arc of table 7.14, then the optimisation model is forced to assign 1/3.5 times that flow to the relative "expensive" and therefore unattractive intra-window arc, too. In this way the optimisation model is encouraged to search for a competing equipment pool that has a less distinct difference between the "peakedness" and therefore of the costs of its inter- and its intra-window arc.

<table>
<thead>
<tr>
<th></th>
<th>Duration [hrs]</th>
<th>Of which: working hours [hrs]</th>
<th>(\text{ep}<em>\text{wtf}</em>{wp})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(t2)</td>
<td>1.3</td>
<td>0.2</td>
<td>(\frac{1.3}{0.2} = 6.5)</td>
</tr>
<tr>
<td>(t3)</td>
<td>0.7</td>
<td>0.7</td>
<td>(\frac{0.7}{0.7} = 1.0)</td>
</tr>
</tbody>
</table>

Table 7.14: Different values of \(\text{ep}_\text{wtf}_{wp}\) for the inter- and intra-window arc.

7.4.4.3. Coupled load unit sub-model and equipment sub-model
The idea of coupling the load unit sub-model to the equipment sub-model for inter-window equipment hopping is completely analogous to the previously discussed coupling for INTRA-window equipment hopping. The only difference is that some more arcs and some more side-constraints are included: Arcs to model inter-window flows as shown in figure 7.100, are new in inter-window equipment hopping and therefore induce new side-constraints to match the flows of equipment over these arcs to flows of load units over the corresponding arcs.

As the same structure is used for both optimisation models, the structure of coupling the load unit- and the equipment sub-model in case of inter-window equipment hopping will not require more discussion.

7.4.5. Infeasibility
Both optimisation models, i.e. INTRA- and inter-window equipment hopping, may yield an infeasible solution, which means that the optimisation problem can not be solved due to a too small search space. Several reasons

\(^{52}\)According to the considered equipment routing.
or this infeasibility can be found, some of them may occur both in intra- and in inter-window equipment hopping, and others of them are particular to inter-window equipment hopping only. The possible reasons for infeasibility are:

1. lack of equipment\(^{53}\);
2. lack of stacking capacity\(^{54}\);
3. invalid inter- and intra-window ratio assumption\(^{55}\).

All these reasons will be discussed one by one.

- **Lack of equipment**

The most obvious reason for an infeasible solution is the lack of equipment at some equipment pool. Although the equipment optimisation model is likely to use all available machines of the limiting pools all time, in some cases the number of equipment will turn out to be insufficient. In that case it is advisable to (1) increase the number of available machines in that pool, or to (2) set the pool's capacity to unlimited, thus allowing the equipment optimisation model to find the optimal number of machines to install.

Another cause for lack of equipment are invalid working hours, that limit the interchange between various equipment pools. Especially for exchange points this aspect is crucial.

- **Lack of stacking capacity**

A little less obvious than a lack of equipment is a lack of stacking capacity. In some cases a relatively small stack turns out a limiting bottleneck, blocking all flows that are to run over that stack. As soon as one or more load units are not in the position to reach the opposite side of the modelled terminal system in time (i.e. at their specified system-exit time) then the entire model is found infeasible. The solution to this problem is quite analogous to that of a lack of equipment: The stack's fixed stacking capacity should be increased, or the stacking capacity should be set to unlimited, thus giving "carte-blanche" to the equipment optimisation model.

As at least one location over which any equipment routing runs must provide unlimited stacking capacity to suit the flows of load units through the system, this reason for infeasibility is likely to occur only for a limited number of locations that have not previously been assigned an unlimited capacity.

Note that the first and the second reason for infeasibility may be strongly related in some cases: It may occur that a limited stack capacity causes delays in flows of load units at some times by holding these load units back for a rather long time. After stacking capacity has been freed, all load units that have been held back are likely to start flowing instantly, thus causing an extra peakedness of flows of load units. This extra peakedness of flows of load units in its turn causes peakedness of flows of equipment, which in some may result in a short of equipment. This lacking of equipment will only occur for those equipment pools that have a fixed, limited capacity.

- **Invalid inter- and intra-window ratio assumption**

In some cases it may turn out that the optimisation model can not find a feasible solution due to an invalid assumption on the ratio between flows over the inter- and over the intra-window arc. Obviously, this type of reason can only occur in case of inter-window equipment hopping as only this optimisation model uses this fixed ratio.

In seldom cases it may turn out that timing of flows of load units is thus unequally divided over the available time periods that a substantial greater or smaller number of load units is to flow over the inter-window arc than estimated by the fixed inter/intra-window ratio. Most of the times such an unequal division of flows over the available time periods will cause the equipment optimisation model to put off or to hasten some flows in order to match the required ratios and to satisfy the required system exit-times of load units. But in some cases there may be no room to shift flows in this way in time, resulting in an infeasible solution.

Whenever this reason for infeasibility is encountered, a feasible solution is searched by using intra-window equipment hopping instead of inter-window equipment hopping. True, when a feasible solution is found in this way it will not satisfy the required fixed inter- and intra-window ratio, but as this ratio was only an estimate this

---

53 Only for pools with a limited capacity.
54 Only for stacks with a limited capacity.
55 Only for inter-window equipment hopping.
is not considered as a disadvantage. What is considered as a disadvantage are the simplifications that are introduced by intra-window equipment hopping that were discussed at section "Inter-window equipment hopping" of this chapter.

### 7.4.6. Conversion of continuous results into discrete results

After equipment optimisation during simulation has ended, its results are exported to the actual simulation model in the form of flows of load units and of equipment as shown in figure 7.103. This simulation model first has to convert these flows of load units and equipment into individual, actual load units and machines before they can be used in the actual simulation process. This conversion corresponds with the lower arc of figure 7.103.

Two types of flows require two types of conversion: One for load units and one for equipment. These differences in conversion for both types of flows are a consequence of the selected unit for both flows, which in its turn is a consequence of the selected relation between these two flows: Flows of load units are considered as the master flows while flows of machines are considered as slave flows. This relation was also found in the different constraints for both sub-models for equipment optimisation: The load unit sub-model was constrained to satisfy the required time for system entry and exit for each individual load unit, while the equipment sub-model was constrained to match flows of machines to flows of load units.

Both types of conversion are discussed sequentially. These two types of conversion do not need to be performed for the entire duration of the planning interval; they only need to be performed for the duration of the planning interval which may never exceed the first planning phase.

**Conversion for load units**

The first type of conversion to discuss is the conversion of flows of load units that resulted from the equipment optimisation model into internal transport movements of actual, individual load units. Two aspects are important when considering this conversion:

1. all flows in the load unit sub-model were integer-type flows;
2. all flows in the load unit sub-model were separate for each individual equipment routing, i.e. for each direction over each closed task and for each individual equipment pool that can be used on each section of that equipment routing.

A consequence of these two aspects is, that it is relatively simple to convert the results of the equipment optimisation model into actual flows of load units: Those load units, flowing over a particular equipment routing, that are nearest to system-exit are to be transported first. This will be explained by an example.

Consider figure 7.104 in which two flows of load units originate from an internal stack, during one particular time window. Both flows are results of the equipment optimisation model. The first, horizontal flow specifies that in total 5 load units are to be transported out of the internal stack during that time window, while the second, vertical flow specifies that 10 load units are to remain stacked between this and next time window.

Supposing that all times for system exit for all 15 load units are known and as shown in table 7.15 then the five ones with the smallest system exit time are transported out of the stack while the other ten remain stacked. From this example it may be deduced that load units with the greatest system exit time are assigned to the stacking arcs while the load units with the smallest system exit time are assigned to the internal transport arcs. At least, for INTRA-window equipment hopping in which only one internal transport arc is defined.
Table 7.15: Example of 15 load units from which 5 are selected for internal transport as shown in figure 7.104, based on their system exit time.

In case of inter-window equipment hopping two different arcs for internal transport are discerned: One intra-window arc and one inter-window arc as shown in figure 7.105. In that case priority is given to inter-window transport over intra-window transport as indicated with both marks \(1\) and \(2\). This priority means that (1) those load units that have to very smallest system exit time are transported over the inter-window arc, that (2) those load units with the mid-smallest system exit time are transported over the intra-window arc, and that (3) those load units with the greatest system exit time are "transported" over the vertical stacking arc. All these numbers of load units are in correspondence with the numbers specified by the equipment optimisation model, for each of these arcs.

When discussing figure 7.104 the primary assumption was that it was known in detail which fifteen load units were to be assigned to either one of the arcs shown in that figure; however, it was not discussed yet how these fifteen load units have been found. This will be solved by referring to the dotted arcs that are shown in figures 7.104 and 7.105: These dotted arcs specify the flows of load units into the internal stack and therefore are to be used to determine precisely which load units are present at the internal stack and need to be assigned to one of the out-going arcs of that stack. All dotted arcs can originate from a limited number of "locations", namely:

1. a location on the same equipment routing located before this internal stack;
2. a location modelling the same internal stack at a previous time window;
3. a location modelling the initial stack size at the current simulation time.

It is at the first and at the last type of location that the precise time of system exit can get known: (1) If the location before the internal stack is a bounding element then all load units flowing over the internal stack are created at that bounding element and therefore will receive a value for their system exit time at that very bounding element. Whenever load units flow from this bounding element, those with the nearest system exit time first, then it is always known which particular load units leave this bounding element and therefore arrive at the internal stack. (2) If the "location before" the internal stack is the initial stack size, then each load unit in this initial lot is known individually and therefore its system exit time.

So, finding the load units matching up to any flow determined by the load unit sub-model simply is done by tracing back each individual load unit to its source (either a bounding element or an initial stack size); after that all load units with the smallest system exit time leave each successive location first, thus allowing to trace each individual load unit through the load unit sub-model.

Conversion for machines
Converting flows of machines into assignments of transport jobs to actual, individual machines is more complicated than the method just discussed for load units. This effect is primarily a consequence of the difference in unit for both types of flows: While in the load unit sub-model each unit of flow corresponds to exactly one individual load unit, in the equipment sub-model each unit of flow corresponds to the average number of machines of a pool that is used during a particular time window for transport over a particular
equipment routing. By consequence there is no way to convert flows in the equipment sub-model directly into transport jobs for actual machines.

Instead of direct conversion a more "fuzzy" conversion process is used for equipment flows. The basic principle for this "fuzzy" conversion process is that whenever a machine finishes a transport job during simulation, it checks whether there is still another job to do based on the results of the last equipment optimisation round. In this way transport jobs are more or less executed in a simple FIFO-order than in a kind of planned time-table order. This process will be given some more attention in the remainder of this section.

Following types of data are used for the "fuzzy" conversion of flows of machines into actual transport jobs:

1. the numbers of load units that are available at each location for transport during the currently simulated time window. These numbers are references to actual, individual load units that satisfy two conditions: (1) They ought to be selected for transport over a particular arc, stemming from that location, during the currently simulated time window. This pre-selection is made according to the rules for conversion of flows of load units as just discussed. (2) They ought to be actually present at that location at the current simulation time.

2. the numbers of transport jobs that each equipment pool still is planned to perform on each section that it may serve, separate per relocation-origin, during the currently simulated time window.

These types of data will be explained in more detail.

The numbers of load units available for transport are stored at each single location in the system in lists as shown in figure 7.106, separated per equipment routing. These lists are called the "available-lists". These load units are all actually present at that location, ready for further transport over that equipment routing.

Next to these lists of available load units, some other lists exist. These lists define which load units still ought to arrive at that location over that particular equipment routing during the currently simulated time window, according to the converted results of the load unit sub-model. But these load units are not present yet at that location. These lists are called the "ought-to-lists".

---

56 If desired, some more complex final assignment method could have been selected. Such a method could use a second optimisation model to convert the results of the first optimisation model (i.e. equipment hopping) into more detailed, actual assignments of equipment to transport jobs. In that case equipment hopping is merely used as a pre-planning tool that provides a rough basis for further, more detailed optimisation and assignment.
The interrelation between both lists is shown in figure 7.107. It shows a flow chart that is executed whenever a load unit arrives at a particular location during simulation. As such a load unit is always transported over a particular equipment routing, it can be found quickly over which arc the load unit is to "leave" the location according to the converted results of the load unit sub-model. This could either be (1) an inter-window transport arc (if present), (2) an intra-window transport arc, or (3) a stacking arc (if present).

After the right arc is found, the load unit is removed from the "ought-to-transport" list and is added to the "available-for-transport" list if it does not concern a stacking arc. In this way individual load units are registered as present for internal transport. Finally, all equipment pools that may be used for transport over the selected arc are warned that a new transport job has arrived.

The numbers of transport jobs that each equipment pool still is to perform during the currently simulated time window is the second type of data that controls "fuzzy" conversion of flows of machines into actual transport jobs. These numbers of transport jobs are based on following parameters, that result from equipment optimisation:

1. number of load units to transport during a time window \( w \) by equipment pool \( p \) over section \( s \); this parameter equals \( \text{lu}_\text{tn}_w,p,s \)
2. number of trips to be made during time window \( w \) by equipment pool \( p \), related to transport over section \( s \) and relocating from location \( o \); this parameter equals \( \text{w}_\text{dt}_w\text{epa}_n,w,p,s,o \).

These parameters are used as follows. As soon as a machine is ready to start working (either by finishing a previous transport job or by starting at the beginning of its working hours) it checks for the need to transport any load unit. This check is made by polling all values of \( \text{w}_\text{dt}_w\text{epa}_n,w,p,s,o \) for the currently simulated time window \( w \) and the machine's own pool \( p \). The machine itself is located at location \( o \) at the time it starts checking. Whenever the value of \( \text{lu}_\text{tn}_w,p,s \) exceeds zero\(^{59}\) then the transport can potentially be accepted by the machine. If the transport can be served within valid working hours (as will be discussed on short notice) then finally the machine chooses that particular transport that is related to the load unit with the smallest system exit time that is registered as available for transport over section \( s \) at the location at the start of that section. Note that in some cases the value of \( \text{lu}_\text{tn}_w,p,s \) may exceed zero but nevertheless may not result in a valid transport assignment as no load unit has been registered as available at the location at the start of section \( s \), at the current simulation time. In case no load unit is found at all for transport, the machine waits until it gets signalled that potentially a new transport jobs is available.

After a load unit has been selected for transport by a machine a number of actions is to be performed. First of all, the load unit is to be reserved for transport by that particular machine which can be considered as the actual assignment action. Second, the load unit is removed from the "available-for-transport" list at the location at the start of section \( s \)\(^{59}\). Third, the values of \( \text{lu}_\text{tn}_n,p,s \) and of \( \text{w}_\text{dt}_w\text{epa}_n,w,p,s,o \) lowered: \( \text{lu}_\text{tn}_w,p,s \) by the number of load units transported (which may be greater than one for MLU-equipment as will be discussed shortly), and \( \text{w}_\text{dt}_w\text{epa}_n,w,p,s,o \) by one for all types of equipment.

\* Conversion for machines - working hours
Some details need some further discussion: MLU-equipment, trailer-type equipment and working hours. This latter aspect is tackled first. Working hours of equipment pools are included in two ways during simulation: (1) To trigger all machines of a pool to start looking for a transport job, and (2) to prevent a machine from selecting a transport job. The first use is controlled by a pool manager that first sets a flag named "In Working Hours" to

\footnotesize{\textsuperscript{57}}{e} at which arc it is registered at the "ought-to-transport" list.

\footnotesize{\textsuperscript{58}}This additional check is needed to correct for rounding-errors that occur in case of MLU-equipment as will be discussed shortly.

\footnotesize{\textsuperscript{59}}The load unit has been removed previously from the "ought-to-transport" list at that location.

---

page 160
TRUE for all machine in that pool and then re-activates them all, making them start looking for a transport job in the way just described. The second use is controlled by the machine itself, as it refuses to select a transport job that ends after its current working-hour period ends. In this contemplation the full transport time \( t_{p,s,o} \) including time needed for relocation from \( o \) to \( s \). This is the only way in which the aspect of equipment relocation is included in the assignment procedure.

**Conversion for machines - MLU equipment**

**MLU-equipment** introduces the problem that the number of load units transported (1) may exceed one and (2) may vary from trip to trip. It has been chosen to solve this problem in a rather simple way, based on the average degree of occupancy of an MLU-pool. This average degree of occupancy is defined as:

\[
\text{average occupancy} = \frac{lu_{w,p,s}}{et_{ce} \cdot lu_{dlutn}_{w,p,s}}
\]  
(7.39)

For each individual transport that is to be made by MLU-equipment the number of load units to be transported is then defined by:

\[
\text{load units to transport} = \text{Trunc} \left( \frac{lu_{w,p,s}}{et_{ce} \cdot lu_{dlutn}_{w,p,s}} + \text{Random}[1] \right),
\]

(7.40)

in which:

\[\text{Trunc} = \text{truncating function}\]

\[\text{Random}[1] = \text{random number in the range } [0, 1).\]

In this way a small variation around the average degree of occupancy is obtained. As it is not guaranteed in this way that the total number of actually transported load units equals \( lu_{w,p,s} \) at all times the additional small deviations of the planned numbers of load units to transport compared to the eventually transported numbers of load units may occur.

**Conversion for machines - pulling equipment**

**Pulling equipment** is somewhat more complicated than normal self-propelled equipment as it must be assigned simultaneously to flows of load units and to flows of trailers. However, this latter, additional assignment is handled in the same way as the assignment of load units was handled: After all potentially acceptable flows of load units have been established, the pulling machine searches for a matching trailer to transport these load units with. If such a trailer is found, then that trailer is marked as reserved by the pulling machine and the number of trailers that was planned to be pulled during the currently simulated time window is decreased. In case of simplified double relocation this planned number is stored in \( \text{epa}_{n,w,p,pull,s,pull,p,trail} \) and in case of non-simplified double relocation it is stored in \( \text{epa}_{n,w,p,pull,s,od,p,trail} \).

The pool to search a trailer from is known on beforehand as each unique combination of trailer and pulling pool is discerned individually in the load unit sub-model by the fictitious equipment pool \( p_{pool, trail} \), which results in a flow of load units in the load unit sub-model named \( lu_{th_{w,p,pool, trail}} \) as previously discussed.
8.1. Introduction

Stacking of load units at terminal complexes is an integral part of internal transport as discussed at Chapter 7: It is used to decouple two subsequent movements of internal transport. Refer to figure 8.1 for an example hereof, in which a load unit is transported between two locations and is stored in between at a stack. The horizontal axis of figure 8.1 shows the position of that particular load unit in the modelled terminal complex, while the vertical axis represents the time that elapses during internal transport and stacking. Clearly, the time that is consumed by stacking is substantially greater than the time consumed by internal transport. This behaviour is characteristic for stacking.

Stacking is encountered at a number of locations in real-life terminal systems and in the model described in this thesis. It is found in total at three different types of locations, as can be deduced from Chapters 5 and 7:
- bounding elements;
- train handle areas;
- internal stacks.

The presence of stacking at the latter type of location is quite obvious; this in contrast to the first two types of locations at which special kinds of stacking are modelled. All three types of locations will be described in detail in this chapter.

**Differences per level of detail**

Stacking at UCM level of detail is approached different from stacking at SDM and DDM level of detail in two ways: (1) The locations at which load units can be stacked as just shown and (2) the way stacking times are determined.

At UCM level of detail stacking is allowed at internal stacks only and therefore is not available for bounding elements and train handle areas. Stacking times at this level of detail are defined as fixed stacking times that depend on the type of load units that is stacked only; there is no distinction between the various closed tasks that run over the internal stack with respect to the stacking time. As Chapter 5 provides a sufficiently detailed description of stacking at UCM level of detail, this level is not further discussed in this chapter.

At SDM and DDM level of detail stacking is allowed at all three types of locations just mentioned, each one having its own characteristics. Stacking times are not defined as fixed times but are determined dynamically during simulation and optimisation based on so-called flow patterns and equipment routings. Refer to Chapter 5 for a description of this principle and to Chapter 7 for more details on flow patterns and equipment routings.

This chapter discusses stacking at bounding elements, train handle areas and internal stacks at SDM and DDM level of detail only.

---

1Stacking at bounding elements and train handle areas is only available at SDM and DDM level of detail.
8.2. Local stacking and normal stacking

In Chapter 7 normal internal transport was discerned from local internal transport. The main difference was that local internal transport was used to model transport within locations as defined at UCM level of detail, while normal internal transport was used to model transport between these locations.

An analogous subdivision is made for stacking: Normal stacking is used to refer to stacking at those locations that allow stacking at all three levels of detail (which are internal stacks) while local stacking is used to refer to stacking at those location that did not allow stacking until SDM level of detail (which are bounding elements and train handle areas). By consequence, local stacking is only defined from SDM level of detail on.

Following three sections describe the way in which (local) stacking is modelled at all three types of locations at SDM and DDM level of detail that allow stacking: Bounding elements, train handle areas and internal stacks successively.

8.2.1. Local stacking at bounding elements

Stacking at bounding elements at SDM and DDM level of detail is optional for each individual location and can be used to model marine stacks or barge stacks at a coarse level of detail. In that case the stack is considered as an indivisible unit only, not divided into separate blocks as internal stacks are, but only discerning between the various equipment routings that run over the bounding element.\(^2\)

If local stacking is used, then the stacking capacity either can be limited to a fixed number of generalised load units or can be specified to be unlimited. This latter option may be needed to ensure that the limited stacking condition as described in Chapter 7 is satisfied.

The pools that can be used for stacking and unstacking load units at a bounding element’s local stack are selected according to the rules previously discussed at Chapter 7 which will be repeated shortly by referring to figure 8.2.

Two different transhipment movements can be discerned related to local stacks: One transhipment movement at the "outside" of the modelled system (which is at the side of the system’s boundary, or sea quay in case of a marine terminal) and transhipment movement one at the “inside” of the modelled system. These two movements are labelled \(\square\) and \(\circ\) respectively in figure 8.2. It is purely coincidental that transhipment movement \(\circ\) represents transhipment into the stack and that transhipment movements \(\square\) represents transhipment out of the stack in this case; if figure 8.2 would have depicted an export flow instead of an import flow then the direction of all horizontal- and transhipment arcs would have been inverted. In that case transhipment movement \(\square\) would have represented transhipment out of the local stack.

Movement \(\square\) can be done by any equipment pool installed at the bounding element’s private network that is able to tranship and to move. Movement \(\circ\) can be done by the same pools used for movement \(\square\) and by any

\[^{2}\text{If a more detailed analysis of such a stack is required then it should be modelled as a separate internal stack instead of a local stack at the bounding element.}\]

\[^{3}\text{or an expanded private network, which is equal to closed internal transport.}\]
equipment pool that is installed at the equipment networks connected to the bounding element and is available for transport over a particular equipment routing.

Analysis after simulation offers an insight in the time behaviour of the local stack size at the bounding element, if required split per closed task or equipment routing.

8.2.2. Local stacking at train handle areas

Stacking at train handle areas at SDM and DDM level of detail is optional for each individual location and is used to model stacking at handle bundles, next to the trains. In contrast to bounding elements, stacking capacity at train handle areas has a disperse character as it is divided over all handle bundles that are defined for a certain train handle area. This disperse character also shows from the way in which the stacking capacity is defined: It is defined as a fixed capacity per handle bundle length unit. In this way equally long handle bundles automatically are assigned an equally great stacking capacity. This will be illustrated by an example.

Consider table 8.1 in which all bundles related to a train handle area are shown. If the maximum stacking capacity for all handle bundles related to that train handle area is estimated to be 15 load units per 100 meter bundle length, then this automatically yields the stacking capacities for all its three handle bundles; no stacking capacity is assigned to the single yard bundle.

<table>
<thead>
<tr>
<th>Bundle</th>
<th>Length [m]</th>
<th>Stacking capacity [load units]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yard bundle</td>
<td>450</td>
<td>--</td>
</tr>
<tr>
<td>Handle bundle I</td>
<td>600</td>
<td>90</td>
</tr>
<tr>
<td>Handle bundle II</td>
<td>750</td>
<td>113</td>
</tr>
<tr>
<td>Handle bundle III</td>
<td>550</td>
<td>83</td>
</tr>
</tbody>
</table>

Table 8.1: Example of stacking capacity per handle bundle in case of 15 load units per 100 meter.

If required, the stacking capacity can be set to unlimited for all handle bundles related to a particular train handle area. This latter option may be needed to ensure that the limited stacking condition as described in Chapter 7 is satisfied.

From figure 8.3 it can be seen that the location of stacking at train handle areas, relative to the system's boundary, is different from the location that was previously found in figure 8.2 for bounding elements: Load units are stored in a local stack of a train handle area (if required) immediately after they have entered the modelled system. This in contrast to local stacking at bounding elements at which load units were stored in a local stack after they had been transported by local internal transport from the sea-quay to that local stack.

![Figure 8.3: Transhipment movements into and out of a train handle area's local stack, marked as ① and ②.](image)

One simplification of the model is that, whenever local stacking is selected for a train handle area, it is applied for all load units that ever enter or leave that train handle area. The effect that in reality some load units are stored next to the train before further transport into the system while others are transported directly after unloading from the train into the remaining system is therefore ignored. Future expansions of the model may mend this drawback, either by an improved version of equipment hopping or by improved discrete simulation.

---

4. or an expanded private network, which is equal to closed internal transport.

5. That is, if both local stacking and local internal transport are used for that bounding element, corresponding to situation A of figure 8.2.
Chapter 8. Stacks

Movement ① is done by equipment pools that are installed at the handle private networks of the train handle area. As discussed previously in Chapter 7 these handle private networks are automatically defined at all rail bundles at which load units can be transhipped, i.e. at which trains can be handled. As to every such rail bundle a different handle private network is assigned it is possible to install any combination of equipment pools at the defined rail bundles. This is clarified by referring to figure 8.4 in which two handle bundles, and by consequence two handle private networks, are defined. At each of these two handle private networks different equipment pools for transhipment from and onto trains can be installed, if desired.7

One special case for transhipment ① is modelled by installing a non-movable transhipment pool at such a handle private network: Such a pool is considered as a pool of gantry cranes over the trains that must always be used for transhipment movement ①, regardless of the presence of other equipment pools at the handle private network. It is allowed to install at most one such non-movable transhipment pool at a handle private network.

Movement ② is done by the same pools used for movement ①, so by all pools that are installed at the handle private networks. In case gantry cranes have installed for movement ① according to the method just described these crane are also available for transhipment movement ②, along with all other transhipment pools installed at that handle private network.

† Simplifications during equipment optimisation
During optimisation of the use of equipment according to equipment hopping some problems may arise for transhipment movements ① and ② which are caused by the diffuse installation of equipment for these movements: At each handle bundle a different set of equipment pools can be installed. Each of these equipment pools can only be used at the handle private network of that particular bundle.

However, during equipment optimisation it is not known at which precise rail bundle each train is to be handled, and therefore it is not known which precise private handle bundle and associated equipment pool can be used for transhipment from and onto the train. This uncertainty is a consequence of the dynamic selection process of rail bundles as described in Chapter 9, in which numerous factors may play a role. By consequence there is no possibility to create different arcs in the equipment- or load unit sub-model for the various handle bundles at a train handle area.

This problem is solved by grouping all equipment pools that are installed at the various handle bundles of a train handle area according to following rules:
1. for each train handle area a new, fictitious equipment pool is created for each equipment type that is installed at one or more handle bundles of that train handle area. This new, fictitious pool is then used for the equipment optimisation model as a substitute for all actual equipment pools of that type installed at the train handle area;
2. for this new, fictitious equipment pool the capacity is set to the sum of all equipment pools installed at the train handle area that represents. If any of these equipment pools has an unlimited capacity then the new, fictitious pool also receives an unlimited capacity.

6Local stacking at train handle areas only refers to equipment that is installed at handle private networks only; equipment installed at the train handle area's main private network is not used for local stacking.
7One possible use hereof is to install different pools of the same equipment type on different handle bundles; each of these pool may have a different capacity, if required.
Obviously, this simplification introduces some deviation between the transhipment movements predicted by the equipment optimisation model and the actual transhipment movements simulated during discrete simulation, which can not be avoided by the current structure of the model.

Figure 8.3 only shows transhipment for load units that are imported into the modelled system at a train handle area, as figure 8.2 did for bounding elements. For load units that are exported out of the modelled system all horizontal- and transhipment arcs are inverted.

- **Need for a separate stack**

In those cases that local stacking is not desired for a train handle area it may be decided to create a separate "private" stack that is strongly related to that train handle area. This can be achieved by inserting an internal stack in one of the physical paths that originate from the train handle area as depicted in figure 8.5 and by prohibiting local stacking.

From an operational point of view such a solution has both advantages and disadvantages, depending on the type of equipment that is used for movements ⊙ and ⊙ of figure 8.3: On one hand a "private" stack would eliminate transhipment movement ⊙ but on the other hand it would create the problem to match transhipment movement ⊙ with subsequent internal transport to the "private" stack. In those cases that movable equipment like reach stackers can be used both for transhipment movement ⊙ and for subsequent internal transport then this latter disadvantage would be eliminated. In those cases that non-movable equipment like gantry cranes are installed at the private handle bundles then this matching problem would be present.

**8.2.3. Stacking at internal stacks**

Internal stacks are generally used to model stacking at SDM and DDM level of detail in a more detailed, different way than at both previously discussed types of locations. But if required, these internal stacks can also be used to model stacking in the same coarse way as found at bounding elements by selecting a special stacking mode.

- **Stacking modes**

Two modes are available to model stacking of load units at internal stacks:

- Detailed stacking and simplified stacking. The main difference between these two modes is whether or not load units are stacked in separate so-called stack blocks. These stack blocks can be considered as small, partial stacks to which priority rules can be attached. In case of simplified stacking these stack blocks are not discerned and therefore no precise information on the position of load units in the stack is available during simulation when simplified stacking is used. In this way simulation is speeded up compared to simulation of detailed stacking.

One correspondence between simplified and detailed stacking is that the capacity of an internal stack is always considered to be unlimited. This in contrast to the local stacking capacity of bounding elements and train handle areas which can be set to a limited value if required. In this way it is always guaranteed that any closed task that runs over an internal stack runs over a location with unlimited stacking capacity -- which is a requirement for the equipment optimisation model described in Chapter 7.

---

8 For both SDM and DDM level of detail.
Detailed stacking
In case of detailed stacking each stack is divided into any number of stack blocks during simulation. These stack blocks can be used to manipulate the precise location within an internal stack at which each individual load unit is to be placed, based on following parameters of the load unit:

- client;
- design class like container or swap body;
- special class like empty and reefer;
- closed task;
- source-bounding element;
- destination-bounding element.

For all possible values of all these parameters a valuation from the REMBRANDT scale can be specified. Based on the score of each individual load unit on all of the just mentioned parameters the best stack block for that particular load unit is selected by using multiplicative AHP. An example will clarify this.

Consider the situation depicted in figure 8.6 in which three different stack blocks are discerned for a particular internal stack. Suppose that the only parameter of interest for stack block selection (i.e. criterion) is the bounding element from which the load units to stack originates. In that case to each stack block a different valuation for the use of that stack block for a load unit originating from a particular bounding element can be assigned as shown in figure 8.6.

If, for instance, a load unit originating from bounding element "Miersk" is to be stacked then the leftmost stack block will be selected as this block results in the highest score of all. Note that this example is quite simple as it is based on one selection parameter only. In reality there may be a need to combine various parameters (criteria) to obtain the best block to stack a particular load unit in. In that case multiplicative AHP as described in Appendix A becomes of importance as it includes and combines any number of parameters and guarantees a natural pay-off behaviour in case of conflicting criteria.

Simplified stacking
In case of simplified stacking no distinction between various stack blocks is made. Instead, the entire stack is considered as one indivisible unit in which only different equipment routings are discerned. In this way a similar stacking behaviour as found at bounding elements is obtained which can be useful if (1) no detailed stacking is needed for analysis or (2) using detailed stacking would yield unacceptable delays in simulation.

---

9 In a more precise way: In case of detailed stacking the subdivision of an internal stacks as modelled is maintained during simulation, while in case of simplified stacking this subdivision is simply ignored.
10 Refer to Appendix A for all details on the REMBRANDT scale and on multiplicative AHP.
11 Although not used during simulation, these stack blocks may still be defined when selecting simplified stacking which is convenient for two reasons: (1) Fast switching between simplified and detailed stacking without re-definition of stack blocks, and (2) improved visual appearance of the stack on-screen.
Equipment pools
From figure 8.7 it can be seen that four transhipment movements are involved with stacking at internal stacks: Two transhipment movements for transport into the stack and two other movements for transport out of the stack. In those cases that load units are transported to the stack by the same equipment pool that is used for actual stacks (which can be the case when using straddle carriers that have a wide service range) then transhipment movements ① and ④ are do not occur.

<table>
<thead>
<tr>
<th>Local stacking = Yes</th>
<th>Closed internal transport</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>①</td>
</tr>
<tr>
<td></td>
<td>A.</td>
</tr>
<tr>
<td></td>
<td>local transport</td>
</tr>
<tr>
<td></td>
<td>B.</td>
</tr>
<tr>
<td></td>
<td>②</td>
</tr>
<tr>
<td></td>
<td>③</td>
</tr>
<tr>
<td></td>
<td>④</td>
</tr>
<tr>
<td>No</td>
<td>①</td>
</tr>
<tr>
<td></td>
<td>C.</td>
</tr>
<tr>
<td></td>
<td>local transport</td>
</tr>
<tr>
<td></td>
<td>D.</td>
</tr>
</tbody>
</table>

Figure 8.7: Transhipment movements into and out of an internal stack.

For transhipment movements ① and ④ all transhipment pools that are installed either at the internal stack's private network or at some connecting network can be used; except the fixed gantries that can be installed at the stack's private network. These fixed gantries are specified, analogous to those installed at train handle areas, as non-movable transhipment pools\(^\text{12}\), and can only be used for transhipment movements ① and ④ which are the actual stacking movements by which the load units are placed in the stack. If no gantry cranes are installed at the internal stack's private network then any other transhipment pool can be used for movements ① and ④.

Transport distances
Depending on the stacking mode one out of two methods to determine the driving distance for stacking equipment is used:

In case of simplified stacking a random distribution is used to specify the transport distance within the stack. This random distribution is used both during actual simulation and for equipment optimisation: During simulation, each time a load unit is to be transported into or out of the internal stack the transport distance for that particular load unit is sampled from the random distribution and used as such. For equipment optimisation the average value of the random distribution is used as an estimate of the transport distance\(^\text{13}\).

In case of detailed stacking geographic information on the stack blocks is used to determine the transport distance. During simulation it is known for each single load unit to or from which stack block it should be transported, and therefore the transport distance is determined as the rectangular distance from the gate of the internal stack to the most remote corner of the selected stack block. For equipment optimisation the average rectangular distance between the gate and all stack blocks is used as an estimate of the transport distance.

Simplifications of the model
At the moment this thesis is written one important aspect of stacks, namely reshuffling and "digging up" load units, is not included in the model yet. Several options to model this aspect could be thought of, but none of them is elaborated sufficiently to be described.

8.3. Initial stack sizes

At start-up of simulation the modelled system already contains a number of load units. These load units have been imported into the modelled system before the simulation started, either by load schedules or by contra schedules that have an arrival time primary to this starting time, and play an important role to get the simulation model quickly in a state of equilibrium. This induces the need to find a reasonable estimate of the location at

\(^{12}\) At most one such pool of fixed gantry cranes can be installed at any internal stack.

\(^{13}\) A future improvement could be to change this average value into the average plus a fixed number times the deviation of the random distribution. In this way a more pessimistic estimation of the transport distance is obtained that models a conservative attitude towards unexpected delays.
which these load units are stacked at simulation start-up which automatically results in the initial stack sizes of all locations. This section describes the way these initial stack sizes are determined.

Four factors play a role in finding the initial stack sizes:

- stacking capacities of all stacking locations;
- preferences for presence according to flow patterns;
- routing transport times according to equipment routings;
- system exit times of all load units present at simulation start-up.

These parameters are not used for all load units in the same way, as the initial stack sizes are determined in two different "rounds": (1) In the first round all urgent load units are assigned to stacking locations as near to the system's boundary as possible to guarantee a smooth and fast system exit. These urgent load units are defined as those load units for which the difference between system exit time and the start-up simulation time is less than their equipment routing’s routing transport time. During this first round all maximum stacking capacities of the stacking locations are respected, but all preferences for presence of flow patterns are ignored. (2) In the second round all non-urgent load units are assigned to stacking locations, taking into account both the maximum stacking capacities and the preferences for presence of flow patterns. During this second round a small optimisation model is used in which two conflicting aspects of the assignment are balanced. This optimisation model, based on linear programming\(^{14}\), reads:

\[
\begin{align}
\text{Max} & \sum_{r,l} (x_{r,l} \cdot f_{p \_ P_{r,l}}) \\
\text{subject to:} & \\
\sum_{l} x_{r,l} &= n_r \\
\sum_{r} x_{r,l} &\leq n \_ max_l
\end{align}
\]

in which:

- \(x_{r,l}\) = decision variable: number of load units running over routing \(r\) that is assigned in the second round be initially stacked at location \(l\).
- \(n_r\) = total number of load units running over routing \(r\) that is to be assigned in the second round to an initial stack.
- \(n \_ max_l\) = maximum stack capacity that is still available in the second round at location \(l\). This stack capacity equals the \textit{original} stack capacity minus the number of load units that already has been assigned to location \(l\) in the first round. In case location \(l\) has an unlimited stack capacity this parameter is set to the total number of load units that is to be assigned in the second round.
- \(f_{p \_ P_{r,l}}\) = preference for presence of load units running over routing \(r\) at location \(l\) as specified at the flow patterns.

Conversion of these anonymous results of the second round into actual assignments of individual load units to stacking locations is done by assigning those load units with the smallest system exit time first to the stacking locations nearest to the system's boundary for system exit until \(x_{r,l}\) has been fully consumed for that nearest location and the one but nearest location is selected.

In case of internal stacks at which detailed stacking is used for each of the load units that is assigned either in the first round or in the second round its proper stack block is selected according to the priority rules of those blocks.

Effect of unlimited stacking capacities
In one particular case all initial load units of the second round, running over some equipment routing can be assigned to the same stacking location. This is the case when (1) that stacking location has an unlimited stacking capacity and (2) the preference for that location is highest of all locations over the equipment routing runs. Although possibly unexpected, this is not concerned an error as it is in correspondence with the formulated problem.

\(^{14}\) Refer to Appendix A.
Chapter 9. Railway aspects

9.1. Introduction

An imaginary visit back to the twenties may result in a "romantic" scenery; a scenery in which steam locs are driving with a dark scent of burnt coals, every now and then blowing their steam whistles. Cattle would be driven out of the wagons by men dressed in rags, yelling and screaming to get the beasts move. Boxes of Ceylon tea and bales of cotton would be carried by the same men, exhausting themselves and overloading their bodies.

Nowadays things run in a different way. As soon as a brightly polished 4.3 MW electrical locomotive enters the station pulling a trail of wagons carrying neat, rectangular containers and swap bodies, the semi-automatic gantry cranes move into position, ready to unload and load the containers and swap bodies, several tens of tons each.

To make all operations run smoothly at a railway station a lot of planning is required. Unlike the previously discussed operations for internal transport (refer to Chapter 7), the operations at a railway station are rigid, returning in the same pattern every day. This is a consequence of the fixed time tables of trains that arrive at a railway station; only the number of containers to load or unload varies between days. Therefore, all operations of trains arriving at rail service centres (which can be considered as railway stations for containers and swap bodies) are modelled in a more rigid way than the operations of equipment for internal transport were: While the latter type of operations was modelled by means of a general structure only, the fist type will include detailed information on precise, mandatory timing of trains and on sequences of operations.

♦ Differences per level of detail

All operations that can be modelled in the terminal design aiding tool for trains being handled at rail service centres are described in this chapter. Important, and always returning in this operational flow, is the infrastructure that is used by trains: The rail tracks and points over which trains run highly restrict the freedom in planning operations of trains. This effect is found when analysing operations of trains and locomotives at DDM level of detail, when detailed routing of locomotives and block safety systems are included in the simulation. At SDM level of detail this effect is less obvious because of some simplifications introduced at that level, needed to advise on the lay-out of the rail service centre at DDM level of detail, and to get a better impression of the actually desired rail infrastructure capacity, without any delays introduced by a block safety system.

At UCM level of detail trains are modelled in a rough way by means of load schedules only\(^1\), without the need to specify details on rail infrastructure. Therefore, none of the elements described in this chapter is used at UCM level of detail.

♦ What is included?

This chapter covers a number of aspects that is found at rail service centres; it covers different types of trains, various operational options for trains, and data on the numbers of load units to be handled, along with their timing. This chapter also covers some discussion of various types of rail points and their impact on the modelled system, and of various types of rail bundles. This chapter does not include emergency protocols in case of a locomotive break-down or a damaged rail wagon; although essential to include in the eventual, actual terminal design, it is left out from the model for the moment, as it is too complicated yet.

9.2. Standard operations at rail service centres

Operations at rail service centres are somewhat more complicated than just loading and unloading trains after they have been rolled under the gantry cranes. As a matter of fact, a great number of additional operations

\(^1\)Refer to Chapter 6 for more details.
complicate processes at rail service centres. One of the most complicating of these additional operations is the use of light diesel engines\(^2\) \((D\text{-}locs)\) for local transport of trains within the range of the rail service centre; this aspect increases complexity of the processes at rail service centres and has a major impact on the required infrastructure.

This section will give a brief overview of this and other additional operations that are found at rail service centres. All of these will return in a later stage in this chapter, when discussing the so-called train concepts, train services and block safety systems.

### 9.2.1. Handling

One of the most straightforward operations at rail service centres is handling, which is simply defined as loading and unloading of standard load units like containers and swap bodies. Handling can occur only at a limited number of locations at the rail service centre: it can only occur at so-called handle bundles which are defined as rail bundles at which handling is allowed. This will be explained.

A rail bundle is defined in this thesis as a (possibly still unknown) number of rail tracks that are functionally related and need to be considered as a unity. Although possibly cryptic, this definition leaves some freedom to the user of the terminal design aiding tool to develop his own view on the definition of rail bundles, freedom which in some cases is needed to increase understanding of the modelled situation. One important aspect to bear in mind when dividing tracks into bundles is, that in the model described here to each bundle a number of allowed operations can be assigned; these are operations that may be performed at any track of that rail bundle. In some cases, in practice, a bundle is defined as a set of rail tracks that can be reached over one commonly shared rail point. In practice it frequently happens that some tracks inside that rail bundle may have a different usage than other tracks inside that same bundle. In this model, however, such a bundle should be split into two or more separate bundles, each bundle having a different combination of allowed operations. This will be discussed later on in more detail.

An example of rail bundles is given in figure 9.1, showing four different bundles that are connected by means of rail tracks. Two of these bundles are shown with gantry cranes placed over them, while the other two are not. These first two bundles, therefore, have been marked as bundles at which trains can be loaded or unloaded (i.e. handled). For that reason these bundles are called handle bundles. Later on a more detailed description of these handle bundles will be given. Figure 9.1 also shows a special rail track that is marked as "system entry/exit". This rail track is connected to the remaining rail network, that lies outside the scope of the model described in this thesis and will turn out to have some special importance for the so-called train services.

For the moment it is only needed to know that several special settings make a rail bundle suitable for transhipment of load units, either loading or unloading, or for some other operation.

### 9.2.2. Local movements

In its most simple form all operations of trains at rail service centres are as shown in figure 9.2: A train enters the modelled train handle area, stops at a handle bundle, is unloaded and loaded by the gantry cranes while its locomotive remains coupled, and leaves the modelled area again with that very locomotive. In this case the box labelled "Internal processes", which represents all processes of the train except system entry and system exit, simply consists of unloading

---

\(^2\) Uncoupled locomotives.
and loading the train -- and no other action whatsoever. In practice, however, a greater number of operations can be found in the box "Internal processes". One (very important) group of these other operations are the local movements of trains. These local movements are discussed in a general way in this section; they will be discussed in more detail at the train concepts and the train services.

Local movements are defined as all movements of (parts of) trains inside the modelled area of any rail service centre except the movements of system entry and of system exit. Almost any train that is handled at any modelled rail service centre causes one or more local movement; only in the special case described in the previous paragraph no local movements are generated.

In the model local movements of trains can only run between two tracks at rail bundles; they can not run between tracks that are not part of a rail bundle. At this point there is some difference between tracks at rail bundles at SDM and at DDM level of detail: Bundles at SDM level of detail consist of an unknown number of tracks, and therefore it is not possible to refer to a particular track of a bundle. By consequence, at SDM level of detail a local movement runs from one rail bundle to the other, without any reference to any particular rail track within that bundle. By using this special approach the model is able to advise after simulation on the number of tracks to install at any rail bundle in the system. When these bundles are modelled in more detail at DDM level of detail as separate tracks, a more detailed simulation and analysis of rail movements is possible. By consequence, local movements of trains at DDM level of detail do run between two individual tracks -- which still must be defined as part of a rail bundle to be a valid reference.

A schematic representation of a local movement of a train is given in figure 9.3, in which a train is relocated from a handle bundle to a yard bundle. The handle bundles can be recognised by the gantry cranes that have been placed over them; the yard bundles are not fit for transhipment of load units.

♦ Train relocation

One of the most important types of local movements is relocation of trains between unloading and loading. When relocation is used, then the train is moved out of the handle bundle at which it has been unloaded to some other, non-handle bundle where it must reside until it is about to be loaded. When the train is about to be loaded, it is moved into the same (or possibly another) handle bundle again, where it is loaded.

This type of operation is commonly accepted to reduce fixed costs of the rail service centre, as handle bundles are generally believed to be more expensive than non-handle bundles. In Germany this type of operation is called Fließverfahren as the trains "flow" in and out of the handle bundles. Refer to figure 9.4 for an example hereof.

The inverse of Fließverfahren, i.e. of trains that use relocation between unloading and loading, is called Bedienverfahren in Germany. In case of Bedienverfahren the trains remain standing at the handle bundle where they are unloaded and loaded concurrently, during the whole handling process, and therefore correspond to the operational scheme of figure 9.2. This concept is mainly used for inland terminals to serve road trucks as quickly as possible by transhipping directly from the trains onto trucks instead of from some stack. The main reason to use this concept is to satisfy the truck drivers as much as possible.

---

3But even at non-handle bundles a full block safety system is to be installed, which is expensive too.
9.2.3. **Coupling**

In some cases all local movements are done by the same locomotive that rolled the train into the modelled system; in those cases that same locomotive, which in most cases is an electrical locomotive (E-loc), is likely to remain coupled all time to the train. In all other cases some or all local movements are done by diesel-powered locomotives (D-locs). It is important to discern between these two types as they allow detailed system analysis after simulation. The primary assumption in the model is namely that all trains that enter or leave the system are pulled by an E-loc, while all locomotives that are installed at the rail service centre, and therefore are used for local movements only, are D-locs.

Coupling and uncoupling of locomotives may occur at a number of places in the total process of a train inside the rail service centre's area: It may occur, for instance, directly after system entry, or between unloading and loading of the train, or after the train has been split into several individual parts; all these special events at which coupling or uncoupling is required will be discussed in detail at the train concepts and the train services. Coupling and uncoupling in the model is, however, *always* executed at rail bundles and never at rail tracks that are not part of a rail bundle. In this way the importance of bundles in the model is increased, facilitating system analysis.

In most cases D-locs are used for local movements because of the electrical wires (catenary) needed for E-locs: These electrical wires are in the way for transhipment by (conventional) gantry cranes, and therefore E-locs can not be used to position trains on their handle tracks. Refer to figure 9.5 for an example of this clash.

**Figure 9.5: Electrical wires blocking transhipment.**

In the model described in this thesis the times needed for coupling and for uncoupling are modelled as a constant time that is consumed by those actions. If required, these two times may be different for D- and for E-locs.

9.2.4. **Brake tests**

Safety is an important issue in rail transport. Especially for trains running over the public rail network safety standards are very high. One of the results of this high demand on safety is the so-called stationary brake test (or simply *break test*) that is compulsory for almost any train. This brake test is used to test whether all brakes of the train are functioning properly by enabling them all and checking (visually, by men walking along the train) whether all brakes are in function. This brake test takes much time for two reasons: (1) The men inspecting the train are to walk along the entire train in both directions to check all brakes, which can be a distance of up to 1.4 kilometres. (2) The brakes work on air pressure that is stored in tanks under the wagons. Unfortunately, these tanks in general leak somewhat so that they are quite empty after a train has been handled at a rail service centre, without the constant presence of an air pumping locomotive. Filling these tanks again takes some time, too. At some rail terminals this time is shortened by installing compressed-air tanks at the terminal yard that allow a quick refill of the tanks.

Two different brake tests are allowed nowadays: A small brake test and a great brake test. The great brake test is the most important one and takes most time of both; about half an hour. During the great brake test the entire train is checked as described above. The small brake test is only a partial test: Only the first few wagons need to be tested.

The main reason to introduce the *great* brake test was the previous use of shunting for most freight trains: Each train that was to leave from a rail service centre was built out of separate wagons, each day anew. As a result, it had to be checked *always* whether all connections for the air pressure system were made properly. Nowadays there is a tendency to use so-called shuttle trains which are to be described in section "Shuttle types" of this chapter. These shuttle trains consist of a fixed set of wagons that are almost never broken apart and shunted, so there is no need to check every day whether all connections were made properly. Nevertheless, because of high
safety standards, this great brake test is in nearly all rail service centres still mandatory. Currently terminal operators try to get permission to abstain from an every-day great brake test.

The small brake test is more limited that the great one, and is used for two purposes: (1) It can be used as a replacement of the great brake test for shuttle trains of a fixed composition, or (2) it can be used to re-check a train that has been standing after the great brake test for over half an hour without an air-pumping locomotive or compressed-air tank attached to it. This latter use prevents that brakes are malfunctioning due to air that has leaked out of the tanks of the wagons after the great brake test.

In the model both brake tests are included, both using a fixed duration. For any train it can be specified which tests are to be performed, and which are not. Refer to section “D-locomotive operations“ of this chapter.

9.2.5. Locomotive storage

As soon as a train enters the rail service centre, pulled by an E-loc, then in most cases the E-loc is uncoupled and replaced by a D-loc for local movements as discussed before. But the now uncoupled E-loc must be kept in "storage" or be used to move another train out of the rail service centre's area. At least in most cases, E-locs are to be relocated to some "storage track" where they remain until further usage. The same holds for D-locs: When they are not in use they in most cases are "stored" at a special track, where they remain until they are used for another local movement.

This aspect is modelled by assigning to each rail bundle a parameter that specifies whether that bundle may be used either for storage of E-locs or for storage of D-locs. In this way the important flow of light engines across the terminal's yard, which has a high impact on the required infrastructure capacity\(^4\), gets more close to reality.

9.3. Train concepts

Every train that might be defined at a rail service centre is characterised by a number of operational settings like the moment of coupling to diesel locomotives (D-locs) or the yard bundle at which the train is to enter the system and wait for unloading. In practice several of trains will be characterised by the same settings, and about three to five different combinations of settings will be used at most rail terminals, thus yielding three to five different operational schemes. In the model a quick variant analysis of a rail service centre is obtained if the behaviour of a number of trains can be changed instantly by changing only one shared parameter that specifies the behaviour of all these trains.

In order to facilitate modelling trains and analysing different operational strategies, the so-called train concept is introduced. Such a train concept can be compared with an equipment type as previously discussed at Chapter 7: It fully defines the operational characteristics of a train that is of (or: refers to) that train concept. A train concept contains data on:

- the type of shuttle train,
- the locomotive operations,
- the handling operations, and
- the system exit/entry operations.

By changing one of these settings of a train concept, all trains that are defined by that train concept change in behaviour. All aspects included in train concepts will be discussed one by one.

9.3.1. Shuttle types - introduction

One of the most elementary settings of a train service is its type of shuttle: Three different settings are allowed in the model, each of them having different characteristics which are to be discussed shortly:

- pure shuttle;
- travel shuttle;

\(^4\)This impact is only modelled properly at DDM level of detail, as only that level includes block safety systems and individual tracks within rail bundles.
• split shuttle.
Shuttle trains are believed to be the only option for future railway transport of standard load units (i.e. containers and swap bodies) as they reduce costs of railway operations by eliminating shunting of wagons. The basic idea of shuttle trains is shown in figure 9.6.

![Diagram of shuttle train concept](image)

**Figure 9.6:** Concept of door-to-door transport.

Figure 9.6 shows a logistic chain in which load units are transported "door-to-door" by means of trucks (2 trips) and by means of trains (1 trip). Two rail service centres are involved in this entire transport chain, needed to tranship load units between trucks and trains twice. Obviously, figure 9.6 depicts only one out of a wide range of possibilities for door-to-door transport; in a number of cases one of the transports by trucks could, for instance, be replaced by a sequence of (1) transport by sea-going vessels, followed by (2) a second transport by trains (at the opposite side of the ocean), and finally ending in (3) transport by trucks again. However, what is important at this point is the type of train intended in figure 9.6.

• Pure shuttle trains
The type of train intended in figure 9.6 is the so-called shuttle train or, according to the definitions used in this thesis, the so-called pure shuttle train. A pure shuttle train drives between two rail service centres only, according to a fixed time table and consisting of a known, always equal, set of wagons. These wagons are never broken apart (except when damaged) and always run over the same route between both rail service centres. As such a shuttle train is never broken apart, no (expensive) shunting is needed, thus reducing the operational costs of the shuttle train. Even more, there is no need to organise and direct flows of (empty) wagons to rail service centres every day anew to create trains that match the requirements of that day, consisting of the right number of the right type of wagons: Shuttle trains always consist of the same, fixed set of wagons, and therefore eliminate the need to compose trains every day anew; every day the same wagons are used. The drawback of this concept is a reduced average occupancy per wagon, but this drawback is believed to be compensated by reduced shunting and overhead costs.

The situation shown in figure 9.6 depicts a somewhat wider part of the logistic chain than intended to model in this thesis; this thesis only aims to model the processes inside one of the two rail service centres (shown as grey bulbs), which is possibly linked to marine terminals or to a DistriPark as found at the Maasvlakte. Nevertheless, the basic idea about pure shuttle trains in this thesis is the same as depicted in figure 9.6: A regular service, arriving at a regular interval according to a fixed time table.

• Micro- and macro-system
The difference between the micro-system (i.e. the rail service centre only) and the macro-system (i.e. the full logistical chain from figure 9.6) is more important when discussing the next two types of shuttle trains: Travel shuttles and split

![Diagram of macro and micro system](image)

**Figure 9.7:** Train behaving as a travel shuttle in the micro system -- not in the macro system.
shuttles. In this thesis a shuttle is considered a travel shuttle only if it behaves as a travel shuttle within the micro system; such a travel shuttle could possibly be regarded as a pure shuttle train from the macro system's point of view. This is illustrated by figure 9.7 showing a train running between Mannheim and the Maasvlakte as a "pure" shuttle, but being handled within the Maasvlakte as a travel shuttle -- at both rail service centres that are present at the Maasvlakte the train is handled. In this thesis such a train is regarded a travel shuttle instead of a pure shuttle.

The inverse situation, shown in figure 9.8 is a train that behaves as a travel shuttle in the macro system, but behaves as a pure shuttle in the micro system, i.e. at the Maasvlakte. Such a train is considered as a pure shuttle in this thesis.

Figure 9.8: Train behaving as a pure shuttle in the micro system -- but as a travel shuttle in the macro system.

✧ Travel shuttle trains
Now, what is a travel shuttle? As can be seen from figures 9.7 and 9.8, a travel shuttle is a shuttle train that is handled at more than two rail service centres; at each of these rail service centres load units are either loaded or unloaded. Although this behaviour is somewhat different from that of a pure shuttle, still a lot of correspondence can be found between a pure shuttle and a travel shuttle: Both types of shuttles travel according to a fixed time table, always run over the same route between the rail service centres at which they are handled, consist of a fixed set of wagons, and are never broken apart except for repair of wagons.

Note that the main difference between the micro- and the macro system is the scale of the travel distances; if the distance between the Central RSC and the Sub-RSC from figure 9.7 were about 50 kilometres, then the micro system would be as shown in figure 9.8, in which the train behaves as a pure shuttle. An important consequence of this increased distance is, that there is (usually) no internal transport as meant in Chapter 7 between the terminals at which the travel shuttle halts.

✧ Split shuttle trains
The third type of shuttle train is the split shuttle. As shown in figure 9.9 the main difference is that such a shuttle is split into a (limited) number of parts which are loaded and unloaded separately, and afterwards are merged again into the same, complete shuttle train as it was before splitting. Analogous to the travel shuttle, it is important to distinct between the micro system and the macro system. As can be seen from figure 9.9, a split shuttle (in the sense of the micro system) may seem to be a pure shuttle to the macro system. And, in reverse, a pure shuttle in the sense of the micro system may appear to be a split shuttle from the macro system's point of view, as is shown in figure 9.10. In this thesis the latter type is concerned to be a pure shuttle instead of a split shuttle.

Figure 9.9: Train behaving as a split shuttle in the micro system -- but as a pure shuttle in the macro system.
Analogous to pure shuttles and travel shuttles, split shuttles always consist of the same wagons (except when in repair) and drive according to a fixed timetable. One important aspect of split shuttles as defined in this thesis is that the location inside the micro system at which the wagons are split in a number of individual sets of wagons, does not necessarily need to be the same location at which these sets of wagons are merged again after unloading and loading; there is some freedom in specifying the precise location at the terminal complex at which the wagons are to be merged again.

This aspect is a typical example of a parameter that is dependent of the type of shuttle that is used: For pure shuttles some other parameters apply, and for travel shuttles even more different parameters apply. All these special parameters will be discussed in the next section.

9.3.2. Shuttle types - options

Each type of shuttle train has a different set of additional operational options that specify its behaviour in the micro system. Therefore, some additional parameters need to be specified; parameters that are specific for each one of the three shuttle types. These additional parameters are discussed for each type of shuttle train.

- **Pure shuttle trains**
  Pure shuttle trains are to be handled at only one rail service centre in the modelled (micro) system. Even in those cases where a pure shuttle is carrying load units that are to be handled at two different rail service centres that have been modelled in the micro system, only one of those is to be selected for handling. An example of such a situation is shown in figure 9.11. (A situation as depicted in figure 9.11 can easily be created by changing a travel shuttle that was modelled to serve to rail service centres into a pure shuttle in order to study the effects of these two train types on the modelled system.)

This selection of the single rail service centre to call at is included in the model as one out five methods. Four of these methods are based on characteristics of the load units that are carried by the train. As will be explained in more detail in section "Loading and unloading" of this chapter, these load units are registered to be carried by some train by assigning load schedules, as defined at UCM level of detail, to this train. The five selection methods are then:

- select the RSC of the first package of load units on-train, which is defined by the first load volume of the first load schedule that is registered at the train;
- select the RSC at which the greatest number of load units is to be loaded;
- select the RSC at which the greatest number of load units is to be unloaded;
- select the RSC at which the greatest number of load units is to be handled, i.e. the sum of the numbers of load units that are to be loaded and of those that are to be unloaded;
- select a fixed, specified, RSC.

For the second, third and fourth method the selected rail service centre may be different for subsequent trains that are defined by the same train service and by the same load schedules, as the numbers of load units that are specified at these load schedules may be random numbers.
This kind of operation introduces an additional problem in the model: As all load units in the model are assigned to \textit{closed tasks} which run between two locations at the system's boundary, an alternative closed task must be selected for those load units that could not be handled at their originally intended rail service centre.

This is illustrated in figure 9.12, showing two rail service centres and two closed tasks, each closed task being spanned between either RSC and a marine terminal. If a pure shuttle train is carrying load units both for the first and for the second RSC, and the second RSC is selected for handling of the \textit{entire} pure shuttle, then all load units that originally were to flow over the first closed task are now commanded to flow over the second closed task.

\textbf{9.12: Selecting an alternative closed task for pure shuttle trains.}

\textbf{Travel shuttle trains}

Travel shuttles require some more additional parameters than pure shuttles do, to model their course in the micro system. These additional parameters will be shown from figure 9.13.

Figure 9.13 shows a typical course of a travel shuttle in the modelled system in which several rail service centres are defined. Supposing that the train is carrying (or is to pick up) load units at all three rail service centres, then some \textit{sequence} must be defined in which these rail service centres are to be called. This is the first parameter.

After the travel shuttle has been handled at one rail service centre, it has to move to another one, i.e. to the next rail service centre in the travel shuttle's sequence. Supposing that it is known at what time the train is to be handled at this next rail service centre, it must be determined at what time the train has to start moving from its \textit{current} rail service centre: Should it (1) wait as long as possible at this current rail service centre, or should it (2) move as soon as possible to the next rail service centre, or should it (3) be relocated to some yard bundle after it has been handled at the current rail service centre, and wait at that very yard bundle before it starts for the next rail service centre in sequence? This is the second parameter.

In the situation depicted in figure 9.13 the travel shuttle is visiting each rail service centre twice: Once for unloading and once for loading. An alternative operational sequence would be to visits each rail service centre, combining unloading and loading in one visit. The relative sequence of loading and unloading when combining them in one visit could be either \textit{separate} or \textit{concurrent} as will be described in section "Handling operations" of this chapter. This is the third parameter.

All options for these three parameters will be discussed in detail, one by one.

\textbf{Travel shuttle trains - RSC sequence}

Several options to determine the sequence in which the various rail service centres are to be called are included in the model. These options are:

- sorting on the load schedules that are included in the train. As each load schedule defined at a train refers to exactly one rail service centre, this sequence can easily be determined. By manipulating the sequence in which the load schedules are included in the train the sequence can be manipulated "by hand".
• sorting on the numbers of load units to unload. Those rail service centres at which the largest numbers of load units are to be unloaded are visited first, thus giving priority to the greater streams.
• sorting on the numbers of load units to load. Those rail service centres at which the largest numbers of load units are to be unloaded are visited last (not first), thus giving the rail service centres with the greater streams more time to prepare for the train.
• sorting on time of unloading. As will be seen when discussing the train services it is known for each rail service centre at which loading and unloading must start and end. The time specified for the start of unloading the will normally be used as a basis for the rail service centre sequence. Therefore this option is likely to be used most of the time.

Note that a travel shuttle only visits a rail service centre whenever at least one load unit is to be handled at that rail service centre. In the particular case that either no load volumes are defined for one handle direction, or that all load volumes for one handle direction result in a total of zero load units, then the particular rail service centre is skipped for that handle direction; for the opposite handle direction that rail service centre may still be called.

♦ Travel shuttle trains - inter-RSC relocation

![Diagram of travel shuttle trains - inter-RSC relocation](image)

**Figure 9.14:** Relocation between rail service centres for travel shuttles.

After a travel shuttle has been handled at one rail service centre and is ready to be handled at another rail service centre, it is very likely that there is some time span between the end of handling at the first rail service centre and the start of handling at the next rail service centre. This is shown in figure 9.14 as the "gap" between both handle actions. The travel shuttle may have to relocate to some yard bundle in this "gap" in order to release transhipment capacity at the first rail service centre. This is the first option. Note that this kind of relocation is different from relocation between unloading and loading as will be discussed at section "Handling operations" of this chapter.

Another option is, not to relocate the train to some yard bundle, but to send it as quickly as possible to the handle bundle of the second rail service centre. This kind of operation releases transhipment capacity at the first rail service centre, but occupies additional transhipment capacity at the rail service centre. This is the second option.

Finally, it can be chosen to let the train remain as long as possible at the first rail service centre's handle bundle, sending it away in the last moment only to arrive just in time for the start of the operations at the second rail service centre. This is the third option.
Chapter 9. Railway aspects

♦ Travel shuttle trains - load mode

The third and last parameter defining the behaviour of travel shuttles is the load mode. Only two options are available for this parameter:

- Either the travel shuttle visits all rail service centres twice as shown in figure 9.13, once for loading and once for unloading; this option is called separate travel loading.
- or the travel shuttle visits all rail service centres only once, combining unloading and loading in one visit; this option is called concurrent travel loading. In that case figure 9.13 turns into figure 9.15.

The advantage of concurrent travel loading is reduction of the number of train movements and therefore simpler railway movements. The disadvantage is that load units at the first and second rail service centre are to be loaded earlier than in figure 9.13, thus forcing the terminals to have the load units prepared for loading sooner than required at figure 9.13.

Note that the settings for concurrent- and separate travel loading are different from normal concurrent and separate loading as will be discussed in section "Handling operations" of this chapter. In advance to this section it can be said that normal concurrent loading for travel shuttles is only used when concurrent travel loading is used, too. Otherwise the setting for normal concurrent loading is ignored.

Also, in advance to section "Handling operations" it can be said that normal relocation between loading and unloading is only used for travel shuttles when concurrent travel loading is used. As with any shuttle type, relocation between loading and unloading is ignored when normal concurrent loading is in use. In that case there are no separate unloading and loading actions to relocation in between.

♦ Split shuttle trains

The last type of shuttle train to discuss is the split shuttle train. This type of shuttle is characterised by only two, relatively simple, parameters: (1) The location where to split the shuttle and (2) the location where to merge the shuttle again.

The location at which the shuttle can be split is, as mentioned before, always a rail bundle and never a rail track that is not related to a bundle. This is a consequence of the importance that bundles play in system analysis after simulation of the model. The precise bundle at which to split can be either one of following options:

- the first bundle at which the train halts, for whatever operation. As will be shown later on, this can either be a bundle at which the train is unloaded, or a bundle at which the train halts to uncouple its E-loc.
- the bundle of the first rail service centre at which the train is handled, i.e. the rail service centre at which the handling operations are to start first of all rail service centres for which the train is carrying load units.
- some fixed bundle, specified by the user of the terminal design aiding tool. This bundle must have been marked as fit for splitting trains and can be specified either at the train's train service or at the first rail service centre at which the train is to be handled.

For merging the train again only two options are available: It must be merged again either:

- at the same location at which it was split before, or
- at some fixed bundle, specified in the same way as the previously discussed fixed bundle for splitting.

9.3.3. System entry- and exit operations

Conventional technology requires trains to be uncoupled from their E-locs with which they entered the system; as can be seen from figure 9.5 conventional transhipment technology, using gantry cranes, does not allow the presence of E-locs at handle bundles because of the required catenary. This implies replacement of the E-loc by a D-locs by means of uncoupling and coupling as described before, and a specification of the location where these E-locs are to be uncoupled. This latter aspect is included in the model as "system entry- and exit operations". These operations are stored in the train concept's yard mode.
System entry- and exit operation in this model simply mean that it can be specified for a train to drive to some known yard bundle upon system entry first, and to wait there until unloading starts. Alternatively, the train can be commanded to drive immediately to the bundle at which it will be handled. In the latter case all operations like locomotive changes (if required at all) are to occur at the very handle bundle; in the first case all are to occur at a yard bundle.

![Diagram showing system entry via a yard bundle and direct to handle bundle.]

**Figure 9.16:** Both types of system entry.

Figure 9.16 will clarify this by showing two trains, behaving in a different way upon system entry. The first train is driven upon system entry onto a (fixed) yard bundle where its locomotive is changed and where it waits until further handling at the handle bundle. The second train immediately drives to the handle bundle where its locomotive is (possibly) uncoupled and where it waits until actual unloading starts. In this latter case either (1) the real-life handle bundle must be made fit to receive E-locs\(^5\) by removable catenary\(^6\) or by alternative transhipment equipment\(^7\), or (2) the train must be pushed onto the handle bundle instead of being pulled.

Operations for system exit are quite analogous to those for system entry: For each train concept it can be specified whether a train is to leave the system directly from its last handle bundle, or that is has to relocate to a (fixed) yard bundle first, before it has to wait for system exit. In the latter case the E-loc by which the train will be pulled out of the system will be coupled (if required) at that yard bundle. Both brake tests, if required, are to occur at this same yard bundle too. This will be discussed in the next section as well.

In short, the operations for system entry and exit, to be specified at train concepts, are:
- is system entry to occur via a yard bundle? If so, then:
  1. specify the fixed yard bundle to enter on.
- is system exit to occur via a yard bundle? If so, then:
  1. specify the fixed yard bundle to leave from.

The settings of these options will influence the required bundle capacities, both for yard bundles and for handle bundles. In case of *Fließverfahren* (which it to be discussed at section "Handling operations" of this section, the option of system entry and -exit via a yard bundle is likely to be used. In case of *Biedenverfahren* the option of direct system entry and -exit, directly from the handle bundle, is to be used.

### 9.3.4. **E-Locomotive operations**

Light engines appear to play an important role in this model and in the actual, real-life systems as they put a high demand on required infrastructure capacity. Especially in those cases where block safety systems are used, each single locomotive movement blocks a number of tracks and points for one or two minutes, thus hindering movements of actual trains. By consequence, light engine movements need to be included in the model to obtain a satisfactory impression of the rail terminal's capacity and required infrastructure. A number of these aspects have been pointed at before, but all will be discussed here.

---

\(^{5}\)In those cases where trains are pulled by D-locs on the public rail network, these trains still arrive in the modelled micro system, pulled by "E-locs". In such cases the difference between "E-locs" and "D-locs" still remains a difference between internal and external locs; only in this special case both internal and external locs are diesel-powered.

\(^{6}\)As SNCF has experimented with at the Trappes Commuter pilot plant.

\(^{7}\)Like the Krupp transhipment system, refer to [16].
In the model a distinction is made between locomotive operations of D-locs, which are to be considered as internal locomotives, and operations of E-locs, which are to be considered as external locomotives. This subdivision will be seen from this section, too. For a thorough analysis of the impact of light engine operations on the required infrastructure a block safety system needs to be included in the model. Such a block safety system, however, is not found until DDM level of detail; and is not found at SDM level of detail. Therefore, both levels will have a different aim: SDM is used to get a rough insight in the tracks needed, in bundle capacity, in the D-locs needed and in the effects of rail-operations on the remaining (Maasvlakte) modelled system. DDM is used to fine-tune the modelled system, based on the results of SDM level of detail.

E-loc operations
The most important setting about locomotive operations is whether E-locs with which trains arrive in the system are to be uncoupled from their trains or not. If not, there is no need for D-locs to be present at all. Otherwise, the complexity of the operations at the terminal yard will increase. After the E-loc has been uncoupled, it is stored at a bundle that is fit for E-loc storage according to some rules to discuss at section "Bundle selection" of this chapter.

Apart from this major setting, three further settings determine E-loc operations per train concept: (1) First of all, it may be specified whether the great brake test should be performed or not; the duration of this great brake test is a constant, user-specified, value which is equal for all great brake tests that are performed during simulation. This great brake test is performed at the last bundle at which the train halts before system exit, so this could be either a handle bundle or a yard bundle, depending on the settings for system exit. (2) Second, it may be specified whether the small brake test is to be done, just upon system exit. This small brake test is performed at the same bundle as the great brake test was performed, and therefore at the last bundle at which the train resides before actual system exit. Analogous to the great brake test, the small brake test lasts equally long for each occurrence during simulation. Nevertheless, the duration of the small brake test will be definitely smaller (about 5 minutes) than that of the great brake test (about 30 minutes). (3) Third, it can be specified whether the E-loc should be coupled (if ever uncoupled only) before or after the great brake test; in this way it can be modelled whether the great brake test requires the presence of the E-loc (i.e. couple before great brake test), or whether the great brake test is to occur by using compressed-air tanks that are installed at the terminal’s yard (i.e. couple after great brake test).

In summary, the settings for E-loc operations as stored at train concepts are as follows:

- uncouple E-loc during operations at the terminal yard?
- perform great brake test?
- perform small brake test?
- couple E-loc before the great brake test of afterwards? Obviously, this setting is only valid when the E-loc has been uncoupled previously.

Next to these aspects regarding operations of E-locs, some other aspects regarding operations of D-locs are included in the model, which will be discussed right now.

9.3.5. D-Locomotive operations

For those trains that require the E-loc to uncouple some additional parameters control the use of D-locs during their stay at the terminal yard. It is possible to specify whether the D-loc should remain coupled all time to the train, or whether the D-loc should be uncoupled and released as soon as any local movement of the train ends. Allowing the D-locs to uncouple between local movements will increase the number of light D-engine trips, and therefore will increase the required infrastructure capacity. On the other hand, it will decrease the number of D-locs needed at the terminal yard as they can be used more efficiently. Whenever a D-loc is uncoupled it is stored at a bundle that is fit for D-loc storage according to some rules to discuss at section "Bundle selection" of this chapter.

So, in short, the only option for D-loc operations is:

- release D-loc after each local movement?

All speeds of trains are included as fixed, constant values in the model. It does not include any acceleration or deceleration of trains due to the great complexity this would imply for the train planning algorithm at DDM level of detail, which does use block safety systems and path reservations of trains. This will be discussed in
more detail is section "Block safety systems" of this chapter. Therefore, only four different speeds are discerned in the model, which are:
- speed of a light D-engine;
- speed of a light E-engine;
- speed of a coupled D-loc, pulling a train of any length and tonnage;
- speed of a coupled E-loc, pulling a train of any length and tonnage.

Some other options apply for locomotive operations, but these options are not stored at train concepts; these further options are discussed at section "Locomotive controls" of this chapter.

### Moments of coupling and uncoupling

At a number of occasions a locomotive needs to be coupled. In most cases, this locomotive is a D-loc. The occasions at which a loco may need to be coupled to a train, and the occasions at which a loco may be uncoupled from a train are shown in table 9.1. This table also shows the type of locomotive that is required to couple. In those cases where the type of locomotive equals "Any", it is free to use either a D- or E-loc, depending (1) on their availability and (2) on the settings of the locomotive manager which are to be discussed at section "Locomotive controls" of this chapter, which define the preference for the use of D-locs over the use of E-locs for local movements. In those cases that E-locs are to remain coupled all time to the trains, then the train will be ready for any (local) movement at any time, and table 9.1 does not need to be used.

<table>
<thead>
<tr>
<th>Action</th>
<th>Uncoupling Required</th>
<th>Coupling Required</th>
<th>Loc-type</th>
</tr>
</thead>
<tbody>
<tr>
<td>System entry</td>
<td>Yes</td>
<td>Yes</td>
<td>Any</td>
</tr>
<tr>
<td>System exit</td>
<td>Yes</td>
<td>No</td>
<td>E-loc</td>
</tr>
<tr>
<td>Relocation(^1)</td>
<td>Yes(^2)</td>
<td>Yes(^3)</td>
<td>Any</td>
</tr>
<tr>
<td>Travel relocation</td>
<td>Yes(^2)</td>
<td>Yes(^3)</td>
<td>Any</td>
</tr>
<tr>
<td>Travel RSC change</td>
<td>Yes(^2)</td>
<td>Yes(^3)</td>
<td>Any</td>
</tr>
<tr>
<td>After splitting</td>
<td>Yes(^2)</td>
<td>Yes(^3)</td>
<td>Any</td>
</tr>
<tr>
<td>After merging</td>
<td>Yes(^2)</td>
<td>Yes</td>
<td>E-loc</td>
</tr>
</tbody>
</table>

\(^1\) = Still to be discussed in section "Handling operations".
\(^2\) = After local movement, if D-loc is to be released after use.
\(^3\) = Before local movement, if D-loc is to be released after use.

Table 9.1: Occasions for coupling and uncoupling of locomotives, provided that E-locs are to be uncoupled upon system entry.

### 9.3.6. Handling operations

Handling at rail service centres is modelled in more detail than handling at marine terminals and barge terminals, which is a consequence of the model's focus on rail service centres. While marine terminals did not really care about loading sequences, except when explicitly modelled by load volumes, rail service centres do include some options to manipulate and control the sequence of loading and unloading of individual load units. By using one or more of these options, that are stored at train concepts, most handling sequences encountered at rail service centres can be modelled. First of all, some of those handling sequences found in reality will be discussed. Afterwards, it will be shown how these real-life handling sequences can be modelled by train concepts.

### Some real-life handling sequences - Fließverfahren

The first handling sequence to discuss is Fließverfahren which is not only found at inland rail terminals, but also at marine bound rail service centres. As shown earlier (refer to figure 9.4) in case of Fließverfahren trains are unloaded quickly, storing all load units in a temporary stack, and are then moved out to some storage bundle where these trains are to wait a few hours for loading. When the trains are to be loaded again, they are moved into the handle bundle again and are loaded as quickly as possible, after which they leave the terminal's yard. The advantage of this type of operation is a reduced number of handle tracks; the disadvantage is a more complex operational process at the rail service centre.
Note especially that unloading was to happen as quickly as possible, not according to some special unloading sequence. As there was no need to use some special sequence for unloading, the train could be unloaded simply head to tail; this type of unloading is commonly called *grazing unloading* and is shown in a schematic way in figure 9.17. The horizontal axis of this figure depicts the time and the vertical axis depicts the position of the gantry crane that is unloading the train (assuming that only one gantry crane is used for unloading). The slant line shows that the gantry crane moves at a more or less constant speed along the train.

**Figure 9.17:** Grazing unloading.

Loading was to happen as quickly as possible too when using Fließverfahren, but could nevertheless have required some special loading sequence: It could, for instance, have been required that load units that belong to one particular owner are grouped together on the train. But because of the sorting character of the temporary stack from which the load units are transported to the train, this loading sequence is already established by the internal transport system that delivers the load units to load to the train. Therefore, the train can be loaded "grazing" too, even when a special loading sequence is required.

One advantage of Fließverfahren is separate loading and unloading of the train, thus ensuring that the train is completely empty when loading starts. This approach prevents problems that might occur during concurrent loading and unloading: In that case a load unit could have to be loaded on a certain, known position on the train, which position still could be occupied by another load unit that has not been unloaded yet. Depending on the freedom in planning the load scheme of the train, and therefore in the freedom in finding a new, non-occupied loading position for the load unit, this problem can be solved by either (1) finding another free position to load the load unit at, or (2) stacking one of both clashing load units next to the train. This problem is illustrated in figure 9.18.

**Figure 9.18:** Clashing load units in case of concurrent loading and unloading.

◆ **Some real-life handling sequences - Bedienverfahren**

The second handling sequence to discuss is *Bedienverfahren*. As found when discussing figure 9.4 this handling concept is characterised by concurrent loading and unloading of load units. It is mainly found at inland terminals where it is used for direct interfacing with trucks that enter the rail terminal, either to pick up or to deliver load units -- load units are transshipped directly between the train and the trucks. The advantage of this approach is (in most cases) a better experience of service of the rail terminal for the drivers of the trucks.

A complicating factor when using Bedienverfahren is the lack of control on the arrival times of road trucks; in the ideal case all trucks arrive equally in time, those one fetching a load unit first, followed by those ones delivering a load unit. But unfortunately, in reality all trucks arrive somewhat peaked in time (refer to the discussion on intensity contra schedules too), and not in a neat sequence of fetching trucks first and after that delivering trucks, but in any random sequence. Because of this random arrival of road trucks for both loading and unloading a situation as depicted in figure 9.18 in which loading and unloading operations clash, is likely to occur.

Another disadvantage of using Bedienverfahren is the decreased efficiency of the use of the gantry crane over the train: For each truck that arrives and needs to be loaded or unloaded, there is a great chance that the load unit that has to be transshipped must be unloaded from, or loaded onto, a rail wagon that is over a hundred metres away from the current position of the gantry crane. This causes additional driving of the gantry crane over the train, thus reducing the efficiency of the transshipment process.

◆ **Parameters of handling operations**

From the above operational schemes a number of parameters that define handling behaviour of rail terminals can be found. These parameters are included in the train concepts with a few standard options, options that
allow quick modelling of most real-life situations. Most of these parameters will be used later on as an important basis to find the precise system entry- and exit time of each individual load unit that is transported by a train. The parameters defining handling operations and their options, as included in train concepts, are:

- Relative position in time of loading and unloading, with options:
  1. separate loading and unloading. If this option is used, then the train may be forced to relocate between unloading and loading to another bundle, thus modelling full Fließverfahren; otherwise the train remains standing at the handle bundle where is has been unloaded, waiting for loading.
  2. concurrent loading and unloading.

- Relative positions of all load units on the train, with options:
  1. random positions; all load units are placed in a random sequence at the train\(^8\).
  2. sorted by client; all load units of the same client are grouped together. There is no control on the eventual sequence of "client blocks" on the train.
  3. sorted by origin; all load units that are loaded from the same origin, i.e. from the same bounding element at the opposite side of the system, are grouped together. There is no control on the eventual sequence of "origin blocks" on the train.
  4. sorted by client and origin; all load units are first sorted into groups of the same client and then sorted into groups of the same origin. In this way a number of blocks referring to the same client, but referring to different origins are placed one behind the other.

This parameter may be different for loading and unloading. Note that for unloading the "origin" must be replaced by "destination".

- Absolute positions of all load units on the train, with options:
  1. evenly spread over the full train; all load units that have been placed in the right order according to the previous parameter are placed at discrete positions on the train, all load units equally divided.
  2. condensed in the head of the train; all load units (placed in the right order according to the previous parameter) are placed one behind the other without interleaving gaps on the train, starting from the head of the train. In this way the tail of the train is likely to remain unloaded.

This parameter may be different for loading and unloading.

- Relative moment of handling, with options:
  1. grazing handling; all load units are handled from head to tail in the sequence they are placed on the train.
  2. random handling; all load units are handled in a random sequence, independent of their place on the train.

This parameter may be different for loading and unloading.

- Freedom in planning the load scheme of the train, which can be used when a clash between loading and unloading load units as shown in figure 9.18 occurs, with options:
  1. freedom to find a new, empty location on the train to load the load unit onto.
  2. no freedom; the load unit that could not be loaded is to kept in a local storage next to the train until the requested location is freed.
  3. no freedom; the load unit that blocked loading is removed from the train and is kept in a local storage next to the train; the load unit that could not be loaded is now transhipped onto the train in the just freed place.

All of these parameters need some more discussion, which will be done by showing their effects on important issues as travel distances of gantry cranes and the use of equipment for internal transport movements.

---

\(^8\)These are only the relative positions of load units, relative to each other; their exact positions are determined by the next parameter.
Separate and concurrent loading
The basic assumption in this section is that handling of trains is to occur by means of conventional technology, that is, by means of gantry cranes. This situation is depicted in Figure 9.19 in which a single gantry crane is stacking on the train. The two arrows indicate the driving direction on the train. Supposing that separate unloading and loading is used, then the MTS trailer depicted in Figure 9.19 too will first be used to unload the train, and then to load it again. This handle sequence is shown in Figure 9.20 along with concurrent unloading and loading. This figure also shows that in case of separate unloading and loading the train may be relocated to be stored at a non-handle bundle for a while.

![Figure 9.19: Using gantry cranes for train handling](image)

<table>
<thead>
<tr>
<th>Separate unloading and loading</th>
<th>Concurrent unloading and loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>Time</td>
</tr>
<tr>
<td>possibly relocate</td>
<td></td>
</tr>
</tbody>
</table>

![Figure 9.20: Separate and concurrent loading and unloading](image)

Clashing load units
In case of concurrent loading and unloading a clash between two load units as shown in Figure 9.18 could occur. In case of such a clash there are three solutions for the deadlock: The simplest solution is to place the loading load unit on the nearest free position on the train, thus possibly creating a new deadlock for another, future, loading load unit. This is illustrated in Figure 9.21. Whenever no free position on the train is available to place the loading load unit on, that load unit is to be stored next to the train, waiting for some (random) position to become free. This option was called the first option at parameter "Freedom in planning the load scheme of the train".

![Figure 9.21: Solving first clash by freedom in planning the load scheme of the train, causing a second clash](image)

---

9 Because of the structure of the model it is always guaranteed that all loading containers will fit on the train. In other terms, it is always guaranteed that the number of positions, available on the train, is sufficient to load all loading containers in the end.
Chapter 9. Railway aspects

The second solution to solve the deadlock was called the second option at parameter "Freedom in planning the load scheme of the train" and is depicted in figure 9.22. In that case the loading load unit has priority over the unloading load unit, and therefore a free position on the train will be created by removing the blocking unloading load unit from the train and storing it next to the train to be picked up later again for internal transport. This will cause the gantry crane to drive a little more than the previous solution (namely: Store the unloading load unit and drive back to the loading load unit) and will also introduce an extra transhipment movement (namely: For the unloading load unit).

The third solution was called the third option at parameter "Freedom in planning the load scheme of the train" and gives priority to the unloading load unit instead of to the loading load unit, and therefore stores the loading load unit next to the train where it will wait until its position is freed. Note these latter two solutions can only be chosen for if freedom in planning the loading positions of load units, as used in figure 9.21, is not allowed.

![Figure 9.22: Solving clash by giving priority to the loading load unit.](image)

<table>
<thead>
<tr>
<th>Time</th>
<th>Time</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Position along the train" /></td>
<td><img src="image" alt="Position along the train" /></td>
<td><img src="image" alt="Position along the train" /></td>
</tr>
<tr>
<td><img src="image" alt="First option" /></td>
<td><img src="image" alt="Second option" /></td>
<td><img src="image" alt="Third option" /></td>
</tr>
<tr>
<td>Freedom in planning</td>
<td>Priority to loading load unit</td>
<td>Priority to unloading load unit</td>
</tr>
</tbody>
</table>

L = handle loading load unit  
U = handle unloading load unit  
○ = getting up load unit  
□ = putting down load unit

![Figure 9.23: Handling and gantry crane movement for all three options for "Freedom in planning the load scheme of the train".](image)

Figure gives an impression of the additional movements and transhipment actions of the gantry crane in case the second or third solution are used. These additional movements are only to occur whenever a clash as shown in figure 9.18 is encountered which is only likely when concurrent loading and unloading is used.

- **Absolute positions on-train**

The chance on the occurrence of such a clash is influenced by the selected option for the relative positions of all load units on the train. If all load units are to be condensed in the head of the train, as is shown in figure 9.24, and if this kind of relative positioning is used both for unloading and for loading load units, then it is very likely that a clash as meant in figure 9.18 is to occur. Referring to figure 9.20, it will be obvious that these clashes can only occur during concurrent unloading and loading.

![Figure 9.24: Condensing all load units in the head of the train.](image)
If, on the other hand, all load units are spread equally over the train, either for unloading or for loading, or possibly for both, then the chance on clashing load units will decrease. This is shown in figure 9.25. This figure rightly suggests that all load units are spread such that one load unit is placed at the very head of the train, and that another load unit is placed at the very tail of the train. All other load units are equally spread in between.

The advantage of using the option to condense all load units in the head of the train is a reduction of the average travel distance for the gantry crane. But in the real-life world it is not always allowed to use this option, for a number of reasons. Some of these reasons are:

- not all wagons are suitable to carry all load units. This can either be caused by limitations in payload of the wagons, or by positions of the "pins" that are used to fix the load units on the wagon. This kind of restriction is not included in the model.
- some wagons may have a common destination when the train is split outside the system's boundary in a number of separate wagon sets. Or, in terms of figure 9.10, the train may be split in the macro system but not in the micro system. Therefore, all load units of the same destination are to be grouped together. This kind of restriction is modelled in a simplified way by using equally spread loading combined with sorting of load units on their client, too.

Any combination of figures 9.24 and 9.25 for loading and unloading direction is allowed. It can, for example, be specified that when the train arrives at the terminal all load units are spread equally over the train because of handle actions on other, previously called rail terminals; this is the unload direction. In the meantime it can be specified that all load units that are loaded at the modelled rail service centre are to be condensed at the head of the train, thus reducing the average travel distance for the gantry crane; this is the load direction.

♦ Relative positions on-train
Some other options than sorting per client are available; it is allowed

1. to sort load units per origin or destination (i.e. the location opposite of the load unit's closed task, opposite from the considered rail service centre; refer to figure 9.26 to an example),
2. to sort them both on client and on origin or destination, or
3. to sort them not at all, locating them in a random sequence on the train.

These options will be explained. Note that all types of sorting described in this section override any sequence of load units as could be derived from the timing at their respective load volumes. In other terms, all information on timing as stored at load volumes is ignored when load units are defined at trains that arrive at a train handle area, both at SDM and at DDM level of detail. At UCM level, on the contrary, this information is preserved and used during simulation. Refer to section "Timing of load units" of this chapter, and to Chapter 6 as well.
If *random* sorting is used, then all load units are placed in a random order, regardless of their origin, destination or client. This kind of "sorting" is normally encountered at rail terminals both for loading and for unloading of trains. The main cause of this frequent occurrence are the restrictions that are imposed on the payload of rail wagons as well as the limited loading schemes per wagon due to the configuration of the "pins" at which load units are to be fixed to the wagon. Random sorting aims to model these and alike effects that result in a hardly predictable sequence of load units on trains. Note that the sequence in which load units were defined at their load volumes is totally ignored. An example of a random sorted unloading train is given in figure 9.27, both for a condensed and for an equally spread load scheme.

If load units are sorted by client then a situation as shown in figure 9.28 is created. This kind of sorting is encountered occasionally in practice on behalf of special wishes of the client of the load units. Sometimes this option is encountered when a train has to be split in the *macro system* (refer to figure 9.10) into a part that belongs to one client only and is driven separately to the "plant" of that particular client, somewhere in the macro system. Obviously, this type of operations is only encountered for those clients that own a substantial number of load units.

If load units are sorted by origin or destination then a situation as shown in figure 9.29 exists. This type of sorting is used to facilitate internal transport at the modelled rail service centre, in combination with *grazing loading*, and can only be used when there are no major physical restrictions imposed on the load scheme by the rail wagons. It is encountered less frequently for unloading of trains. It is not even required for *split shuttles*, as all load units on such trains are *always* automatically sorted by rail service centre first, and *then* possibly sorted by origin of destination, relative to the various rail service centres at which each separate part of the split train is handled. This is shown in figure 9.30.

Figure 9.30 shows a split shuttle that carries load units for two rail service centres (RSCs) that are both modelled in the micro system (not shown in figure 9.26). The separate parts of the entire train that are eventually to be driven to the two rail service centres are treated as individual parts, each being sorted for itself. Nevertheless, both separate parts are sorted in the same way, as the type of sorting to use is related to their mutual train's concept.

In case both sorting by client *and* sorting by origin or destination are used, then all load units are *first* sorted on client and *then*, per block of equal clients, on origin or destination. The choice for this sequence is not based on
the assumption that the interest of clients is superior to the interest of the rail service centre, thus yielding a client-friendly operation.

- **Relative moment of handling**

  While the previous two parameters, relative- and absolute positions on-train, determined where load units were to be placed on a train, both for loading and for unloading, they did not specify when these load units were to be loaded or unloaded. This aspect is defined by the relative moment of handling, which can be either grazing or random.

  Grazing loading or unloading is shown in figure 9.31. All load units are handled, or at least are intended to be handled, in the sequence they are located at the train. The direction, both for loading and for unloading is from head to tail. This type of loading and unloading is likely to be found at rail terminals that have an internal transport system which can be controlled such that a fluent transport to and from the gantry crane is obtained. Note that, in order to determine the precise timing of load units some additional information is needed as the total duration of the loading or unloading operation. This aspect is discussed at section "Timing of load units" of this chapter.

  Random loading or unloading is somewhat more "fuzzy" than grazing loading or unloading and is mainly found at rail terminals that operation according to the concept of Bedienverfahren and deliver directly to road trucks. It will be clear that random loading causes more driving of the gantry crane than grazing loading does. The eventual sequence is totally random, all operations being spread equally in time. Figure 9.32 gives an example of random loading.

  Note that both for random and for grazing handling the moment at which the load unit is to be loaded or unloaded is the final system entry- or exit time of the load unit at the rail service centre. This requested time is respected as much as possible during simulation and is also used as a basis for the equipment optimisation model (refer to Chapter 7), but it is not always guaranteed that this time and loading sequence of the train are exactly satisfied. In some cases the equipment optimisation model may find it more cost-effective to change the sequence in order to increase the degree of utilisation of an MTS trailer. This will mainly occur for multiple-load unit equipment types (MLU-types).

### 9.4. Train services

Based on the train concepts so-called train services can be defined. These train services model a real-life train service that is to enter and leave the system according to some fixed timetable, and is to load or unload a number of load units at one or more rail service centres in the micro system.

This section describes all aspects that are included in the model, which are the train's timetable including disturbances, its length, its numbers of load units to handle and their timing, its concept and its bundle selection rules. All these aspects influence the behaviour of the train in the modelled system.

### 9.4.1. Timetables

Trains use to drive according to fixed timetables, even trains that are handled at rail service centres. The reason for this behaviour lies in the strong relation of even those trains with operations on the public rail network: Eventually all freight trains that are handled at any trail terminal are to arrive at, and leave from that terminal.
over the public rail network. And in most cases this public rail network has a pretty high occupancy, leaving little space for delays in train movements. Even more, if some slight delay is allowed at a particular part of the public rail network, then on a next part of that same network that has a higher occupation, the same delay will turn out to cause further delay of the train. Therefore, timetables of trains are quite strict, even for freight trains.

In the model timetables are specified by three parameters: the time of system entry, the random disturbance on this, and the time of system exit. Note that there is no disturbance on the time of system exit: The train is forced to leave the modelled system always in time.

![Example of system entry- and exit track.](image)

**Figure 9.33:** Example of system entry- and exit track.

These times for system entry and system exit specify the moment in time at which the train either pops up at the train service's system entry track, or should disappear at the train service's system exit track. An example of these two tracks is given in figure 9.33. Each single train service may have a different track for system entry and a different track for system exit, although in practice most train services are likely to use the same track for system entry, and the same track for system exit.

At DDM level of detail, which includes a block safety system, special care must be given to prevent an invalid situation in which trains are allowed to enter or leave the modelled rail yards at a higher frequency than allowed in practice. This will be illustrated by referring to figure 9.33 in which a total of three full sections are inserted between the system entry track and the start of the rail yards, and another, single section is inserted between the end of the rail yards and the system exit track. All four sections are part of the public rail network connecting to the rail yards, and therefore have a considerably longer block length (i.e. one section) than found at the rail yards. In this way there is a guarantee that trains may never enter or leave at a higher frequency than allowed in practice. Even more, the effects of delays in train entry due to local rail operations within the yards that use critical sections needed for train entry, become more close to reality as in such case a queue of trains, waiting on the public network to enter the yards, appears which are standing at an intermediate distance of about 1500 meters. This causes further delays in system entry of following trains.

There is a slight difference in the way the time for system entry and the time for system exit are handled: While the system entry time is always respected (including its disturbance) and therefore trains always pop up in time at their respective system entry tracks, the system exit time is not guaranteed to be respected: In those cases where delays have occurred at the modelled yards, either because of transhipment or because of insufficient rail infrastructure capacity, trains will leave the system too late. However, this amount of time is registered and is used as an indicator for the terminal's performance, at analysis.

### 9.4.2. Operations

Most operational parameters of any train are stored at its associated train concept; each train is related to precisely one train concept determining its shuttle type, its locomotive operations, its yard characteristics and its handling characteristics. Train concepts therefore highly influence the operations at the modelled rail service centre(s).

---

\(^{10}\)In the Netherlands this block length is approx. 1500 meters long (Kreutzberger [15]).
An example of an operational scheme is given in figure 9.34. It shows all operations of a pure shuttle train that is to enter and leave the system via a yard bundle, and that is to relocate between unloading and loading to save transhipment capacity. In this example it is assumed that the E-loc with which the train enters the system is to remain coupled to the train all time it resides in the system. In this way a relative simple operational procedure is obtained, with only 13 different operations. Note that this example models the concept of Fließverfahren as discussed previously.

A number of additional operations may be inserted in this operational scheme, thus complicating the operations at the rail terminal. It may, for instance, be required that the E-loc that pulled the train into the system from its system entry track on, is replaced by a D-loc for local movements. In that case two additional operations are inserted before operations C and K: Uncoupling the E-loc and coupling the D-loc before operation C, and the reverse of these before operation K.

Furthermore, it could be required that the D-loc is to uncouple from the train after each local movement. In that case two more additional operations need to be inserted for each local movement: One before the local movement to couple a D-loc to the train, and another one after the local movement to uncouple the D-loc from the train and to store the now light D-engine at the appropriate storage bundle.

In order to control the great number of different operations during simulation a structure consisting of tasks is defined as described in Van Zijderfeld [38] and [39] that allows relatively simple sequencing of distinct operations, along with checks on pre-conditions.

Timing of rail movements
Timing of train movements as shown in figure 9.34 requires a lot of planning and is handled different at SDM level than at DDM level of detail. This is a consequence to the presence or lack of a block safety system in that level of detail.

From a simulation point of view, timing of train movements is simple at DDM level of detail than at SDM level of detail. This is caused by the fixed usage of infrastructure, which allows to predict well at what time a train should start driving over a set of rail paths to guarantee its in-time arrival at its next destination. Naturally, this guarantee can only be given when a train starts to check for available rail paths sufficiently long before its planned arrival or departure. Therefore a so-called claiming safety margin is needed, which is a fixed amount of time that all train movements are planned ahead; all train movements that are to start between the current simulation time and the current simulation time plus the safety margin are made a reservation of infrastructure capacity, thus guaranteeing their arrival time.

9.4.3. Loading and unloading

Trains are handled different under different real-life conditions: In some cases there is a need to handle them as quickly as possible, and in other cases there is no need at all to hurry that much. This highly depends on the system that is situated behind the trains that are handled: In case of a marine Maasvlakte system load units may be moved as quickly as possible from the trains to the marine terminals, as the number of trains that is to be handled with the limited transhipment capacity is great. In case of a small inland rail terminal there may be no need at all to hurry in unloading trains as it is known that the trucks that are to pick up the load units will not

---

11This is only valid in case of planned claiming, as will be discussed shortly. In case of ad hoc claiming, this safety margin can not be used, and therefore there is no guarantee that a train will arrive in time.
arrive than after five hours only. These different operational circumstances should be allowed to model for a proper analysis of these different types of systems.

One complicating factor when modelling these different operational circumstances and their strategies in detail is the large amount of data that is needed to do so. As seen in section "Train concepts" a number of operational parameters is available in the model that specify the sequence of handling of individual load units; but not the precise timing of these handle actions.

This section describes timing of handle actions of trains. It will be shown that quite some information is needed to model timing of trains properly, but it is considered as inevitable. Furthermore, it is shown in what way the numbers of load units to be handled by trains are specified in the model.

♦ Chosen approach

The approach chosen for in this model is to specify starting times of loading and unloading, and the corresponding ending times as fixed, known times. All load units to handle in this time span are equally spread over it, thus resulting in the precise system entry time and system exit time of each individual load unit that is handled at trains. These thus found system entry- and exit times of each individual load unit are used as a basis for the equipment optimisation model as described in Chapter 7 ("equipment hopping") and is respected as much as possible, but it is not guaranteed to be satisfied for a full hundred percent of accuracy.

This chosen approach results in a rather rigid timing behaviour of trains and their handling actions, but on the other hand it is believed to be a good handhold in timing all other train operations, some of which are shown in the example of figure 9.34, especially for complex train types as split shuttles and travel shuttles.

Another approach would have been to leave the optimal timing of loading and unloading trains up to the simulation system, thus finding automatically the optimal flows of trains through the system. But unfortunately, this would result in a clash with the equipment optimisation model as described in Chapter 7: This latter model focuses on finding the optimal use of equipment for any internal transport movement in a complex internal transport system, using the fixed and known times for system entry and -exit of each individual load unit as a basic input data. So, if scheduling of trains was to be optimised too, then it would not be known on beforehand at what time these load units are to be transhipped onto or from trains are to be handled. By consequence, these load units' system entry and -exit times could not get known in the normal way\textsuperscript{12} to the equipment optimisation model, and therefore could not be used in the equipment optimisation model. Although such a complicated coupling of two separate optimisation models is very challenging, it is not used for the moment. Further future expansions of the model may, however, include such a structure.

♦ Numbers of load units

The first step in defining loading and unloading procedures for trains is defining the numbers of load units to handle. In the model there is a strict relation between load schedules and load volumes at UCM level of detail in this respect. With respect to timing of load units, which will be discussed shortly, there is no relation at all between trains, load schedules and load volumes.

To each train service a number of load schedules can be assigned, each load schedule being related to one train handle area only. In this way it is automatically determined at which rail service centres the train is to be handled: At those rail service centres that are associated to those load schedules that are included by the train's service. This aspect especially is of importance for split shuttles and for travel shuttles: In this way it is determined which rail service centres are to be served by the shuttle trains.

The numbers of load units to handle are simply defined by the numbers of load units that are defined at each load schedule that is incorporated by the train service. In those cases where more than one load schedule that is incorporated in a train service refer to the same rail service centre, the total number of load units to handle at that rail service centre is the piece-wise sum of all these individual load schedules.

\textsuperscript{12} i.e. the way in which all other load units, that are transhipped at bounding elements rather than at train handle areas, are known to the equipment optimisation model.
Timing of load units

In the model timing of handling of load units is specified by a number of parameters that are part of each individual train service. In case of concurrent loading and unloading of the train two times as many parameters are needed to specify timing of load units as in case of concurrent loading and unloading: in case of separate loading and unloading both handle actions may start and end at different moments in time, loading never starting before unloading. In case of concurrent loading and unloading both handle actions start and end at the same moments in time.

Each handle action is timed by two parameters. In the model two different ways of specifying timing are allowed, both requiring two parameters only. In the first way following parameters are used:

- start of loading or unloading (fixed time);
- end of loading or unloading (fixed time),

which results in a very strict timing of handling of trains. Regardless of the number of load units they are carrying, handling is always to start and end at the same time. By consequence, sometimes trains are to carry more load units than usual and cause a high, peaked demand of transhipment and internal transport capacity. And sometimes trains are to carry less load units than usual, resulting in a very moderate demand for transhipment and internal transport capacity. This is the first way to model timing of load units; the second way to model timing is by using following two parameters:

- start of loading or unloading (fixed time);
- fixed, estimated transhipment capacity,

which results in an equal starting time for all trains, but for a different ending time, depending on the number of load units that is to be handled. The ending time of handling is simply determined as the start time plus the number of load units to handle times the fixed, estimated transhipment capacity. By using this option a more smooth and predictable usage of transhipment and internal transport capacity is obtained, as trains with a great number of load units are allowed a more long transhipment duration. Another consequence of using this option is the less predictable flow of train movements, as different trains of the same service are likely to finish transhipment at different moments.

The fixed, estimated transhipment capacity is a rough estimation of the transhipment capacity that will be available for handling the trains. Note that this estimated transhipment capacity is specified as an available capacity per train rather than per handle bundle. This may result in a difference between the actual, assigned transhipment capacity and the estimated available capacity.

Note that all timing aspects as defined at the load schedules that are incorporated in the train service are totally ignored at SDM and at DDM level of detail. Therefore, if there is no desire to simulate the system at UCM level of detail there is no need to specify any timing aspect at all at those load schedules and load volumes that are defined at train handle areas, thus reducing modelling effort at UCM level of detail.

Timing of load units - concurrent handling

In case of concurrent loading and unloading some special rules apply. First of all, the duration of the total handle action is found by multiplying the sum of the numbers of load units to load and to unload by the fixed, estimated transhipment capacity. Second, all load units for both handle directions are spread equally in time over this total duration, which may yield different net transhipment speeds for both handle directions in case of different numbers of load units to handle. This is clarified in figure 9.35 in which concurrent handling is depicted at a fixed, estimated transhipment capacity of 15 load units per hour. In case the train that is to be handled carries 12 load units for unloading and 4 load units for loading, then this total of 16 load units results in a total handle duration of (16/15)=1.07 hours. The net average speed at which load units are unloaded equals (12/16)*15=11.25 load units per hour, while this net average speed for loading equals (4/16)*15=3.75 load units per hour.

![Figure 9.35: Timing of load units in case of concurrent handling.](image-url)
Locations on-train
Combining with the options already discussed at section "Train concepts" on defining the positions of load units on trains, it can now be found at what time these load units are to be loaded or unloaded. In this way the full loading and unloading scheme of trains is defined.

Consider for instance a pure shuttle train that is to be handled concurrently, and that has different settings for the load units' positions on-train for loading and for unloading direction: Suppose that for loading direction load units are spread equally over the entire train and are to be handled in random sequence. Also suppose that for unloading direction load units are to be condensed in the head of the train and are to be loaded in "grazing" sequence. No sorting of load units on either clients or on destination is applied.\(^\text{[13]}\)

This situation is depicted in figure 9.36, which shows that in total four load units are to be unloaded, and that five load units are to be loaded. Supposing that in this case it has been chosen for to specify a fixed start- and end time of handling activities, than the time span available for transhipment has a fixed value, say 45 minutes. Therefore, within these 45 minutes all nine load units are to be handled in the sequence shown in figure 9.36.

If handling of the train is to start at 12:30 hours, then the times for system entry- and exit of the load units at the train's rail service centre are as shown in following table:

<table>
<thead>
<tr>
<th>Loading</th>
<th>Nr</th>
<th>System exit time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>12:30</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>12:41</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>12:52</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>13:03</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>13:15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unloading</th>
<th>Nr</th>
<th>System entry time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>12:30</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>12:45</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>13:00</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>13:15</td>
</tr>
</tbody>
</table>

Table 9.2: Example of timing of a pure shuttle's load units.

Train length
Another aspect that is of importance is the length of the train in case of a split shuttle: Of which length should the separate parts in which the entire train is split be? This length is determined as a fraction of the entire train's length, relative to the maximum of the number of load units that is to be handled at each rail service centre that is served by the split shuttle. So, if the split shuttle is to unload 5 load units at the first RSC, and is to load 1 and to load 15 at the second RSC, then the part that sets for the first RSC will be \((5/(5+15))\)=25% of the entire train's length, while the part that sets for the second RSC will be \((15/(5+15))\)=75% of the original train's length.

\(^{[13]}\)Note that this influences the handle times for individual load units, but does not influence the handle sequence of the entire train.
9.4.4. Bundle selection

Sometimes there is a need or desire to handle trains of specific services at particular rail bundles only. In other cases, it is desired to relocate all trains of a particular service to a special bundle that is reserved for that service only. This type of operational rules can be modelled too by some simple settings that can be defined at each individual train service.

For each rail service centre at which a train is to be handled, it can be specified whether some fixed bundle that is related to that particular rail service centre should always be selected for four different activities of any train of that service, when the train is present at that rail service centre. These four activities for which a fixed rail bundle can be specified are:

- handling;
- splitting;
- merging;
- relocation.

Only those bundles can be selected that are marked as fit for the various operations. In a number of cases there is no need to establish such a fixed bundle for a particular activity; in such cases the decision on the precise bundle to select is left to the rail service centre, as discussed at section "Bundle selection" at page 205.

9.5. Infrastructure

Infrastructure in the rail model at SDM and DDM level of detail is divided into three different types of elements: rail tracks, rail points and rail bundles. Especially the rail bundles play an important role for following reasons:

- several operational options can be specified at rail bundles;
- almost any rail movement starts or ends at a rail bundle;
- analysis of rail infrastructure after simulation heavily relies on the subdivision in rail bundles. This is a consequence of the above two aspects.

This section will handle these aspects along with a discussion of the various point types and their impact on the performance of any rail service centre.

9.5.1. Rail bundles

Most important of all rail infrastructure elements are rail bundles. As mentioned in the very beginning of this chapter rail bundles are defined as "a number of rail tracks that are functionally related and need to be considered as a unity". This definition will be clear after the various parameters defining the allowed operations for any rail bundle have been discussed.

Each rail bundle in the model can be given any of following functions:

- handling, i.e. transhipment of load units;
- splitting, i.e. splitting and merging of split shuttles;
- D-loc storage, i.e. storage of D-locs when not coupled and temporarily not needed for further operations;
- E-loc storage, i.e. storage of E-locs when not coupled and temporarily not needed for further operations;
- relocation, i.e. relocation (1) of trains between unloading and loading or (2) of travel shuttles between two subsequent rail service centres.

Furthermore, a bundle at which handling is allowed requires another parameter, which is a reference to another bundle at which trains are to be relocated after loading or unloading. Whether or not this fixed reference is used for relocation of trains depends on further settings of the train's train service, and of the bundle's rail service centre. This aspect will be discussed in detail in section "Bundle selection" of this chapter.

By using these parameters for rail bundles the flow of trains and light engines over the modelled rail infrastructure can be controlled in a rather detailed way without the need to explicitly mark for each train at what bundles it is to relocate, or where it is to be handled. In those cases where a rail service centre contains more than one bundle at which certain functions can be performed there is a need to select the best out of the available bundles. For this purpose some standard rules and a two- or three-level assignment structure is included in the model that will be described in section "Bundle selection" of this chapter, too.
Chapter 9. Railway aspects

In most situations it will be advisable to assign one or two (small) rail bundles to D- and E-loc storage only, to get a good insight in the numbers of light engines that are idling during the course of the simulation. Especially for E-locs it is interesting to assign one bundle to E-loc storage only, thus obtaining a good insight in the presence of light E-engines in the system in time, and therefore in the presence of trains that have uncoupled their E-locs. On the other hand, assigning a rail bundle to D-loc storage only will give a good and instant insight in the time-pattern of the utilisation of D-locs.

♦ Rail bundles - ownership

Both in reality and in the model, most bundles will be related to (or "owned by") a particular rail service centre that is included in the system under study; only trains that are served for that particular rail service centre may be handled at those bundles. This aspect of "ownership" of bundles is included in the model.

In the example in figure 9.37 two yard bundles and three handle bundles are shown, marked with a gantry crane. Two of these handle bundles seem to be related to the "First RSC" and the third one appears to be related to the "Second RSC". Because of this fixed relation between bundles and rail service centres it can be controlled in a simple way at which set of rail bundles trains can be handled: They can be handled only at those bundles that are "owned" by the rail service centre at which they are to be handled. For travel shuttles and split shuttles several handle bundles may be valid.

In the model it can be selected for each single rail bundle whether it should be related to some rail service centre or not. If it is not related to any rail service centre then handling is not allowed at that bundle, as it would be ambiguous for which rail service centre the load units that are transported by the train are to be unloaded or loaded. By consequence, for these bundles that are not owned by a rail service centre, there is no need to specify a fixed bundle to which train should be relocated after unloading, as trains can not be unloaded there. By consequence, only non-handle bundles (i.e. yard bundles) can be without an owning rail service centre.

♦ Rail bundles - differences between SDM and DDM

Rail bundles are modelled with a different aim at SDM and at DDM level of detail: At SDM level of detail they are modelled to find out during simulation what number of tracks should be installed at them, while guaranteeing a required service level. For this reason bundles at SDM level consist of an unknown number of tracks -- this number of tracks is to be determined after simulation on basis of the analysis results.

Figure 9.38 gives an example of a bundle at SDM level of detail. It shows how the bundle actually is spanned between two rail points, one on H- and one on V-side of the bundle. For each of these two points it may be chosen what external connections they allow: On track, two tracks or no track at all. In the latter case the bundle will not be connected at that side to the remaining rail network.

At DDM level of detail rail bundles are modelled in more detail, by a number of separate, individual tracks and points. Because of this more detailed modelling of the rail infrastructure DDM level of detail allows the inclusion of a block safety system in the model, and to study its effects on the required rail infrastructure capacity.

9.5.2. Rail points

Rail points are included in the model at a meso level of detail: All allowed transitions of trains over the points are modelled, but their precise dimensions and maximum train speeds are ignored. Even more, there are no checks that the modelled infrastructure is geographically one hundred percent in correspondence with the rules.
that apply for this aspect. This simplification has been chosen for to keep focus on the logistical aspects of rail movements rather than on civil engineering aspects, in agreement with the Dutch Railways.\(^\text{14}\)

All points included in the model are divided into five groups, analogous to the real-life situation. Within each group a number of individual point types can be found, each point type having its own angles for transition and dimensions. These groups are:

- normal points;
- symmetrical points;
- full English points;
- half English points;
- cross points.

Points are modelled by seven parameters. Six of these are shown in figure 9.39. All seven parameters are:

- number of tracks at H-side;
- angle between track H-1 and horizontal;
- angle between track H-2 and horizontal;
- number of tracks at V-side;
- angle between track V-1 and horizontal;
- angle between track V-2 and horizontal;
- allowed transitions.

The most important aspect is the list of allowed transitions which defines in which directions trains can cross the point. The importance of this aspect will be shown by a small example.

<table>
<thead>
<tr>
<th>Point type</th>
<th>Number of tracks</th>
<th>Slants(^\text{15}) (ratios)</th>
<th>Transitions(^\text{16})</th>
<th>Graphical representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>2</td>
<td>1:34.7 1:9</td>
<td>H-1 - V-1</td>
<td>(\text{H-1} \rightarrow \text{V-1})</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1:20 1:9</td>
<td>H-1 - V-1</td>
<td>(\text{H-1} \rightarrow \text{V-1})</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1:12 1:9</td>
<td>H-1 - V-1</td>
<td>(\text{H-1} \rightarrow \text{V-1})</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1:9</td>
<td>H-2 - V-2</td>
<td>(\text{H-2} \rightarrow \text{V-2})</td>
</tr>
<tr>
<td>Symmetrical</td>
<td>2</td>
<td>1:15 1:9</td>
<td>H-2 - V-1</td>
<td>(\text{H-2} \rightarrow \text{V-2})</td>
</tr>
</tbody>
</table>

\textbf{Table 9.3:} \hspace{1cm} \textit{Examples of point types.}

Table 9.3 shows some examples of point types and their allowed transitions. Based on these point types (and their prices in the real-life world) a terminal’s yard could be designed as shown in figure 9.40\(^\text{17}\). The purpose of this structure is to allow train movements between bundles A and B on one hand, and bundles C and D (which is more remote than bundle C) on the other hand. When only normal points and cross points are used,

\(^{14}\)Further detailing of these aspects could be a good idea for further development of the terminal design aiding tool.

\(^{15}\)Some examples only.

\(^{16}\)Inverse direction is always allowed, too.

\(^{17}\)Note that point angles and -dimensions are not quite as they are in reality; for reasons of clarity these parameters are exaggerated somewhat.
Chapter 9. Railway aspects

This problem could have resulted in the solution shown in Figure 9.40. This solution requires in total five points of two types: Four normal points and one cross point.

An alternative solution is shown in Figure 9.41 in which the five points shown in Figure 9.40 are replaced by one single full English point. This single full English point allows all transitions that are required in this case, consuming less space than the previous solution did which in some cases can be a decisive factor for this latter one. Most likely, operational and initial costs of this second solution are less than those of the first solution, too, and therefore makes it even more attractive. (Note that this aspect of cost analysis is included in the model, too, and is available after simulation has ended.)

Nevertheless, the second solution does have one drawback: It does not allow concurrent movement between bundles A and C on one hand, and between bundles B and D on the other hand, thus reducing the potential capacity of the rail infrastructure. Whether this restriction is of importance for the modelled terminal fully depends on the timing of movements of trains and light engines. If these movements can be timed such that they never need to occur simultaneously, this second solution will be fine. If, on the other hand, it is likely that movements from A to C and from B to D are to occur simultaneously, then the first solution may turn out to be better, resulting in a better performance of the terminal’s design. An example of such a simultaneous movement is depicted in Figure 9.43.

Note that these latter effects only show clearly through at DDM level of detail, as only at that level block safety systems are included in the model. Without such a system, that is used for infrastructure claiming any number of trains can drive simultaneously at one track or point.

9.5.3. Rail tracks

Rail tracks outside rail bundles are the least important of all three types of infrastructure elements: Rail points, rail tracks and rail bundles. In this model rail tracks are only used as connections between points and bundles, providing a “skeleton” to define rail routes between these locations.

There is, however, one important group of rail tracks. These are the rail tracks that are referred to by train services their tracks for system entry and their tracks for system exit. These are the tracks at which the trains "appear" according to their time schemes, popping in from the public rail network. Each train service may have a different track for system entry and for system exit, thus modelling different "gates" of the terminal’s yard to the public network. But in most cases only one such a "gate" to the public rail network will be modelled.

Analogous to rail points, at SDM level of detail any number of trains can drive simultaneously at one track because of a lack of a block safety system at that level of detail. At DDM level of detail such a block safety system is included in the model.
9.5.4. Rail routes

Based on the modelled rail infrastructure a number of *rail routes* can be defined. These rail routes simply define the possible routings of trains through the modelled rail yard that can be used during simulation for train movements. Rail routes are defined between two rail tracks rather than between two rail bundles, as it must also be possible to define routes between bundles and the train services' tracks for system entry and for system exit. Although it might be possible to create these rail routes automatically, based on the infrastructure, this has not been chosen for: Each individual rail route must be pointed out by hand, or must be composed out of a number of already existing rail routes. This is done to keep optimal control on the routings that are defined in the model.

An example of a typical rail route is shown in figure 9.44 in which a train is moved from bundle B to bundle C by driving over bundle A. This remarkable movement is a consequence of the limited freedom of movement on rail infrastructure. A generally accepted name for such a movement is a "sawing" movement. These movements are a further impediment in finding automatically all rail routes in the system: Sawing movements are hard to detect automatically in the modelled infrastructure, and not all possible sawing movements are allowed.

*During simulation for each train movement a fit train route is to be selected. This selection is done by the so-called route manager that can assign one of the available routes according to one of following selection criteria:*

- assign route with shortest distance;
- assign route with shortest time, i.e. the earliest time of arrival at the train's destination;
- assign route first encountered;
- assign route random.

There is no difference between the first two route selection methods at SDM level of detail, as there is no need to claim infrastructure. Instead, at SDM level a train can start moving at any time, independent of the number of trains that is already moving at its route's tracks. At DDM level of detail, one route may result in a longer driving distance than the other, but nevertheless may cause the train to arrive earlier at its destination. This is a result of the inclusion of infrastructure claiming and of block safety systems.

9.5.5. Block safety systems

Block safety systems have been introduced in real-life rail networks to improve safety of rail movements. This holds for public rail networks and for most rail terminal yards. Basic characteristic of block safety systems is the division of the entire rail infrastructure in a number of *sections*\(^1\) that are individually electrical isolated. Through this isolation it is possible to detect the presence of (part of) a train by means of a low-voltage current that flows via the train's axis between both rails of the section. As will be shown in this section, these block safety systems may have a high impact on the infrastructure usage. For this reason they are included in the model when performing detailed analysis at DDM level of detail.

This section gives a short overview of the most important elements that play a role in block safety systems at railway yards, after which several ways of using these block safety systems and their impact on rail infrastructure usage is discussed. Stress is put on the differences between block safety systems at railway yards, and block safety systems at public rail networks\(^2\).

Important is the difference between yard bundles and handle bundles. Generally block safety systems are not installed at all at handle bundles as normally they are not used for bypassing trains as yard bundles may:

---

\(^1\)Not to be confused with *sections* for equipment at found at equipment routings.

\(^2\)For information on block safety system for public rail networks, refer to Dijkstra [7], Kreutzerber [15] and Van Marle [20], the first and latter also giving a brief description of future developments in this field in the Netherlands, Germany and France.
Normally trains only arrive at handle bundles to be transhipped, and therefore are to stop at a track of the handle bundle and consequently drive at very low speed.

As just mentioned rail infrastructure is divided into individual sections (or rail sections in this thesis to discriminate from sections in equipment routings). In the real-life world these rail sections always start and end at a weld (or insulated joint) above which a signal may, but does not necessarily need to, be placed. In the model described in this thesis sections only start at signals. This is clarified in figure 9.45 which also shows three different modes for visibility of signals. Visibility of signals plays an important role for the definition of rail paths.

Rail paths are defined both in real-life world and in this thesis as a set of subsequent rail sections in a rail route that can be driven without encountering a signal facing the driving direction. This is clarified in figure 9.46 in which three signals, four rail sections and three rail paths are shown. It can be seen from this figure how the direction of visibility of the signals determines which sections are incorporated in a rail path.

Based on these rail paths part of the rail infrastructure can be reserved (or claimed) for a particular train movement. Each section in a rail path is claimed by a train at the same moment, but is released at a different moment as shown in figure 9.47. This figure depicts the moments at which all four rail sections in all three rail paths from figure 9.46 are claimed and released.

The general rule for claiming infrastructure is that an entire rail path must be claimed from the moment on that the first axis of the train (generally the pulling locomotive) arrives at the first rail section in the rail path. After that, each individual rail section in the rail path is released as soon as the last axis of the train (generally the last wagon) leaves the section.

---

20 Signals at rail yards are different from those at public rail networks: they are triangular and close to the ground, which is feasible because of low driving speeds at rail yards. The colour system for signalling is equal to that of signals at public rail networks.

21 The role of rail paths is much greater at rail yards than at public rail networks. This is caused by the much more complex configuration of rail yards, at which numerous points are concentrated in a relative small area, thus creating a situation in which several rail routes intersect and therefore can all be blocked at once by a single train movement (which, even worse, is to occur at a low driving speed). The main factor that determines that capacity of public rail networks, on the other hand, is the block length, in combination with the mix of trains (possibly driving at different speeds) that uses the rail infrastructure. Refer to Kreutzerberger [15] and Van Marie [20].

22 Or normally some time before.

23 In contrast to public rail networks, the minimum "dead time" between two train movements at public rail networks is very small, which is a consequence of the low driving speeds at railway yards.

24 Even though the actual block safety system (as used in the Netherlands) discerns three colours for signals (red, green and yellow), the model only discerns two colours (red and green). This simplification is allowed as the function of the third colour, yellow, is to command
From this general description two effects shown in figure 9.47 can be explained: (1) There always is some overlap in claiming of two subsequent rail paths as the last axis of a train will be still at some section the train's previous rail path at the moment the first axis of the train enters the first rail section of the next rail path. (2) All rail sections in a rail path are claimed from the same moment on, but are released at different times.

For train movements during simulation of the model at DDM level of detail each train and each single light machine must claim infrastructure before it can start driving. As rail infrastructure capacity is limited this will inevitably cause delays in railway operations thus influencing all train operations. From reality a number of methods for claiming infrastructure can be found. These methods differ (1) in the moment claiming starts and (2) the number of subsequent rail paths that is claimed. Their names as defined in this thesis are:

1. ad hoc, stepwise claiming;
2. ad hoc, integral claiming;
3. planned, stepwise claiming;
4. planned, integral claiming.

All of these claiming methods will be discussed in more detail as will their specific characteristics and field of application.

As can be seen from the above list all four infrastructure claiming methods consist of one out of two possibilities for the start of claiming (ad hoc and planned) and of one out of two possibilities for rail path sequencing (stepwise and integral).

**Ad hoc claiming**

As the name ad hoc claiming as a contrast to planned claiming suggests, infrastructure is, in case of ad hoc claiming, not claimed until the moment the train is originally planned to start driving. Generally, ad hoc claiming will induce additional waiting times as there is no guarantee that a valid set of rail paths can be claimed at the moment the train actually wants to start moving. Therefore some planning safety margin will be used that makes the trains start moving earlier than strictly needed, and therefore start claiming infrastructure earlier than strictly needed. This results in the effect that in those cases a valid set of rail path is immediately available after the train's request the train will start driving a little earlier than originally planned. This more or less unpredictable behaviour is characteristic for train movements at relative small rail yards that lack a highly sophisticated infrastructure planning system. Even for many bigger rail yards as found at rail terminals this way of claiming comes near to practice.

**Planned claiming**

In contrast to ad hoc claiming, planned claiming searches for a valid set of rail paths for the next train movement at the very moment the previous train movement has ended. In this way it is guaranteed that the train can actually depart at the time it was planned to leave. The basic assumption for this type of infrastructure claiming is that the next destination of the train is already known at the moment the train arrives at its current destination. Occasionally this assumption is not true, especially in those cases a train's next destination is a handle bundle where it is to be unloaded; whenever this handle bundle is to be determined dynamically during simulation based on momentary occupancies of all handle bundles it can not be determined beforehand and automatically ad hoc claiming will take effect.

Planned claiming is used in case of public rail networks or in case of large, complex railway yards. As planned claiming is not needed (and too expensive) for small railway yards it will rarely be used for rail terminals.

---

the drivers to drive "at sight, and prepare to stop at the next signal" - which is useful at public rail networks with high speeds (at which the driver should reduce his driving speed to 40 km/hr after a yellow signal, in the Netherlands), but not at railway yards at which driving speeds normally are very low.
Stepwise claiming
In case of stepwise claiming any set of subsequent rail paths is valid for a train movement, even if it does not guarantee a smooth, undisturbed train movement without intermediate stops. So, in case of stepwise claiming a set of subsequent rail paths is valid even if the train is forced to stop and wait before a next rail path can be entered. An example of such an intermediate stop is shown in figure 9.48 which is based on figure 9.47.

The entire set of rail paths has been claimed and reserved for the train before the train leaves. Therefore the intermediate stops are also planned and taken into account when claiming infrastructure capacity, thus avoiding deadlocks of trains moving in the opposite direction over the same sections. In case of the example of figure 9.48 the rail section before signal 3 is planned to be claimed longer than shown in figure 9.47.

![Diagram explaining stepwise claiming](image)

Figure 9.48: *Delay due to intermediate stop in case of stepwise claiming (based on figure 9.47).*

Stepwise claiming can be used for large but not centrally controlled railway yards. Generally it can not be used for public rail networks.

Integral claiming
In case of integral claiming the train is not allowed to have intermediate stops. In some cases this may cause additional delay to the desired moment of departure of the train as there is less chance to find a continuous path without interruption through a complex and busily used rail network. Nevertheless, integral claiming will be the normal claiming method for railway yards.

Selection of one out of the four claiming mechanisms has effects on the required infrastructure capacity and on delays of trains. Especially integral claiming in combination with high demands on smooth rail operations may yield an increased demand for rail infrastructure capacity. Analogous, the subdivision of the rail infrastructure into rail sections and rail paths with signals highly influences the need for rail infrastructure capacity. Therefore this task is to be done by a person (probably pragmatic) well known with the problems encountered at rail infrastructure definition.

Effects of block safety systems on required infrastructure capacity
The effects of block safety systems on the required infrastructure capacity is considerable, especially in those cases of high rail traffic intensity. Also, the precise subdivision of the infrastructure into sections is of great importance, as is the method used to claim infrastructure. This will be illustrated by some examples.

---

25Claiming of infrastructure has been programmed in TORCMOD but can not be activated and used yet. During this programming some special rules were developed to handle uncertainties in times of departure of trains -- train which could be blocking rail paths for future movements of other trains. To cope with this problem a train was polled to return an estimated time of departure (ETD) which was respected by the claiming mechanism. After that reported ETD any claim (which was still "soft") of another train could be planned. Before that latter train actually started to move, all (still "soft") claims were checked for validity by asking all trains on the claimed infrastructure to extend their ETD. If all trains reported to still agree with their original ETD all claims were locked (made "hard") and the actual train movement could start. If, unfortunately, after that moment any train present at the now "hard" claimed infrastructure did not manage to leave before its reported ETD and could not find a way out of the claimed infrastructure anymore (before being overdriven by the just departed train) then this train received the status of phantom train after which it became "invisible" for the infrastructure claiming mechanism. Such a phantom train was to be reported as an alarming simulation result, indicating a serious problem in the modelled railway system, which was expected to be corrected by the user of the model.
Figure 9.49 shows a simple configuration in which two parallel tracks are connected by two normal points, which are shown in thick lines. Supposing that both tracks are bi-directional (trains can drive in both directions over the tracks) the shown location of signals, which are the only elements that may indicate the start of a new section in this model, result in a rather large section that covers both points and part of both tracks as shown by the dotted lines. The implication of this configuration is, that all traffic over the upper track temporarily blocks all traffic over the lower track, irrespective of the actual use of both points. An improved situation would therefore be as shown in figure 9.50.

Figure 9.50 shows one signal positioned at the middle of both points, thus creating two independent sections.26 The advantage of this construction is that there is no interference between train movements over both parallel tracks as long as there is no change of track, thus increasing the capacity of the rail infrastructure. Obviously, any train changing tracks over both points (and the signal of figure 9.50) still must claim both parallel tracks at one time, which is quite logical.

The claiming method used is another important factor that influences the capacity of the rail infrastructure. Especially integral claiming reduces the number of train movements that can be made on the available infrastructure as all sections in the route are claimed and held from the beginning of the train movement on. Whether or not the advantage of more smooth and simpler train movements pay off against the reduction of capacity depends on numerous factors like prices of additional infrastructure, intensity of rail traffic and available space. No general conclusions can be drawn at this point.

9.6. Further operational options

For simulation of the rail bound model that has been described until now some additional settings are required, both at SDM and at DDM level of detail. These settings control selections that will be made by the model and special actions that facilitate analysis and optimisation of the modelled system. All these parameters are discussed in this section.

9.6.1. Bundle selection

Important for the behaviour of the simulated railway system is the way rail bundles are selected for various activities of a particular train, such as splitting or relocation between unloading and loading. The selection method used in this model has a rather hierarchic structure in which a valid bundle to use may be returned by one out of following types of elements, commonly in this same order:

- train concept;

26 In an actual situation this would be implemented as an insulated joint, without a signal.
- train service;
- train handle area; rail service centre;
- rail bundle (for relocation of trains only).

This hierarchic structure will be shown by explaining gradually figure 9.51 in which a valid bundle is searched for splitting of a split shuttle. Other activities for which a bundle could be searched are handling of trains, merging of split shuttles, relocation between unloading and loading, relocation of travel shuttles between two subsequent rail service centres, D-loc storage, and E-loc storage.

<table>
<thead>
<tr>
<th>Decision level</th>
<th>Question: Split at which bundle?</th>
<th>Returned answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train concept</td>
<td>Can determine?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>First stop</td>
</tr>
<tr>
<td></td>
<td></td>
<td>First RSC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No, ask service</td>
</tr>
<tr>
<td>Train service</td>
<td>Can determine?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed bundle</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No, ask RSC</td>
</tr>
<tr>
<td>Rail service centre</td>
<td></td>
<td>Lowest average utilisation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lowest momentary utilisation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed bundle</td>
</tr>
<tr>
<td></td>
<td></td>
<td>User's selection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Random</td>
</tr>
</tbody>
</table>

Figure 9.51: Example of a hierarchic bundle selection.

By using this hierarchic approach it is possible to model almost any bundle assignment strategy that is found in practice.

**Selection at train concepts**

The first step in the hierarchic decision process as depicted in figure 9.51 is to request the train's associated train concept to return an advice on the bundle to select. As can be found from the section on train concepts, these concepts only specify a rail service centre rather than a specific rail bundle for following activities:

- splitting of a split shuttle;
- merging of a split shuttle;
- handling of a pure shuttle.

By consequence, whenever a train concept has returned a rail service centre to select for a particular activity, then this rail service centre needs to be polled for further details on the rail bundle to select. In figure 9.51 this structure is shown by the dotted line, indicating that the rail service centre that was selected by the train's train concept has to decide on the eventual rail bundle to split the train at. Obviously, this eventually selected bundle should be fit for splitting trains.

In some cases (i.e. the bundle to split a train and the bundle to merge a train) it may be specified that the bundle to select is determined by the train's service rather than by the train's concept. In that case the train concept can not return a valid bundle, and the second step in the hierarchic decision process is invoked.

**Selection at train services**

The second step in the hierarchic decision process, when the train's train concept has failed to return a satisfactory answer, is to poll the train's associated train service on the bundle to select. As discussed at the section on train services, these services may return a fixed rail bundle that is coupled to the rail service centre that is visited. In case of handling this "visited rail service centre" is simply the one at which the train is to be handled, but in case of splitting this "visited rail service centre" is the train's very first load schedule's rail service centre. A train service can specify a fixed bundle for one of following activities:

- handling of trains;
- relocation of trains between unloading and loading;
- relocation of travel shuttles between two subsequent rail service centres (same rule as previous one);
- splitting of split shuttles;
- merging of split shuttles.

All these fixed bundles can be different for each rail service centre at which the train is to be handled, which is of importance for split shuttles and travel shuttles, as well as for pure shuttles that carry load units for more than one rail service centre.

---

27 Which has the same decision process as relocation between unloading and loading.
For all these activities, and for all rail service centres to call upon by the train service, it may be specified that the service should "comply with the rail service centre's choice" and should thus invoke the third step in the hierarchic decision process rather than returning a fixed bundle.

**Selection at rail service centres**

The third step in the hierarchic decision process, when the train's service failed to return a valid bundle, is to poll the rail service centre at which the train is to be handled for a valid bundle reference. In some cases this step in the decision process may be executed *directly* after the first step upon the results of the train's concept decision rules.

A rail service centre may return an answer on the bundle to use for following activities:

- handling of trains;
- relocation of trains between unloading and loading;
- relocation of travel shuttles between two subsequent rail service centres (same rule as previous one);
- splitting of split shuttles;
- merging of split shuttles;
- E-loc storage;
- E-loc storage.

For all but one these activities the rail service centre *must* return a valid bundle as the final selection, as can be seen from figure 9.51 too. Only for the selection of a bundle for relocation the rail service centre may return no valid bundle as an answer, but instead invoke the fourth step in the hierarchic decision process.

Selection of bundles by rail service centres has a more dynamic character than the previously discussed selection by train services: While train services could refer to fixed bundle only, fixed per rail service centre, rail service centres may assign a rail bundle according to one out of following rules:

- lowest average utilisation, i.e. select the bundle that has the lowest average presence of trains until now;
- lowest present utilisation\(^\text{28}\), i.e. select the bundle that has the lowest current presence of trains;
- fixed bundle, i.e. select a fixed specified bundle;
- random, i.e. make a random selection out of all available bundles;
- user's selection, i.e. ask the user during simulation which bundle to select.

These rules apply for all just mentioned activities and can only return a bundle that is fit for the required activity. But for relocation only one additional selection rule may be selected, which is:

- bundle's selection.

This latter selection rule, which is unique for relocation only, asks the rail bundle at which the train that requested a relocation bundle is currently being handled to specify another rail bundle to which the train must be relocated. This relocation bundle is a fixed referenced bundle of the train's current handle bundle as will be discussed shortly.

Important to note is, that these rules are dependent on train's associated train concept. So, for two different trains of two different concepts that are handled at the same rail service centre, two different selection rules may apply.

**Selection at rail bundles**

The fourth step in the hierarchic decision process is only used for the selection of a bundle to relocate a train on after handling. This bundle may be selected either for a normal train between unloading and loading or for a travel shuttle between two subsequent rail service centres. As can be seen from figure 9.52 such a fixed reference is defined at a handle bundle only; each handle bundle refers to exactly one bundle that is fit for train relocation, which does not necessarily need to be owned by the same rail service centre as the handle bundle itself. This fourth step is never reached in the example of figure 9.51 as it concerned the selection of a bundle for splitting rather than for train relocation.

---

\(^{28}\) Note that, in order to determine this present utilisation, all trains that already have set off for that bundle but have not arrived yet, are to be included in the utilisation, too.
In contrast to selection at rail service centres, selection of bundles at rail bundles does not distinguish between different train concepts: Each train of each concept will be assigned the same relocation bundle.

9.6.2. Locomotive controls

During simulation light engines play an important role in the rail service centres' operations. To complete these operations of locomotives some additional aspects need to be discussed, concerning the initial presence of locomotive, automatic creation of locomotives, and priorities to use either D- or E-locomotives. All these aspects are discussed in this section.

♦ Initial presence
At the moment simulation of the modelled system starts, a number of D-locs (which are intended for local movements only) are already present at the rail yards of the various rail service centres. These locomotives are said to be the initial locomotives. Any number of these initial locomotives may be created and placed at some rail track, anywhere in the modelled rail yards. Each locomotive may have a different initial position. Normally, all these initial locomotives will be D-locs. Nevertheless, it is possible to specify that some particular one must be E-locs instead.

♦ Automatic creation
During simulation new D-locs may be created automatically when needed for smooth operations at the rail terminal. This creation is controlled by the so-called loc-manager through three settings:
• whether it is allowed to create automatically new D-locs during simulation;
• the threshold for waiting trains to create new D-locs;
• the initial track for newly created D-locs.
If the loc-manager is allowed to create new D-locs during simulation then the creation threshold specifies the amount of time that trains must have been waiting for the assignment of a D-loc for a local movement before a new D-loc may be created. In this way it can be controlled that the rail terminal offers a certain level of internal service to the railway operations: It is guaranteed that no train will have to wait longer than the specified creation threshold for a D-loc, after which time a loc will be created dynamically. After simulation it is shown how many D-locs have been created for trains that have been waiting up to the creation threshold for the assignment of a locomotive, thus getting automatically a good impression of the total number of locomotives to install at the rail terminal.

♦ Locomotive bundle selection
A further setting of the loc-manager is the rule to select the bundle from which a locomotive for (local) movements must be obtained. Five settings for this selection rules are allowed:
• nearest in distance;
• nearest in time;
• highest concentration;
• lowest concentration;
• random.
At SDM level of detail the results of "nearest in distance" are equal to those of "nearest in time" because of the lack of a block safety system at that level. At DDM level of detail a more remote bundle may turn out to result in a faster assignment of a locomotive than a near one because of already planned other train movements.

When selecting to retrieve always locomotives from the bundle with the lowest concentration of locomotives the system will show a tendency to show a peaked presence of locomotives at loc-storage bundles: Some have a high concentration and others a definitely low concentration. When selecting the reverse selection method, i.e. to retrieve them from bundles with the highest concentration, a more flat pattern of the presence of locs will be obtained, in which all locs are more or less equally spread over all bundles. In this way the average driving distance will be reduced thus increasing the rail service centre's capacity.

♦ Locomotive priority
The final setting of the loc-manager is the priority that must be given to D-locs over E-locs for local movements. Or, when referring to table 9.1, the priority that must be given to D-locs for train movements that require the presence of a locomotive of "Any" type to start.
Two settings for the locomotive priority are allowed: (1) Whether to simply prefer D-locs over E-locs for local movements if both are available at the found bundle from which to obtain a locomotive, as just discussed, or (2) to totally exclude E-locs from usage for local movements. This latter option will be used most of the times, because of the required presence of a catenary at the entire yard if E-locs may be used for local movements, too.
Chapter 10. Application of the model for future situations at the Maasvlakte

10.1. Introduction

A theoretical terminal design model as described in Chapters 5 through 9 requires a number of examples to obtain a good insight in the way the various model elements are interrelated and can be used to model practical situations. The best way to study these examples is by performing hands-on analysis of several example cases with an implementation of the terminal design model. Unfortunately, this way is not always available and therefore this chapter provides a description of the use of the implementation of the terminal design model for some example cases.

These example cases will be based on the future situation at the Maasvlakte as described roughly in Chapters 2 and 3 of this thesis, and will cover most model elements described in those chapters. For each of these example cases two different aspects are described:

- model specification, which defines the example case\(^1\), and
- model analysis, which describes results on the behaviour of the example case, after simulation\(^2\).

To each of these aspects a different section is devoted. In total five example cases are defined and then analysed.

Both model specification and model analysis will be illustrated partially by referring to tables that contain information on input- and output data, and partially by using some screen-dumps of the currently available implementation of the terminal design model\(^3\). Especially this latter type of information will help to get a better insight in the coherence of the model elements.

\* Confidential information

As information on the future situation at the Maasvlakte is partially confidential, not all details on the input data can be given. This especially is true for data on train schemes and numbers of load units flowing through the system. Therefore in some cases the used data will be somewhat different from the actually predicted data for the Maasvlakte, without any further notice. Still, these various scenarios give a good impression of the effects they have on the outcome of the model.

10.2. Model specification

A large-scale terminal complex as the future Maasvlakte system is an ideal example situation for a theoretical terminal design model as described in this thesis as it includes all model elements. It includes, for instance, a number of marine terminals with "private" stacks, and a large rail service centre. Between all these terminals and stacks various flows of load units occur, each one having different priorities and timing characteristics. These flows are to be served by a wide variety of equipment pools of various types that have different service ranges: Some equipment pools are used only locally inside particular stacks, while other equipment pools are used for long-distance inter-terminal transport across the entire Maasvlakte system. Next to these different options for configuration of the Maasvlakte system different options in economic scenarios exist, implying different cases for the numbers of load units that are to cross the system's boundary of the modelled Maasvlakte system.

\(^1\) Data on model specification is also referred to as input data.
\(^2\) Data on model analysis is also referred to as output data.
\(^3\) Which is named TORCMOD. It is programmed in Borland Pascal 7.0 © for Windows™, is entirely object-oriented and comprises 91,199 lines of code. It allows fast switching between the three levels of detail by a single menu command and instant simulation of the modelled terminal complex. For simulation the Must scheduling algorithm is adapted, converted into an object-oriented environment.
Table 10.1 gives a survey of the different cases that are described in this section, and analysed in next. This survey is based on two characteristics of each example case: (1) the economic scenario and (2) the operational scenario. This latter operational scenario in its turn is split into two parts: A part describing the scenario for train operations and another part describing the scenario for equipment configuration. In the remainder of this section all scenarios will be described in detail. Note that not all cells in table 10.1 represent an example case that will be specified and analysed; only the five shaded cells actually do represent one.

In order to prevent false expectations it is stressed that currently only part of the terminal design method is fully implemented. In the ideal situation the implementation would cover both model specification and model analysis for all model elements. But at this moment only model specification is implemented for nearly all model elements, while simulation and analysis are implemented for rail operations at SDM level of detail only. Therefore, the different scenarios for equipment configuration will have no effect on the analysis results at this stage of implementation; they are solely discerned to illustrate the effects of these scenarios on the coherence between model elements.

### 10.2.1. Economic scenarios

In total three different economic scenarios are discerned:

1. GSM scenario,
2. Technique scenario, and
3. Forecast scenario.

All three scenarios have a different background, as will be discussed shortly, and define totally different numbers of load units that are to cross the system’s boundary of the modelled Maasvlakte system. By consequence these different economic scenarios will yield totally different results after analysis of the modelled system. No precise timetables of the various train scenarios will be given, but only the total number of shuttle trains that is to be handled each day.

1. The GSM scenario is based on a forecasting model developed and used by the port authorities of Rotterdam. This forecasting model is commonly known as Goederenstromenmodel and is widely accepted as an authoritative study of freight streams. Based on the results of this Goederenstromenmodel and some additional information on tonnage per container the number of trains in this scenario is established at 20 pure shuttles a day.

2. The Technique scenario is based on the so-called technical capacity of the future Maasvlakte system when fully developed. The technical capacity is defined as the total capacity of all future terminals under the assumption that these future terminals will be equipped with techniques as currently developed for the Delta-SeaLand terminal. Refer to Chapter 3 for a brief description hereof. The main difference between this and both other scenarios is that it is not based in a direct way on freight stream analysis, and therefore actually represents a kind of max-max situation; it is virtually impossible to handle more load units at the Maasvlakte than specified by the Technique scenario. The number of trains in this scenario is established at 46 pure shuttles a day.

3. The Forecast scenario is based on a model developed by the author of this thesis which is a refinement of the GSM scenario. The Forecast scenario includes freight stream analysis at the level of individual rail terminals and the effect of the introduction of shuttle trains on these freight streams. Refer to Van Zijderveel [37] and [36], and to Chapter 9 for a description of shuttle trains. The number of trains in this scenario is established at pure 18 shuttles a day.

---

4Simulation at UCM level of detail is implemented too, but system analysis is not for the moment. Therefore no further attention will be paid to UCM level of detail in this chapter.

5A full description of Goederenstromenmodel can be found in GSM6 [31].

6286 Different rail terminals in total.
10.2.2. Operational scenarios

All three operational scenarios incorporate two different aspects: (1) operations of trains and (2) operations of equipment. A survey of these operational scenarios is given in table 10.2 in which two equipment scenarios and three train scenarios are discerned. All of these partial scenarios will be described in more detail in this section. As simulation and analysis of trains is implemented at the time this thesis is written, but simulation and analysis of equipment is not, the results after the simulation will only show an influence of the different train scenarios.

<table>
<thead>
<tr>
<th>Operational scenario</th>
<th>Train scenario</th>
<th>Equipment scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>A</td>
<td>②</td>
</tr>
<tr>
<td>II</td>
<td>B</td>
<td>②</td>
</tr>
<tr>
<td>III</td>
<td>C</td>
<td>②</td>
</tr>
</tbody>
</table>

Table 10.2: Definition of operational scenarios.

- Train scenarios

Three train scenarios require further definition as shown in table 10.3 from which two different attitudes towards brake tests and train relocation can be seen. Scenario A will be used to determine the effects of performing brake tests on the train operations. As discussed in Chapter 9 these brake tests may become superfluous in future as a result of the shuttle train concept. Scenario B is used as a reference for both other scenarios, as operational scenario II is used as a reference for both other operational scenarios. Scenario C is used to determine the effects of relocation between unloading and loading for all trains instead of only part of the trains.

<table>
<thead>
<tr>
<th>Train scenario</th>
<th>Brake tests</th>
<th>Relocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>No</td>
<td>Partial</td>
</tr>
<tr>
<td>B</td>
<td>Yes</td>
<td>Partial</td>
</tr>
<tr>
<td>C</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 10.3: Definition of train scenarios.

As follows from table 10.3 ("partial" relocation) two different attitudes towards train relocation between unloading and loading are discerned in base scenarios A and B, yielding two different train concepts. The difference between these two train concepts is that one train concept requires train relocation while the other does not. The subdivision of trains into these two train concepts for all three economic scenarios is as shown in table 10.4.

<table>
<thead>
<tr>
<th>Economic scenario</th>
<th>Number of trains assigned to each concept</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Concept 1</td>
<td>Concept 2</td>
<td>Total</td>
</tr>
<tr>
<td>Do relocate</td>
<td>Do not relocate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSM</td>
<td>9</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>Technique</td>
<td>27</td>
<td>19</td>
<td>46</td>
</tr>
<tr>
<td>Forecast</td>
<td>7</td>
<td>11</td>
<td>18</td>
</tr>
</tbody>
</table>

Table 10.4: Subdivision of trains over both train concepts in train scenarios A and B.

Other characteristics that hold for all train scenarios are the compulsory use of D-locs for local movements, direct entrance of trains on the handle bundles, performance of the great brake test only, re-coupling to E-locs before the great brake test, and immediate release of D-locs that are no longer used for local movements. A limited number of trains is loaded and unloaded concurrently, while the majority is unloaded first and then loaded against

- Equipment scenarios

Both scenarios for equipment, ② and ③ respectively, will be explained by referring to figures 10.3 through 10.5. The first two figures show the Maasvlakte system modelled in TORCMOD in two views in which all marine terminals and their stacks can be recognised, along with their "private" stacks. These figures also show the networks for internal transport through the entire Maasvlakte system and the rail service centre with its handle- and yard bundles.

---

7 So train scenario B is adopted as the base scenario for train operations.
8 In those cases that brake tests are performed. The great brake test lasts 30 minutes in the example case.
9 If available, refer to the demonstration diskette that has been supplied with this thesis.
10 These scenarios for equipment do not affect the analysis results.
Equipment scenario 1 is shown in figure 10.4 and can be described as follows:

1. All transport between the various terminals (ITT) is done by means of one pool of AGVs. These AGVs do not enter the terminals themselves, but are always decoupled from transport at these terminals by means of an internal stack.

2. All transport at the terminals themselves is to occur either by means of "private" pools of AGVs (at marine terminals) or by means of "private" pools of MTS-trailers and trucks (at the rail service centre). It is assumed that these AGVs can only be used for local internal transport at the terminal at which they are installed, which means that even internal transport between to adjacent marine terminals is to occur by means of the ITT transport system.

3. All internal stacks are equipped either with automated stacking cranes (at marine stacks) or with manually operated gantry cranes (at the rail service centre).

Through the modelled system a number of closed tasks and equipment routings is defined which define the characteristics of equipment assignment to particular transport. An example of such equipment priorities, as stored at equipment routings, is given in figure 10.2.

Equipment scenario 2 is shown in figure 10.5; compared to equipment scenario 1 it is mainly different in the equipment pool that is used for ITT: It uses a multi-trailer system instead of a pool of AGVs. Although this has little consequences for the marine terminals, it does have major consequences for the rail service centre as the trailers of this pool are allowed to be loaded and unloaded directly by the gantry cranes installed over the trains. This means that the trailers are allowed to enter the private network of the rail service centre and implies that local stacking at the rail service centre is allowed, this in contrast to equipment scenario 1. By consequence there is no need for a separate internal stack at the rail service centre, which is therefore omitted in figure 10.5.

An example of a flow pattern specifying the preference for presence of load units at various locations for one closed task is given in figure 10.1. From this example it shows that load units flowing over this particular closed task have a high preference to reside at the internal stack at marine terminal "ECT Delta 1" while all other locations have a lower preference for presence. In this way the flow of load units through the modelled system is controlled.

---

11 Abbreviation of: Inter-Terminal Transport.
Figure 10.2: Example of an equipment routing.

An example of an equipment routing is given in figure 10.2. This clearly is part of equipment scenario 0, as internal transport between the stack of the rail service centre and the stack of marine terminal "ECT Delta 1" is covered by AGVs. Only one equipment network is available for this transport and only one equipment pool is installed at this equipment network. Therefore, there would be no effect for the setting of the preference to use this pool of equipment\textsuperscript{12} on the highlighted section, as the equipment optimisation model (when implemented) would be allowed to select this single pool only.

Figure 10.3: Example of the Maasvlakte system modelled in TORCMOD.

\textsuperscript{12}As shown in the lower-right bottom as "0. Indifferent".
Figure 10.4:  *Equipment scenario 1.*

Figure 10.5:  *Equipment scenario 2.*
10.3. Model analysis

Each scenario from Table 10.1 has been simulated for a 28-hour period. Some results of these various experiments are shown in Tables 10.5 and 10.6.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Number of D-locs</th>
<th>Number of yard tracks</th>
<th>Number of handle tracks</th>
<th>Number of yard tracks</th>
<th>Number of handle tracks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technique-I</td>
<td>4</td>
<td>24</td>
<td>16</td>
<td>31</td>
<td>19</td>
</tr>
<tr>
<td>Technique-II</td>
<td>4</td>
<td>24</td>
<td>16</td>
<td>31</td>
<td>19</td>
</tr>
<tr>
<td>Technique-III</td>
<td>5</td>
<td>24</td>
<td>16</td>
<td>30</td>
<td>22</td>
</tr>
<tr>
<td>GSM-II</td>
<td>2</td>
<td>7</td>
<td>7</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Forecast-II</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 10.5: Analysis results of example scenarios (part 1/2).

Table 10.5 shows the number of D-locs and the number of tracks to install at the rail terminal according to the various simulation runs. The number of required D-locs is based on a maximum waiting time of 15 minutes before assigning a D-loc to a train, while the number of required tracks is based both on a service level of 90 and of 100%.

Both service levels are used to indicate the influence of the "peakedness" of train movements at the modelled rail terminal: it shows that this influence is considerable and that therefore a considerable amount of investment is put in offering the last 10% of service to train movements. This might in a real-life situation be an indication that some changes should be considered in the internal time-table of the rail terminal, as well as in the external time-table.

When comparing the various test cases in Table 10.5 it is found that the effects of brake tests (operational scenario 1 versus operational scenario 2) is nil. Although possibly surprising, this merely is a logical consequence of the fact that
1. trains in all test cases are to drive from the public network immediately onto the handle bundles,
2. trains therefore wait at handle bundles until their planned system exit time, and
3. during this inevitable waiting perform break tests at the handle bundles.
By consequence, under these conditions no effects are to be expected from abstaining from brake tests. Omission of brake tests would, however, reduce the required number of handle tracks in those cases where (1) trains are to leave from yard bundles instead of from handle bundles, and (2) brake tests are to be performed immediately after handling finishes.

Table 10.5 shows that the total number of tracks to install at the terminal is less than the total number of trains in each scenario in 10.4, but not in a great extent. This is caused by the simultaneous presence of a substantial part of all trains, as can be seen from Figure 10.8 too.

Table 10.5 also shows the enormous impact of the various economic scenarios on the numbers of tracks to install, which is quite obvious, actually. For analysis of the GSM-II and Forecast II example case some modifications of the modelled terminal have been made to obtain a realistic situation: Initially, both example cases have been analysed while using five handle bundles and five yard bundles used for all three Technique example cases, and as shown in Figure 10.3. After analysis the number of tracks to install at all three bundles appeared to be thus low that it was decided to reduce the number of bundles in order to get a more realistic model. In this way the numbers of handle bundles and yard bundles were reduced to three for the GSM-II.

---

13 Refer to Chapters 5 and 9, and to Van Zijderveld [35] for more details on this service level.
14 The external time-table generally is very inflexible due to (1) restrictions imposed by the available rail paths at the connecting public rail network, and (2) marketing demands imposed by service to "customers" in the form of times of arrival and departure of trains.
15 Changes of this type require another, additional type of tool which offers the possibility to shift trains in time such that the planned and roughly estimated bundle track usage is shown immediately and thus can be smoothed by hand as far as possible. Such a "planning tableau" was originally intended to be implemented in TORCMOD as a highly powerful addition. Unfortunately time lacked to do so.
example case, and to two for the Forecast example case. This adaptation was simply done by changing the allowed abilities for two and three handle bundles respectively in both example cases; for those handle bundles that were to be eliminated it was specified that handling was not allowed any more. Through the strong interrelation between handle bundles and yard bundles in the example cases, as shown in figure 10.3, the corresponding yard bundles were eliminated from the simulation process, too.

When comparing the GSM-II scenario with the Forecast-II scenario it is particular to note that the two additional trains for the first scenario require two additional yard tracks while they do not have an influence on the number of handle tracks. This is mainly a consequence of the fact that both trains are relocated between unloading and loading and apparently must be handled at such times that they no not affect the number of handle bundles.

The increase of D-locomotive movements in case of the Technique scenarios is great and is mainly a consequence of the large number of trains that are to relocate between unloading and loading.

When comparing operational scenario II with operational scenario III it is found that the effects of train relocation between unloading and loading on the required number of tracks is limited, even though full relocation in this case appears to result in a more peaked, and therefore less appealing, usage of rail infrastructure. Note that these results are highly depending on the precise trains (i.e. which train services precisely) that are to relocate at operational scenario II. The effects of various alternatives at this point could have been determined more easily when time had been available to construct a "planning tableau" as mentioned just before.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Personnel costs\textsuperscript{16}</th>
<th>Number of trips</th>
<th>Percentage of idling time</th>
<th>Percentage of driving time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Uncoupled</td>
<td>Coupled</td>
</tr>
<tr>
<td>Technique-I</td>
<td>13440</td>
<td>152</td>
<td>68.2</td>
<td>13.1</td>
</tr>
<tr>
<td>Technique-II</td>
<td>13440</td>
<td>152</td>
<td>68.2</td>
<td>13.1</td>
</tr>
<tr>
<td>Technique-III</td>
<td>16800</td>
<td>258</td>
<td>57.5</td>
<td>17.8</td>
</tr>
<tr>
<td>GSM-II</td>
<td>6720</td>
<td>50</td>
<td>79.2</td>
<td>8.8</td>
</tr>
<tr>
<td>Forecast-II</td>
<td>3360</td>
<td>28</td>
<td>88.4</td>
<td>5.5</td>
</tr>
</tbody>
</table>

Table 10.6: Analysis results of example scenarios (part 2/2).

Table 10.6 shows some details on the usage of D-locs and on the costs involved with this usage. All cost parameters are totally fictive and as shown in the screen-dump of figure 10.7. When comparing the number of trips of operational scenario II with those of operational scenario III it shows that the additional relocation of trains causes a substantial number of extra locomotive movements. When considering the total costs, the "advantage" of an increased average usage of locomotives is compensated by the costs of an additionally required locomotive (table 10.5).

\textsuperscript{16}During full 24-hours simulation period.
Another type of analysis is shown in figure 10.8 which provides an overview of the course of all trains in the modelled system for the Tech-II scenario, thus giving a quick impression of the operations at the modelled terminal complex. For this course analysis a set of activities has been chosen to show, such that a complete insight in the most important activities is obtained. Figure 10.9 shows the same activities for one particular train only, thus providing a good insight in the course and actions of that single train. This kind of detailed analysis is useful to control proper model functioning on one hand and to improve insight in processes at rail terminals.

Complementary to analysis as shown in figures 10.8 and 10.9 which considers the course of activities without a direct relation to location of these activities (i.e. bundles or tracks) is another kind of analysis that considers the presence of trains at one particular location only (i.e. bundle or track) which results in data on required infrastructure capacity as shown in table 10.5. These two analysis approaches are both required to obtain a full picture of the characteristics and behaviour of the modelled rail terminal. Figure 10.6 shows their coherence.

---

17 The lower seven graphs from figures 10.8 and 10.9 can show any out of twenty-six different activities. Six pre-formatted sets of these seven activities are available for quick analysis.
Chapter 10. Application of the model for future situations at the Maasvlakte

♦ Speed of simulation
The duration of the simulation for all example cases at SDM level of detail is shown in table 10.7. These time spans, covering a simulation period of 28 hours, can be called very acceptable. The simulation performed only includes rail operations at SDM level of detail; it does not include internal transport nor equipment optimisation according to equipment hopping. Especially this latter aspect is expected to slow down the simulation process considerably, when implemented.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Duration of simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technique-I</td>
<td>28 sec</td>
</tr>
<tr>
<td>Technique-II</td>
<td>28 sec</td>
</tr>
<tr>
<td>Technique-III</td>
<td>33 sec</td>
</tr>
<tr>
<td>GSM-II</td>
<td>&lt; 10 sec</td>
</tr>
<tr>
<td>Forecast-II</td>
<td>&lt; 10 sec</td>
</tr>
</tbody>
</table>

Table 10.7: Duration of simulation.

♦ Conclusions
The effects of omitting brake tests is nil in case trains are to leave from handle bundles and their system exit times are not altered. The additional costs to handle the last 10% of all train movements are relatively high, such that alterations in the internal and external time-tables (for as far as allowed) are attractive to study. Relocation of all trains is, under the given conditions, not attractive as it results in a more peaked rail movement behaviour and the need for one extra locomotive at the terminal.

♦ Model validation
Important and in the meantime difficult is validation of a model as described in this thesis. It is important as it must be known whether the results produced by the model are valid, and it is difficult as it is hard to obtain reliable information on real-life terminal systems. Even more, because of the only partial implementation of the model, proper validation is not possible except for rail operations at SDM level of detail. Although time lacked to gather information on existing rail service centres, there was still the option to compare the results of TORCMOD with the results of other, accessible studies on rail service centres. Therefore, it has been chosen to validate the model by referring to a study made by Tebodin consulting engineers in 1993 concerning the design of the future Maasvlakte rail service centre.

As this Tebodin study is considered confidential and therefore can not be described in this thesis, it must be trusted by the reader that this comparison was made with care. It was found that the results produced by the model were in good correspondence with the results of the Tebodin study. Based on this comparison it could be stated that the simulation results produced by TORCMOD are reliable, this statement also being supported by numerous other experiments with small, manageable systems.

---

18 On a 486-66DX2, 8Mb, LB machine under MS-Windows 3.1™.
19 Refer to Chapter 4.
20 Which scope was limited to rail infrastructure only.
Figure 10.8: Train course analysis of all trains in scenario Tech-II.
Figure 10.9: Train course analysis of one particular train in scenario Tech-II.
Conclusions and recommendations

This thesis gives a general structure to model and analyse a wide variety of terminal types for transhipment of standard load units. One of the main characteristics of this structure is its capability to model and analyse terminal complexes, comprising multiple terminals of various types that are interrelated through a mutual internal transport system. Through this interdependence of all terminals in the modelled complex, special measures are needed for a smooth operation of all these terminals. The general structure described in this thesis provides some mechanisms to facilitate modelling, optimisation and analysis of this interdependence. This structure is based on model elements called load schedules, contra schedules, equipment routings and flow patterns. Due to the only partial implementation of this structure, it is hard to draw conclusions on the validity of these model elements.

Contemporary terminal design is characterised by extensive calculations by hand, supplemented with simulation analysis. For each simulation analysis a new simulation model is created and implemented. Two major drawbacks hereof and of the calculations by hand are:
1. the great amount of time and assets involved with analysis of a terminal complex, and
2. the small number of alternative scenarios for terminal design that are analysed, which is a mere consequence of the first drawback. In some cases this may yield a sub-optimal terminal design.

The general structure described in this thesis may help to eliminate these drawbacks to some extent.

As a consequence of the contemporary limited possibilities to analyse different scenarios for terminal design in a short time very little experiments with alternative designs are made, while particularly this type of analysis increases understanding of, and insight in, the processes that are found at terminal complexes. It is this lack of understanding of the processes at terminal complexes that impedes further development of new terminal layouts and operational strategies. Implementing and using the model described in this thesis may be a first step on a way to more intelligent and sophisticated design tools and design methods for terminal complexes.

Further development of alike terminal design methods and design tools is required, as this thesis only aims to provide a basis for further research. Further development should primarily focus on more sophisticated control rules for aspects like equipment assignment and flow planning, and on more detailed transhipment processes, especially for rail bound transhipment. In addition, during this further development the chances for a more "automated" approach of terminal design should be examined, an approach in which a computer model is inquired more regularly than nowadays for advice. But above all, rail operations at DDM level of detail should be implemented as to allow more detailed and accurate analysis of rail operations and of rail infrastructure usage.
Conclusions and recommendations
Appendix A. Decision techniques

A.1. Introduction

The success of real-life logistic systems strongly depends on the decisions that are made; wrong decisions can ruin a basically well-structured logistic system. The same holds for computer models: If the computer model's structure is fine, but the decisions are poor, then the value of the computer model's results is restricted. Moreover, the decisions made in a computer model should not only be good, i.e. be what they are supposed to be, but they also should come as near as possible to the decisions that would have been made in the real-life system.

This goal can be achieved by including three aspects of real-life decision making in the computer model. The first aspect is called human assessment, which is no more than translating human-related appraisals (which are semantic) into computer-related appraisals (which are numerical). Several methods exist. Section A.2 describes some of these methods and indicates which method has been selected as the best. The second aspect is called multi criteria analysis (MCA), which is a technique to combine for each alternative its scores on various criteria to one overall score; this is further explained in section A.3 and in the definitions below. Several methods for MCA exist. Section A.3 describes some of these methods, and indicates which method has been selected as the best. The third aspect is called random selection, which is a technique to pick the "best" choice from a set of alternatives, imitating human behaviour. The importance of random selection cannot be stressed enough; it has a major impact on the final decision for an alternative. Section A.4 describes how random selection works and what options are available.

Definitions

In the remainder of this Appendix following definitions will be used:

- proposal = element that must be matched with one of the available alternatives
- alternative = element that can be assigned to the proposal; it can be either chosen for, or rejected
- criterion = attribute or aspect of an alternative that should be taken in consideration when choosing pro or contra assigning the alternative to the proposal
- score = valuation of a criterion of an alternative for the current proposal
- final score = overall score of an alternative; a combination (by means of MCA) of all scores of the alternative on all relevant criteria (possibly means of human assessment).
- choice = chosen alternative from a limited set of alternatives of one type, based on their final scores. The choice is assigned to the proposal.

An example will clarify these definitions. Consider a situation where a container must be stored in a stack that is divided into ten storage blocks. Our decision technique must determine which of the ten blocks must be selected.

In this case, the proposal is the container, including all relevant data for the current decision. This data may specify, for instance, its type, its owner, and its colour. The alternatives are all ten storage blocks. All these ten storage blocks may be selected based on the container's type, the container's owner, and the storage block's number of free slots; the container's colour is of no interest.

The scores are the valuations for the container on all three criteria, at all ten storage blocks; each storage block may have its own preference to store a container of, say, owner "Miersk", and may have its own preference to store any container, depending on its current number of free slots. Some of the scores will be determined using human appraisal; for instance, the second storage block may have a "Weak positive" valuation of any container belonging to "Miersk".

The choice is the finally selected storage block, taking into account all individual scores, combining them by means of MCA.
A.2. Human assessment

Why using human assessment in a computer model, anyway? A computer model is called a "computer model" as it is supposed to model the real-life system. In this real-life system human beings normally play an important role; they decide what is to be done. This yields the conclusion that (1) the decision behaviour of a computer model should imitate that of human beings, and (2) the values of parameters influencing the decision behaviour should be specified in human language and not in computer language. A good computer model should meet both demands, and therefore should include human assessment.

♦ Method requirements

What requirements should a method modelling human assessment in a large-scale logistic model meet? First, the user of the model must be offered a limited set of words that he can use to express his appraisal for a specific criterion at an alternative. An example hereof is "Very good". Second, the behaviour of the method modelling human assessment must imitate human behaviour; for example, whenever an alternative scores strong negative ("Very bad") on one criterion, then the alternative should score average positive ("Good") on many other criteria, should the alternative ever be selected. Third, the method modelling human assessment should allow the model to make automatic repeated decisions, once all relevant decision parameters have been set.

The first requirement yields the conclusion that the method should offer a limited set of words expressing the user's appraisal on one criterion for the use of an alternative for a proposal. The number of available words must be balanced: On one hand the user must be offered enough nuances to model real-life system behaviour, and on the other hand the user must be protected from an unmanageable set of extremely slight nuances. Based on Otto [22] a number of nine different semantic expressions have been selected for use. Table A.1 shows their names.

The second requirement yields the conclusion that the values related to the semantic preference judgements showed in table A.1 must form a geometrical series. According to psychologists geometrical series result in a natural behaviour of human assessment methods. Other examples of geometrical series are the well-known scales for sound and light intensity. Therefore, in this thesis this point of view is adapted. Refer to Otto [22], page 18.

The third requirement yields the conclusion that the user should specify the parameters for the decision process (i.e. the valuations of all alternatives on all criteria) only once, after which the decision process does not need to prompt the user anymore to make a decision. For this reason only so-called absolute assessments can be used; relative assessments and interval assessments can not be used. This will be explained below.

Table A.1: Selected semantic preference judgements for human assessment.

<table>
<thead>
<tr>
<th></th>
<th>1. Very strong preference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2. Strong preference</td>
</tr>
<tr>
<td></td>
<td>3. Average preference</td>
</tr>
<tr>
<td></td>
<td>4. Weak preference</td>
</tr>
<tr>
<td></td>
<td>5. Indifference</td>
</tr>
<tr>
<td></td>
<td>6. Weak negative</td>
</tr>
<tr>
<td></td>
<td>7. Average negative</td>
</tr>
<tr>
<td></td>
<td>8. Strong negative</td>
</tr>
<tr>
<td></td>
<td>9. Very strong negative</td>
</tr>
</tbody>
</table>

♦ Available methods

Three different types of human assessment methods can be discerned: Relative assessment methods, absolute assessment methods and interval assessment methods. These three types will be described briefly. A more detailed description can be found in Otto [22] or Lootsma [19].

Relative assessment methods are based on a comparison of all alternatives on one criterion at the time, each time a decision must be made. This approach makes that relative assessment methods cannot be used in a computer model. Following example will clarify this.

Suppose that a new crane must be ordered, and that we must choose from three different types named A, B and C; these are the alternatives. One of the important aspects of a crane is maintenance; this is the criterion. If we use the relative AHP assessment method as defined by Saaty and described in Otto [22], page 15, then the decision process is as follows. (Some other relative assessment methods can be found in Voogd [32]).
In that case we need to compare all cranes (alternatives) on one aspect (criterion), using the semantic preference judgements and the associated values shown in table A.2. The ratio \( W_i / W_j \) is the relative score of alternative \( i \) on one criterion, compared to alternative \( j \). The value associated with this ratio is used to determine the total score of each alternative on one criterion.

Suppose that the relative scores of all three cranes on maintenance are as follows:
- A compared to B: Weak preference for B \( \Rightarrow W_B / W_A = 3 \)
- A compared to C: Little preference for C \( \Rightarrow W_C / W_A = 2 \)
- B compared to C: Weak preference for B \( \Rightarrow W_B / W_C = 3 \),

then we can construct a matrix showing the relative preferences for the alternatives on the criterion maintenance, as shown in formula (A.1),

\[
A = \begin{pmatrix}
W_A & W_A & W_A \\
W_B & W_B & W_C \\
W_A & W_B & W_C \\
W_C & W_B & W_C \\
\end{pmatrix}
\]  
(A.1)

or, if we fill in the just found relative preferences of all three types of cranes on the criterion maintenance:

\[
A = \begin{pmatrix}
1 & 3 & 1 \\
3 & 1 & 3 \\
2 & 1 & 3 \\
\end{pmatrix}
\]  
(A.2)

Now something odd has happened: Although the relative preferences seemed to be just fine, matrix \( A \) in formula (A.2) turns out to be inconsistent. This inconsistency can easily be shown: Just fill in the appropriate values in the logical formula (A.3):

\[
\begin{pmatrix}
W_A \\
W_B \\
W_C \\
\end{pmatrix} = \begin{pmatrix}
1 & 3 & 1 \\
3 & 1 & 3 \\
2 & 1 & 3 \\
\end{pmatrix} \begin{pmatrix}
W_A \\
W_B \\
W_C \\
\end{pmatrix}
\]

and the result is the invalid expression:

\[
\frac{1}{3} = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}
\]

which proves that matrix \( A \) shown in formula (A.2) is inconsistent.

What is the reason of this inconsistency? There are two reasons. First, the relative preferences shown in formula (A.2) were obtained by converting semantic preference judgements into numeric values. These semantic preference judgements in their turn were obtained by interviewing a human being who had to decide in a rather "fuzzy" manner on his preferences for all three cranes. By consequence, there was no possibility at all to safeguard any consistency between the semantic preference judgements. Second, the values associated with the semantic preference judgements as shown in table A.2 are not consistent. A small example will clarify this. Suppose, for example, that crane A obtained weak preference over crane B, and that crane B in its turn obtained weak preference over crane C. Then crane A would automatically have obtained absolute preference over crane C. This is proven in formula (A.5). It will be clear that this behaviour of AHP is somewhat in contradiction with our expectations. This example yields the conclusion that any human assessment method should use geometrical series to be valuable.

\[
\frac{W_A}{W_C} = \frac{W_A}{W_B} \cdot \frac{W_B}{W_C} = 3 \cdot 3 = 9 = \text{Absolute preference}
\]

Can this inconsistency be avoided? No, it cannot be avoided, but there is a well-known technique to make sure that the inconsistency in matrix \( A \) is at an acceptable low level; it makes sure that the difference between the matrix \( A \) as given by the user of the decision technique corresponds sufficiently well with the unknown, but fully consistent matrix \( A' \). Therefore, the technique ensures that matrix \( A \) is a sufficiently accurate estimate of

<table>
<thead>
<tr>
<th>( W_i / W_j )</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Equal importance</td>
</tr>
<tr>
<td>3</td>
<td>Weak preference</td>
</tr>
<tr>
<td>5</td>
<td>Strong preference</td>
</tr>
<tr>
<td>7</td>
<td>Demonstrable preference</td>
</tr>
<tr>
<td>9</td>
<td>Absolute preference</td>
</tr>
</tbody>
</table>

Table A.2: Relative assessment scores for relative AHP
Appendix A. Decision techniques

the perfect matrix \( A' \). This method is called the Perron-Frobenius eigenvector method, or the Saaty eigenvector method.

How does the Perron-Frobenius eigenvector method work? Suppose that matrix \( A \) were consistent (and therefore equal to matrix \( A' \)); in that case the number of eigenvalues of matrix \( A \) would equal \( \dim(A) \). One of these eigenvalues would equal \( \text{trace}(A) \), and all other eigenvalues would equal zero. As matrix \( A \) is a reciprocal matrix (and by consequence all elements on the diagonal equal 1) the value of \( \text{trace}(A) \) would equal \( \dim(A) \). By consequence, the single non-zero eigenvalue of matrix \( A \) would equal \( \dim(A) \).

Or, in other terms, if matrix \( A \) were consistent, and if we bear in mind the definition of matrix \( A \) in formula (A.1), then we would find that there is exactly one value for \( n \) that satisfies formulas (A.6) and (A.7). This value would be the single (non-zero) eigenvalue of the consistent matrix \( A \).

\[
(A - n \cdot I)w = 0 \quad \text{(A.6)}
\]

in which:

\[
w = \text{vector containing the scores of all alternatives on one criterion}
\]

or:

\[
Aw = n \cdot w \quad \text{(A.7)}
\]

However, if matrix \( A \) is not consistent, then there will be more than one non-zero value for \( n \) in the last two formulas. These values are the new eigenvalues \( \lambda_i \) of matrix \( A \). The Perron-Frobenius eigenvector method uses the largest of these eigenvalues, which is called \( \lambda_{\text{max}} \), as the best available estimate of the eigenvalue \( n \) that would be found if matrix \( A \) were consistent.

The values of the eigenvalues \( \lambda_i \) are calculated by solving (in case of our example):

\[
\det(A - \lambda I) = \det\begin{vmatrix} 1 - \lambda & \frac{1}{3} & \frac{1}{2} \\ 3 & 1 - \lambda & 3 \\ 2 & \frac{1}{3} & 1 - \lambda \end{vmatrix} = 0 \quad \text{(A.8)}
\]

The largest of these eigenvalues, \( \lambda_{\text{max}} \), is then used to determine the actual vector \( w \) containing the scores of all three alternatives on the criterion maintenance such that:

\[
(A - \lambda_{\text{max}} \cdot I)w = 0 \quad \text{(A.9)}
\]

or, in case of our example:

\[
\begin{vmatrix} 1 - \lambda_{\text{max}} & \frac{1}{3} & \frac{1}{2} \\ 3 & 1 - \lambda_{\text{max}} & 3 \\ 2 & \frac{1}{3} & 1 - \lambda_{\text{max}} \end{vmatrix} \begin{bmatrix} W_A \\ W_B \\ W_C \end{bmatrix} = 0 \quad \text{(A.10)}
\]

Is there any guarantee that the scores in the estimated vector \( w \) are close enough to the actual, correct values that would have been found if matrix \( A \) were consistent? Yes, the relative AHP method requires that any vector \( w \) that has been calculated as just described, satisfies the condition:

\[
\frac{\lambda_{\text{max}} - n}{n - 1} \leq 0.10 \quad \text{(A.11)}
\]

This condition has been determined by Saaty as a heuristic rule. If the vector \( w \) does not satisfy this condition, then the user of the decision method must give a new, more accurate assessment of all alternatives on the currently investigated criterion. This process is repeated until condition (A.11) is satisfied.

Now the scores of the alternatives on the criterion of maintenance are known. Scores on other criteria must be determined in the same, interactive, way.

From this example it will be clear that relative assessment methods are not fit for use in a computer model; each individual decision requires a lot of interactive questions to the user of the model. Therefore, these methods cannot be used for often repeated decisions, as found in the computer model described in this thesis.
Appendix A. Decision techniques

Absolute assessment methods are based on a comparison of each criterion of all alternatives with an imaginary standard alternative; the user of the method must indicate for each criterion, of each alternative what the preference for a specific proposal will be. This approach makes that absolute assessment methods can be used in a computer model. Following example will clarify this.

Suppose that a container must be stored in a stack; this is the proposal. This proposal must be matched with a choice; this is a storage block in the stack. Each storage block in the stack has its own preferences for specific criteria; these are the relevant attributes of the container. Absolute assessment methods tell us immediately what preference each storage block has for the container.

If we use the absolute AHP method as described in Otto [22], page 19, in combination with the so-called REMBRANDT scale, then the decision process will be as follows. The REMBRANDT scale assigns to each of the semantic preference judgements shown in table A.1 a numeric value, forming a geometrical series. These numeric values are shown in table A.3. The REMBRANDT scale should always be used in combination with multiplicative AHP, which is a MCA method that will be discussed in the next section.

Note that the values shown in table A.3 actually are relative valuations, relative to the imaginary standard alternative. However, if all valuations of all criteria are relative to this imaginary standard alternative, then they are mutually relative, too.

Suppose that in our example we consider two storage blocks named I and II, and that a storage block is selected on basis of its appraisal for the owner of the container owner only. Suppose that the modelled system discards only two different owners of containers, named "Nin-Jan" and "Miersk".

In that case we can specify the appraisal of each storage block for each owner. If we compare the owner of a container with the appraisals at the storage blocks, then we can determine the best storage block for the container.

Suppose that the appraisals for the two owners at the two blocks are as shown in table A.4. In that case any container belonging to "Miersk" would be assigned to storage block I, and any container belonging to "Nin-Jan" would be assigned to storage block II.

Note that this example considers a very simple case in which the owner of the container is the only relevant criterion. In reality many more criteria will be included in the final decision, making the necessity of the REMBRANDT scale clear. This will be discussed in detail in the next section.

<table>
<thead>
<tr>
<th>Storage block</th>
<th>Owner</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&quot;Nin-Jan&quot;</td>
</tr>
<tr>
<td>I</td>
<td>Indifference</td>
</tr>
<tr>
<td>II</td>
<td>Indifference</td>
</tr>
</tbody>
</table>

Table A.4: Example of appraisals using the absolute AHP method.

Interval assessment methods are based on a subdivision of the range between extremes in a number of intervals of equal strength. Each interval has a semantic name like "+++", "--" or "Good". These intervals are converted into numeric values by means of a so-called utility function \( U \). Each criterion that is considered in the decision process has a different utility function assigned to it. Therefore, the number of different utility functions may be rather big. If we bear in mind that in practice it can take up to one full day to establish the correct course of one single utility function, then it will be clear that interval assessment methods are not preferred for use in a model as described in this thesis. Nevertheless, interval assessment methods may be valuable for major investment decisions. An example of an interval assessment method, UTA, can be found in Jacquet-Lagreze [11].
Appendix A. Decision techniques

♦ Selected method
Because of its capability to model repeating decisions by setting all decision parameters only once, and of its strong relation with multiplicative AHP (to be discussed in the next section), the REMBRANDT scale is selected for implementation in TORCMOD.

♦ Summary
Human assessment can be modelled in several ways. All available methods can be divided into three groups: that of relative assessment, that of absolute assessment, and that of interval assessment.

The only group of methods that can be used in a logistic computer model as described in this thesis is that of absolute assessment. The reason for this is, that this group allows repeated choices without any human intervention, once all decision parameters have been set.

For use in TORCMOD the absolute AHP method, in combination with the REMBRANDT scale, has been chosen. This is strongly related with the choice for multiplicative AHP as the best MCA method, as is discussed in the next section.

A.3. Multi criteria analysis

What is Multi criteria analysis? Multi criteria analysis (MCA) is a technique to determine the best out of a set of alternatives, where each alternative is assessed on basis of more than one criterion. Note that this definition is very broad; it does not necessarily induce the use of numeric values, but remains open for use of simple "better-than" relations.

A great number of MCA methods is available. This section gives a description of the three main groups of MCA methods, and shows which groups can be used and which can not. Each group contains several methods. Some of the most interesting methods are described in detail, and finally the best one is selected.

♦ Method requirements
What requirements should a MCA method in a large-scale logistic model meet? First, it should accept a mix of totally different criteria. These criteria may differ in their value ranges, in their unit and in their direction; this is explained in detail below. Second, it should take into account the differences in importance of the criteria as not all criteria will have the same weight in the final score. Third, it should show a natural pay-off behaviour when two or more criteria of an alternative are in contradiction; this is explained below. Fourth, it should allow automatic repeated decisions, once all relevant decision parameters have been set. Note that this final requirement is analogous to the last one given for human assessment.

The first requirement, accepting a mix of totally different criteria, can be clarified by a small example. Consider the situation that a man wants to buy a house. Suppose that he thinks only two aspects (criteria) of the house relevant, namely the price and the future neighbours. Then the value range, the unit and the direction of these criteria could be as shown in table A.5; note the enormous difference between these two criteria. The meaning of the value range and of the unit will be clear without any further explanation. The meaning of the direction, on the other hand, may need some explanation: If the direction of a criterion is "upwards" then a high score on this criterion is preferred over a low score; if, on the other hand, the direction of a criterion is "downwards" then a low score on this criterion is preferred over a high score. A good MCA method copes with all these differences between criteria.

<table>
<thead>
<tr>
<th></th>
<th>Value range</th>
<th>Unit</th>
<th>Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>100.000 .. 250.000</td>
<td>US$</td>
<td>Downwards</td>
</tr>
<tr>
<td>Neighbours</td>
<td>+++ .. +++</td>
<td>??</td>
<td>Upwards</td>
</tr>
</tbody>
</table>

Table A.5: Example of differences in criteria.

The second requirement, taking into account the differences in importance of the criteria, can be illustrated by expanding our housing example. Suppose that two different persons want to buy a house. The first person is a travelling salesman who earns little money and is seldom at home. The second person is a rich married man who takes care of the children and the house. As the travelling salesman is seldom at home he will care less for his neighbours than the married man will; therefore, the travelling salesman will give a lower importance to
good neighbours than the married man. On the other hand, the travelling salesman may desire to spend less money on his house than the married man; therefore, the travelling salesman will give a higher importance to a low price than the married man. Due to these differences in importance the travelling salesman may buy a totally different house than the married man. This small example yields the conclusion that (1) the importance of the criteria can have a major impact on the final scores of the alternatives, and (2) the importance of each criterion is very subjective; it strongly depends on the person using the decision model. This last aspect should always be borne in mind when using a computer model.

The third requirement, providing a natural pay-off behaviour when two or more criteria of an alternative are in contradiction, can be shown by slightly expanding our housing example. Suppose that we do not have just one criterion called "Neighbours", but that we want to discern between our future neighbours on the left hand side, and them on the right hand side. In that case both types of neighbours would have the same importance, but their values could differ. Suppose that our future neighbour at the left hand side is an absolute nerd, while the one at the right hand side is like an angel. Then it is most likely that we do not have a clear opinion on our neighbours as a whole; the sum of our neighbours would score something like "indifferent", but certainly not something like "very pleasant neighbours" or "extremely boring neighbours". This aspect of a natural pay-off must therefore be included in a good MCA method.

The fourth requirement, allowing automatic repeated decisions, yields the conclusion that the user should specify the parameters for the decision process (i.e. the valuations of all alternatives on all criteria) only once, after which the decision process does not need to prompt the user anymore to make a decision.

*Available methods*

All multi-criteria analysis methods can be divided into three groups. The first group is the American group which contains methods that try to calculate one overall score per alternative, and use that overall score as a basis for selection of the best alternative. This group is also called MCDM, which stands for Multi Criteria Decision Making. The second group is the French group which contains methods that are based on pair-wise comparisons of alternatives on various criteria. These comparisons result in so-called preference relations between the alternatives, based on a system of binary out-ranking methods. Refer to Lootsma [19]. This group is also called MCDA which stands for Multi Criteria Decision Aid. The third group is the interactive group which contains methods that try to find the best alternative through a number of interactive questions to the user. It should be clear that this group can not be used in a computer model as described in this thesis, and therefore is not further elaborated.

The first group, which is the American group, focuses on finding a mathematical function that calculates the overall score of an alternative, based on the scores of all criteria of the alternative. This function is called the value function \( V \). The value function \( V \) of alternative \( i \) is defined as (refer to Otto [22]):

\[
V_i = V(g_1(i), g_2(i), g_3(i), \ldots, g_n(i))
\]

(A.12)

in which:

\[ g_j(i) = \text{score of alternative } i \text{ on criterion } j. \]

An alternative \( i \) is preferred over alternative \( j \) if and only if \( V_i > V_j \). Each criterion has a weight factor associated with it; this weight factor determines the importance of the criterion. The precise use of these weight factors is hidden in the value function \( V \), and therefore differs between the American methods.

The second group, which is the French group, focuses on finding a partial pre-order structure of all alternatives, which expresses "better-than" relations between alternatives. An example of such a partial pre-order structure is shown in figure A.1 (Otto [22]). It shows the relative preferences between five alternatives named A through E. In the example of figure A.1, alternative A is preferred over alternative B, alternative B over alternatives C and D, and finally alternatives C and D over alternative E. Alternatives C and D are equally preferred.

![Example of a partial pre-order structure](image-url)
Appendix A. Decision techniques

Key elements in the French group are the binary relative preferences; these are just expressions for the relative assessment of two alternatives on one criterion. Some examples are given in table A.6. These binary relative preferences are transformed into a degree of credibility which expresses the degree that one alternative outranks another alternative, or not. Some French methods discern more different types of binary relative preferences than others.

<table>
<thead>
<tr>
<th>Relative preference</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a P b</td>
<td>a is strictly preferred over b</td>
</tr>
<tr>
<td>a I b</td>
<td>a and b are equally preferred</td>
</tr>
<tr>
<td>a R b</td>
<td>a and b cannot be compared</td>
</tr>
<tr>
<td>a S b</td>
<td>a is preferred at least as much as b</td>
</tr>
</tbody>
</table>

Table A.6: Example of binary relative preferences in French methods.

Just like in the American group, most methods in the French group assign to each criterion a weight factor \( p_i \); this weight factor determines the importance of each criterion. The use of the weight factors in the French group will be illustrated by discussing the ELECTRE I method, which is part of the important set of ELECTRE methods [22].

The ELECTRE I method focuses on dividing all alternatives into two groups: One group contains all attractive alternatives, and the other group contains all unattractive alternatives. A more mathematical definition of this objective is: To find a subset \( N \) of alternatives such that any alternative that is not in \( N \) is outranked by at least one alternative in \( N \), and that the alternatives of \( N \) are incomparable.

Like all French methods, the ELECTRE I method compares all alternatives mutually, on one criterion at the time. Suppose that alternatives \( a \) and \( b \) are compared under criterion \( j \). Then the ELECTRE I method tests whether \( a \bowtie b \) under \( j \) is true, or, in terms of the so-called preference function \( g_j \), whether \( g_j(a) \geq g_j(b) \). The results of all these tests will be used in formulas (A.14), (A.15) and (A.16).

Suppose that each criterion \( j \) has a weight factor \( p_j \). Then the total weight factor \( P \) would be:

\[
P = \sum_{j=1}^{n} p_j
\]  
(A.13)

Now the ELECTRE I method determines the so-called concordance index \( c(a,b) \), as the ratio of the weight factors of all criteria at which \( a \) is at least equal to \( b \), divided by the sum of all weight factors. Or, in a formula:

\[
c(a,b) = \frac{1}{P} \sum_{j \in j: g_j(a) \geq g_j(b)} p_j
\]  
(A.14)

Before the relation \( a \bowtie b \) can be evaluated, the user of the ELECTRE I method must specify the veto threshold \( v_j(g_j) \). This threshold is used to guarantee that all alternatives that score clearly too low on criterion \( j \) are totally excluded from further examination. So, if the difference in value between \( g_j(b) \) and \( g_j(a) \) exceeds the veto threshold \( v_j(g_j(a)) \) then alternative \( a \) is marked as unattractive. Or, in a formula, alternative \( a \) is marked as unattractive if the condition:

\[
g_j(a) + v_j(g_j(a)) < g_j(b)
\]  
(A.15)

is satisfied.

The last step in the ELECTRE I method is to test whether \( a \bowtie b \) is true when taking all criteria into account.

This test requires one more parameter: The concordance threshold \( \hat{c} \) which must be specified by the user of the ELECTRE I method. The proposition \( a \bowtie b \) is said to be true if the concordance index at least equals the concordance threshold, *and* if no veto condition is imposed on alternative \( a \). Or, in a formula:

\[
a \bowtie b \Leftrightarrow c(a,b) \geq \hat{c} \land \forall j \ g_j(a) + v_j(g_j(a)) \geq g_j(b)
\]  
(A.16)

The ELECTRE I method tests all relative preferences between all alternatives, which finally results in a partial pre-order structure as shown in figure A.1. All alternatives \( a \) that did not satisfy condition (A.16) have been placed in the group of unattractive alternatives.
Appendix A. Decision techniques

• Comparison of methods

These descriptions of the American and the French methods yield the conclusion that several major differences between the American and French methods exist. First, the American methods use (internally) only numeric values, while the French methods use binary outranking. Second, the American methods include all scores on all criteria in the final score, while the French methods show a more hierarchic approach. This hierarchic approach of French methods allows the total exclusion of less important or even contradicting criteria from the final "score". Nevertheless, some American methods approach this highly desired behaviour. Note that the total exclusion of alternatives due to contradicting criteria can also be achieved by removing them on beforehand from the set of alternatives, before starting the actual MCA method. When choosing either for an American or a French method all these differences should be taken into account.

Which group of MCA methods is preferred? The American group or the French group? Unfortunately, it is hard to make a well-founded choice between the American and the French group; both groups contain several good MCA methods that are fit for use in a large-scale logistic model. This dilemma, however, appears to be a general problem: Just look at the names of both groups and note that the American group is mainly used in English-speaking countries, while the French group is mainly used in French-speaking countries. The choice for one of both groups of MCA methods merely seems a matter of culture than one of rational arguments; in spite of many attempts from the side of "American" users to convince "French" users of the merits of the American methods, the "French" users still stick to the French methods. And vice versa.

• Selected method

Now then, given this dilemma, which MCA group should be chosen for use in TORCMOD? It has been chosen for the American methods for following reasons. First, the American methods are more easy to implement in a computer program than the French methods. This is caused by the purely numeric approach of the American methods; the binary outranking relations of the French methods (P, I, S) are more difficult to implement in a computer program. Second, the American methods are easier to understand for the user, as they abstain from hard-to-grasp elements like the concordance threshold. Third, the American methods open the path towards random selection which is discussed in next section; random selection uses the numeric final score of an alternative as a basis to determine the chance that the alternative is selected. As the French methods in general do not assign a numeric score to each alternative, they cannot be used in combination with random selection.

Which method from the American group should be chosen for use in TORCMOD? It has been chosen for the multiplicative variant of AHP, which will be discussed in detail below. The abbreviation AHP stands for Analytic Hierarchy Process.

• Comparison of three American methods

Although an almost infinite number of American methods exist, three of them have been selected as potentially interesting. The first method is SMART as defined by Von Winterfeldt and Edwards [34]. The second method is AHP as defined by Saaty [24]. The third method is the multiplicative variant of AHP as defined by Lootsma [19]. Each of these methods will be described below.

The SMART method is the most simple method. Actually, the SMART method is not just a MCA method; it is a method that combines human assessment and MCA. It directly rates the scores of all alternatives on all criteria, and then it determines the final scores of all alternatives as a weighted sum of these scores. Although the SMART method is very simple to implement, it is not fit for decisions that include many alternatives, as the user of the SMART method has to keep in mind all alternatives at the same time. A more detailed description of the SMART method can be found in Von Winterfeldt and Edwards [34].

The AHP method is a little more complex than the SMART method. The AHP method is often used, as it is still relatively simple and produces reasonable results. The AHP method determines the final score of an alternative by summing the alternative's scores on all criteria, multiplied by the weight factor of the criteria. Or, in a formula:

\[ V_i = \sum_{j=1}^{n} p_j \cdot g_j(i) \]  \hspace{1cm} (A.17)
Appendix A. Decision techniques

in which:

\[ V_i = \text{final score of alternative } i \]
\[ p_j = \text{weight factor of criterion } j \]
\[ g_j(i) = \text{score of alternative } i \text{ under criterion } j. \]

The values of the weight factors and the scores of the alternatives per criterion that are used in formula (A.17) are not the values that are specified by the user of the AHP method, nor the values that were obtained from numeric criteria. A special conversion process is used to determine the values of \( p_j \) and the values of \( g_j(i) \) that are used in formula (A.17). The first step in this process is to convert the weight factors as specified by the user of AHP into new weight factors by using the Perron-Frobenius eigenvalue method as described in formula (A.1) through (A.10). In this way the original weight factors and scores are replaced by estimates of the actual, correct values that would have been found if matrix \( A \) in formula (A.1) were consistent. The same holds for the scores of all alternatives on one criterion. Although the Perron-Frobenius eigenvalue method removes the inconsistency in the original weight factors and scores, and the resulting new values will be close to the actual, correct values, there is no guarantee at all that the new values are the best possible new values; there is no optimisation of any kind. Refer to formula (A.11). The second step in this process is to normalise the resulting weight factors and scores, such that:

\[ \sum_{i} p_i = 1 \quad (A.18) \]

and:

\[ \sum_{i} g_j(i) = 1 \quad \forall j \quad (A.19) \]

In this way the effects of differences in value range and in unit are removed from the final scores.

Although this AHP method produces reasonable results, it has been criticised for various reasons. First, it is criticised for the simple, linear scale to quantify human assessment. Psychologists state that a scale to quantify human assessment should be exponential. Refer to Lootsma [18], [19]. Second, it is criticised for using the Perron-Frobenius eigenvector method to determine the weight factors \( p_j \) and the scores \( g_j(i) \). This eigenvector method gives no guarantee at all that the estimated values are the best. Third, it is criticised for summing the scores on all criteria, which leads to unexpected pay-off behaviour.

The multiplicative variant of AHP is the most complex of the three methods, and its results are considered better than those of AHP and SMART. The multiplicative AHP method determines the final score of an alternative by multiplying the alternative's scores on all criteria, raised to the power of the weight factor of the criteria. Or, in a formula:

\[ V_i = \prod_{j=1}^{n} r_j(i)^{p_j} \quad (A.20) \]

in which:

\[ r_j(i) = \text{transformed score of alternative } i \text{ under criterion } j \text{ (to be explained below).} \]

While the normal AHP method required that the weight factors \( p_j \) were determined with the Perron-Frobenius eigenvector method, the multiplicative AHP method uses logarithmic regression to determine the values of \( p_j \).

What is logarithmic regression? Suppose we have determined the relative importance of all criteria 1 through \( n \). The results could be written in a matrix as shown in formula (A.21), showing all relative importance ratios. Compare this matrix with the one in formula (A.1). This matrix \( A \) is reciprocal, too, which means that \( a_{1,2} \) is the inverse of \( a_{2,1} \).
Appendix A. Decision techniques

\[
A = \begin{bmatrix}
1 & a_{1,2} & \ldots & a_{1,n} \\
a_{2,1} & 1 & \ldots & a_{2,n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n,1} & a_{n,2} & \ldots & 1
\end{bmatrix}
\]  \hspace{1cm} (A.21)

Based on this matrix \( A \) the weight factors \( p_j \) of each criterion must be determined such that:

\[
\sum_{i \neq j} (\ln a_{i,j} - \ln p_i + \ln p_j)^2
\]  \hspace{1cm} (A.22)

is minimal. This equation can be solved by calculating the normal vectors on \( A \); in case of four different criteria this problem would look like:

\[
\begin{bmatrix}
1 & -1 & & & \\
1 & -1 & \ln p_1 & & \\
1 & -1 & \ln p_2 & \ln a_{1,3} & \\
1 & -1 & \ln p_3 & \ln a_{2,3} & \\
1 & -1 & \ln p_4 & \ln a_{2,4} & \\
1 & -1 & & \ln a_{3,4} &
\end{bmatrix}
\]  \hspace{1cm} (A.23)

or:

\[
A' P' = B'
\]  \hspace{1cm} (A.24)

As this system is overdetermined the last element from \( P' \), \( p_4 \), is set to 1. By consequence, \( \ln p_4 \) equals zero, and therefore the last row of \( P' \) and \( B' \), disappear, just like the last column of \( A' \). The values of \( P' \) can now be determined by solving:

\[
P' = \left( A' A' \right)^{-1} A' B'
\]  \hspace{1cm} (A.25)

Now the weight factors \( p_j \) of all criteria are known and can be used in formula (A.20).

While the normal AHP method required that the scores on the criteria are normalised per criterion, the multiplicative AHP method does not need to normalise the scores by using formula (A.18) as any differences in scale between criteria are automatically removed through multiplication. However, there is one complication. This complication is, that the multiplicative AHP method does not simply use the scores per criterion as found and used in normal AHP (refer to formula (A.17)), but that the multiplicative AHP method first transforms the scores per criterion using logarithmic regression and then uses the resulting \( r_j \) scores per criterion in formula (A.20).

How to convert scores \( g_j(i) \) into scores \( r_j(i) \)? Suppose that four alternatives must be assessed on three criteria. Then the scores of all alternatives on all criteria are known. However, these scores must be transformed into new scores using logarithmic regression, one criterion at the time. So, if we consider the first criterion \( j \), then the relative values of the scores of all alternatives on this criterion are placed in matrix \( A \), analogous to formula (A.21).

\[
A_j = \begin{bmatrix}
1 & g_j(1) & g_j(1) & g_j(1) \\
g_j(2) & 1 & g_j(2) & g_j(2) \\
g_j(3) & g_j(3) & 1 & g_j(3) \\
g_j(4) & g_j(4) & g_j(4) & 1
\end{bmatrix}
\]  \hspace{1cm} (A.26)
Appendix A. Decision techniques

in which:

\[ A_j = \text{matrix with scores of all relative scores of all alternatives on criterion } j. \]

Now the actual scores \( r_j(i) \) that are used in formula (A.20) are calculated by solving formula (A.25), where the values in \( P_j' \) represent the scores of all alternatives on criterion \( j \).

♦ Human assessment and numeric scales in the multiplicative AHP method

Do all scores have to be converted by logarithmic regression before they can be used for multiplicative AHP? No, there is one exception to this rule: Scores that are obtained from geometrical series do not need to be converted by logarithmic regression; these scores can be used directly in formula (A.20), without any conversion by logarithmic regression. This feature is used by the REMBRANDT scale that was discussed in section A.2. Refer to Lootsma [18].

The use of this feature is not restricted to human assessment only. It can also be used to convert scores on numeric criteria into scores that can be used directly in formula (A.20), without any conversion by logarithmic regression. This direct conversion can only be used for those criteria that have a known and limited value range. A good example hereof is a criterion that represents a percentage: The value range is known, and restricted to the interval [0..100%].

How does direct conversion work? Direct conversion replaces the original scores on a numeric criterion with one of the values of the REMBRANDT scale.

But which value of the REMBRANDT scale should be used? Direct conversion relates each semantic preference judgement of the REMBRANDT scale with a specific part of the value range of the criterion. Such a part is called a sub-range. An example of sub-ranges is given in table A.7. For each of the nine sub-ranges, this table shows the related semantic preference judgement, the lower bound, the upper bound and the length. The length of a sub-range is simply the difference between the upper bound and the lower bound.

<table>
<thead>
<tr>
<th>Semantic preference judgement</th>
<th>Sub-range</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower bound</td>
<td>Upper bound</td>
</tr>
<tr>
<td>1. Very strong preference</td>
<td>0.0 %</td>
<td>11.1 %</td>
</tr>
<tr>
<td>2. Strong preference</td>
<td>11.1 %</td>
<td>22.2 %</td>
</tr>
<tr>
<td>3. Average preference</td>
<td>22.2 %</td>
<td>33.3 %</td>
</tr>
<tr>
<td>4. Weak preference</td>
<td>33.3 %</td>
<td>44.4 %</td>
</tr>
<tr>
<td>5. Indifference</td>
<td>44.4 %</td>
<td>55.6 %</td>
</tr>
<tr>
<td>6. Weak negative</td>
<td>55.6 %</td>
<td>66.7 %</td>
</tr>
<tr>
<td>7. Average negative</td>
<td>66.7 %</td>
<td>77.8 %</td>
</tr>
<tr>
<td>8. Strong negative</td>
<td>77.8 %</td>
<td>88.9 %</td>
</tr>
<tr>
<td>9. Very strong negative</td>
<td>88.9 %</td>
<td>100.0 %</td>
</tr>
</tbody>
</table>

Table A.7: Example of direct conversion of a numeric criterion for multiplicative AHP.

So, if the original score on a numeric criterion is known, then this value is replaced with the value from table A.3 that corresponds with the REMBRANDT semantic preference judgement that is related to the sub-range in which the original score on the numeric criterion lies. The upper bound is an open bound, except for the last semantic preference judgement. This will be clarified with an example.
Appendix A. Decision techniques

Consider the situation that a transportation job must be assigned to one out of three pools of equipment. Suppose that the only criterion that is taken into account in the decision process is the momentary percentage of free capacity of a pool: if a pool is fully occupied then the momentary percentage of free capacity equals 0%; if, on the other hand, a pool is fully idling then the momentary percentage of free capacity equals 100%.

Suppose that in our example an equipment pool has a free capacity of 44.4%. In that case the semantic preference judgment related to the numeric score is "Indifference"; the upper bound of "Weak preference" is open, and therefore "Weak preference" is not selected. Refer to table A.7. The value associated with the REMBRANDT semantic preference judgment "Indifference" equals 1 (refer to table A.3), and therefore the score of the equipment pool will be 1 instead of 44.4%. This score can be used in formula (A.20) without any further conversion.

Should all sub-ranges be equally long, as shown in table A.7? No, there is a great amount of freedom in varying the lengths of the sub-ranges, just as long as the sub-ranges satisfy the previously given demands.

A good example of the use of varying sub-ranges is shown in figure A.2. It shows the complement of the lower bound of the sub-ranges for three different attitudes towards risk: Risk neutral, risk avoiding and risk seeking. The values corresponding to the risk neutral attitude are taken from table A.7; the values corresponding to the risk avoiding and the risk seeking attitude depend on the user of the decision method and on the type of risk that is modelled.

Figure A.2: Example of different types of direct conversion.

Summary
Multi criteria analysis (MCA) is a technique to determine the best out of a set of alternatives, where each alternative is assessed on basis of more than one criterion.

There are three different groups of MCA methods: The American group, the French group and the interactive group. Both the American and the French group contain several methods that can be used in a large-scale logistic model. Mainly because of its fully numeric approach, the American group is preferred over the French group.

For use in TORCMOD the multiplicative AHP method has been chosen. This is strongly related with the choice for absolute AHP, in combination with the REMBRANDT scale, as the best human assessment method.

A.4. Random selection

The final step in the decision process is random selection. While human assessment has been used to specify semantic appraisals for several criteria, and MCA has been used to determine the final score of each alternative, random selection will be used to determine the single alternative which is to be selected as the "best", and to be assigned to the proposal. This selected alternative is called the choice.

Why making things more complicated than they are? Why don't we simply select the alternative with the highest final score as the best? This question will be answered with an example.
Suppose we had to feed an elephant by giving him one out of five baskets containing food; each basket containing a different mix of apples, bread and peanuts. Therefore, each basket has its own overall appraisal for use by the elephant. These overall appraisals, or final scores, are shown in table A.8. One might say that the elephant should get basket V, as that one has the highest final score of all. In the real-life system, however, we would have a problem: The guard who has to decide which basket to use, does not know the exact final scores of each basket; he only has some vague indications which baskets are good and which are not. Therefore, basket V may have a good chance of being selected, but sometimes the guard may decide to give the elephant basket IV, or even basket III; the guard does not decide in the same exact way as a computer would, due to a lack of exact information. To model this human selection behaviour we need random selection.

Note that random selection is only needed to model real-life systems in which human beings decide what to do; it is not needed to model real-life systems in which automatic systems decide what to do, based on exact information. In that case it is sufficient to simply select the alternative with the highest final score as the best.

**Definition**
In short, random selection is a mathematical technique to include the "fuzzy" character of the human decision process in a computer model. Random selection uses the results of MCA as input data, and selects one alternative as the "best" by using random numbers.

**Requirements**
Random selection has several aspects. Each aspect highly influences the behaviour of the random selection process. The first aspect is called the exponential growth. The exponential growth determines the weight that is given to the highest scores; the higher the exponential growth, the more important the highest scores. The second aspect is called the skip level. The skip level determines the minimum demand for an alternative to be interesting for selection. The higher the skip level, the less alternatives are interesting. The third aspect is called the zero-based option. The zero-based option is used to put even more stress on the highest scores, in combination with a non-zero skip level. This is explained below.

All three aspects will be clarified by gradually expanding the example of the elephant. Suppose we consider random selection in its simplest form. In that case each basket would have a chance of being selected that is linear to its final score; a basket scoring 8 would have twice as much chance of being selected as a basket scoring 4. In short, a basket i scoring $s_i$ has a chance $p_i$ of:

$$ p_i = \frac{s_i}{\sum_j s_j} \quad (A.27) $$

in which:

- $s_i =$ final score of alternative $i$
- $p_i =$ chance of alternative $i$ of being selected

of being selected. This type of random selection is called linear random selection as the chance of being selected is linear with the alternative's final score. If we want to put some more stress on the highest scores, then we must transform (A.27) into:

$$ p_i = \frac{s_i^2}{\sum_j s_j^2} \quad (A.28) $$

This type of random selection is called quadratic random selection as the chance of being selected is (almost) quadratic with the alternative's final score.

This brings us to the first aspect, called exponential growth. In general, we can say that the exponential growth of the selection method is determined by the exponential growth factor $p$. In case of linear random selection $p$
equals 1, and in case of quadratic random selection \( p \) equals 2. The chance of each alternative of being selected equals:

\[
p_i = \frac{s_i^p}{\sum_j s_j^p}
\]

in which:

\[
p = \text{exponential growth factor}.
\]

Returning to our elephant, we would like to determine the effect of the exponential growth factor on the chance each basket has of being selected. These values have been calculated for three different settings of the exponential growth factor, and are shown in table A.9. This table shows clearly that the relative chances change dramatically when the exponential growth factor increases. If we compare basket II and V, their relative chance changes from 2 in case of linear random selection to almost 8 in case of 3rd power random selection.

The second aspect, the skip level is used to excluded a number of uninteresting alternatives from the selection process. In case of our elephant we might state that we are not interested at all in any basket that scores less than 80% of the best basket. In that case the skip level would equate 80%; all alternatives that score less than 80% of the best alternative are excluded from the selection process. If we apply a skip level of 80% to our elephant problem, then the chances for all alternatives change dramatically. See table A.10. It appears that baskets I through III are not interesting anymore, and are skipped.

From this description it yields that the skip level is defined as the minimum percentage of the highest final score that each alternative must score, to be included in the decision process.

The third aspect, the zero-based option is used in combination with a non-zero skip level to put even more stress on the highest scores. How can this be done? Suppose that we use a skip level without using the zero-based option. Then the formulas describing the decision process are very much like those shown in (A.31); the only difference would be that all alternatives scoring less than the minimum required score are skipped. If, on the other hand, we do use the zero-based option, then the formulas for those alternatives that are not skipped look like:

<table>
<thead>
<tr>
<th>Basket</th>
<th>Score</th>
<th>( p = 1 ) Linear</th>
<th>( p = 2 ) Quadratic</th>
<th>( p = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>2.5</td>
<td>8.9 %</td>
<td>3.5 %</td>
<td>1.3 %</td>
</tr>
<tr>
<td>II</td>
<td>4.0</td>
<td>14.3 %</td>
<td>9.0 %</td>
<td>5.2 %</td>
</tr>
<tr>
<td>III</td>
<td>6.0</td>
<td>21.4 %</td>
<td>20.2 %</td>
<td>17.6 %</td>
</tr>
<tr>
<td>IV</td>
<td>7.5</td>
<td>26.8 %</td>
<td>31.5 %</td>
<td>34.3 %</td>
</tr>
<tr>
<td>V</td>
<td>8.0</td>
<td>28.6 %</td>
<td>35.9 %</td>
<td>41.6 %</td>
</tr>
<tr>
<td>Sum</td>
<td>28.0</td>
<td>100.0 %</td>
<td>100.0 %</td>
<td>100.0 %</td>
</tr>
</tbody>
</table>

Table A.9: Example of exponential growth factor \( p \).

<table>
<thead>
<tr>
<th>Basket</th>
<th>Score</th>
<th>( p = 1 ) Linear</th>
<th>( p = 2 ) Quadratic</th>
<th>( p = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>2.5</td>
<td>--- %</td>
<td>--- %</td>
<td>--- %</td>
</tr>
<tr>
<td>II</td>
<td>4.0</td>
<td>--- %</td>
<td>--- %</td>
<td>--- %</td>
</tr>
<tr>
<td>III</td>
<td>6.0</td>
<td>--- %</td>
<td>--- %</td>
<td>--- %</td>
</tr>
<tr>
<td>IV</td>
<td>7.5</td>
<td>48.4 %</td>
<td>46.8 %</td>
<td>45.2 %</td>
</tr>
<tr>
<td>V</td>
<td>8.0</td>
<td>51.6 %</td>
<td>53.2 %</td>
<td>54.8 %</td>
</tr>
<tr>
<td>Sum</td>
<td>28.0</td>
<td>100.0 %</td>
<td>100.0 %</td>
<td>100.0 %</td>
</tr>
</tbody>
</table>

Table A.10: Example of skip level 80%.

<table>
<thead>
<tr>
<th>Basket</th>
<th>Score</th>
<th>( p = 1 ) Linear</th>
<th>( p = 2 ) Quadratic</th>
<th>( p = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>2.5</td>
<td>--- %</td>
<td>--- %</td>
<td>--- %</td>
</tr>
<tr>
<td>II</td>
<td>4.0</td>
<td>--- %</td>
<td>--- %</td>
<td>--- %</td>
</tr>
<tr>
<td>III</td>
<td>6.0</td>
<td>--- %</td>
<td>--- %</td>
<td>--- %</td>
</tr>
<tr>
<td>IV</td>
<td>7.5</td>
<td>40.7 %</td>
<td>32.1 %</td>
<td>24.5 %</td>
</tr>
<tr>
<td>V</td>
<td>8.0</td>
<td>59.3 %</td>
<td>67.9 %</td>
<td>75.5 %</td>
</tr>
<tr>
<td>Sum</td>
<td>28.0</td>
<td>100.0 %</td>
<td>100.0 %</td>
<td>100.0 %</td>
</tr>
</tbody>
</table>

Table A.11: Example of zero-based option (skip level 80%).
Appendix A. Decision techniques

\[ P_i = \frac{(s_i - c_{\text{skip}} s_{\text{max}})^p}{\sum_j (s_j - c_{\text{skip}} s_{\text{max}})^p} \]  

(A.31)

in which:

- \( c_{\text{skip}} \) = skip factor
- \( s_{\text{max}} \) = maximum score of all alternatives.

In fact, if the zero-based option is used, then all final scores are reduced by the minimum required score (which equals \( c_{\text{skip}} s_{\text{max}} \)), and then these reduced values are used as if they were the actual final scores, that were determined with MCA. In this way the distinction between the scores above the minimum score becomes more evident.

Just compare table A.10 with table A.11 and all will be clear. Note that the relative importance of basket V, which scores just a little better than basket IV, is definitely higher if we apply the zero-based option.

![Figure A.3: Example of zero-based selection (exponential growth 3, skip level 80%).](image.png)

This can also be seen from figure A.3. It shows the scores of basket IV and V using 3rd power random selection with a skip level of 80%, both for zero-based and for non-zero-based decisions.

- **Usage of zero-based option**

  When should the zero-based option be used? The zero-based option should be used mainly in those cases where the skip level is set to about 70% or more. This reason for this is twofold. First, the zero-based option is needed to re-scale the highest final scores; in this way the distinction between those final scores above the skip level is increased. Second, the zero-base option is needed to obtain a smooth entrance of final scores that are close to the skip level; this is explained below.

  The first reason, re-scaling all final scores above the skip level, if clarified by looking at figure A.3. This figure shows clearly that introducing the zero-based option increases the distinction between basket IV (scoring 7.5) and basket V (scoring 8.0) in case of a skip level of 80% (which is 6.4). Although it has not been investigated explicitly, it is believed that this increased distinction is essential for modelling real-life decisions by human beings.

  Note, however, that re-scaling can cause some odd effects when the skip level is extremely high. An example will clarify this. Consider the situation that basket IV scores 7.8 instead of 7.5, that basket V still scores 8.0, that the skip level is increased to 95%, that the exponential growth factor is set to 3, and that the zero-based option is used. In that case basket IV would have a chance of just 11.1% of being selected, while basket V would have a chance of 88.9%. Normally, human beings won't be able to make such a clear distinction between an alternative scoring 7.8 and an alternative scoring 8.0. Therefore, extremely high skip levels should be used with care.
The second reason, obtaining a smooth entrance of final scores that are close to the skip level, is clarified by slightly modifying our elephant example. Suppose that basket III would score 6.5 instead of 6.0. In that case it would score just a little better than the skip level of 80% (which is 6.4), and therefore it would get a non-zero chance of being selected. If we would not use the zero-based option, then the results would be as shown in table A.12. Just compare these results with those found in table A.10, and note that a slight increase of one final score can dramatically change the results of random selection, if we do not use the zero-based option.

But, if we do use the zero-based option for our slightly modified elephant example, then the results would be as shown in table A.13. Just compare these results with those found in table A.11, and note that the zero-based option has almost eliminated the effect of the slight increase of the final score of basket III. This behaviour of the decision process is much more natural than that shown in table A.12.

The necessity of smoothed entrance of final scores that are close to the skip level can also be seen from figures A.4 and A.5. Both figures show the chances of selection for baskets III, IV and V in two cases: the first case is when basket III scores 6.0, and the second case is when basket III scores 6.5. Both cases are based on linear random selection, combined with a skip level of 80%. The difference is that figure A.4 shows the results when the zero-based option is not used, while figure A.5 shows the results when the zero-based option is used. It should be clear that the latter results seem to be more natural.

<table>
<thead>
<tr>
<th>Basket</th>
<th>Score</th>
<th>p = 1 Linear</th>
<th>p = 2 Quadratic</th>
<th>p = 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>2.5</td>
<td>--- %</td>
<td>--- %</td>
<td>--- %</td>
</tr>
<tr>
<td>II</td>
<td>4.0</td>
<td>--- %</td>
<td>--- %</td>
<td>--- %</td>
</tr>
<tr>
<td>III</td>
<td>6.5</td>
<td>29.5 %</td>
<td>26.0 %</td>
<td>22.7%</td>
</tr>
<tr>
<td>IV</td>
<td>7.5</td>
<td>34.1 %</td>
<td>34.6 %</td>
<td>34.9%</td>
</tr>
<tr>
<td>V</td>
<td>8.0</td>
<td>36.4 %</td>
<td>39.4 %</td>
<td>42.4%</td>
</tr>
<tr>
<td>Sum</td>
<td>28.5</td>
<td>100.0 %</td>
<td>100.0 %</td>
<td>100.0 %</td>
</tr>
</tbody>
</table>

Table A.12: Necessity of entrance smoothing (modified example).

<table>
<thead>
<tr>
<th>Basket</th>
<th>Score</th>
<th>p = 1 Linear</th>
<th>p = 2 Quadratic</th>
<th>p = 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>2.5</td>
<td>--- %</td>
<td>--- %</td>
<td>--- %</td>
</tr>
<tr>
<td>II</td>
<td>4.0</td>
<td>--- %</td>
<td>--- %</td>
<td>--- %</td>
</tr>
<tr>
<td>III</td>
<td>6.5</td>
<td>3.6 %</td>
<td>0.3 %</td>
<td>0.0 %</td>
</tr>
<tr>
<td>IV</td>
<td>7.5</td>
<td>39.3 %</td>
<td>32.0 %</td>
<td>24.5%</td>
</tr>
<tr>
<td>V</td>
<td>8.0</td>
<td>57.1 %</td>
<td>67.7 %</td>
<td>75.5%</td>
</tr>
<tr>
<td>Sum</td>
<td>28.5</td>
<td>100.0 %</td>
<td>100.0 %</td>
<td>100.0 %</td>
</tr>
</tbody>
</table>

Table A.13: Example of entrance smoothing by using the zero-based option (modified example).

![Figure A.4: Non-smoothed entrance of basket III.](image)

![Figure A.5: Smoothed entrance of basket III by using the zero-based option.](image)

These examples yield the conclusion that use of the zero-based option is strongly preferred when modelling human decision behaviour. Otherwise we could be confronted with some unexpected side-effects.

- Summary

Computer models that include human decision behaviour should also include random selection. Random selection uses the results of MCA as input data.
Appendix A. Decision techniques

There are three different ways to control the behaviour of random selection. First, the exponential growth rate determines the weight that is given to the highest scores. Second, the skip level determines the minimum demand for an alternative to be interesting for selection. Third, the zero-based option puts even more stress on the highest scores, in combination with a non-zero skip level.

Human behaviour should be modelled by using a skip level somewhere between 70% and 90%, in combination with the zero-based option. Extremely high skip levels should be avoided, and care should be taken when combining them with the zero-based option.
Appendix B. Simulation techniques

B.1. Introduction

At a number of occasions there is a need to obtain information on the efficiency of factories or other logistical systems. Such information is needed, for instance, when it is suspected that such a system is functioning in sub-optimal way and could be improved; or it is needed when a new logistical system it to be developed and there is a need to get an impression on the performance of that new, not yet existing, system. This latter type of use is commonly accepted when designing and analysing new logistical systems and concepts.

Next to these simulation techniques some other analysis methods exist, like queuing theory or spreadsheet calculations. These latter techniques are called static techniques, in contrast to simulation which is called a dynamic technique. The main difference between static techniques and dynamic techniques is the explicit inclusion of the factor time in the analysis method: While static analysis techniques do not discern a continuum of time periods, dynamic techniques do. These latter techniques will be given some more attention in the appendix.1

An example of the use of simulation will be discussed by referring to figure B.1. This figure shows two barrels of milk, one of which (the upper one) overflows into the second one (the lower one). The size of this overflow varies in time according to some random processes that lie outside the scope of the here discussed simulation model.

The milk that flows into the lower barrel is collected there, but in the meantime flows freely out of the lower barrel again under influence of gravity only. As the simulation model discussed here only focuses on the variation of the level of this lower barrel of milk, two flows need to be included in the model: One is the flow of milk into the lower barrel, and the other one is the flow of milk out of the lower barrel. These flows are marked flow\_in(t) and flow\_out(t) respectively in figure B.1. Because of some physical processes, this latter flow appears to be dependent on the current level of milk in the lower barrel according to following relation:

$$\text{flow}_{\text{out}}(t) = C \cdot \sqrt{\text{level}(t)},$$  \hspace{1cm} (B.1)

where C is a constant. The level of milk in the lower barrel is determined by the relation:

$$\text{level}(t) = \text{level}_{t=0} + \frac{1}{A} \int_{\tau=0}^{t} (\text{flow}_{\text{in}}(\tau) - \text{flow}_{\text{out}}(\tau)) \cdot d\tau,$$  \hspace{1cm} (B.2)

in which:

$$A = \text{surface of the lower barrel}.$$

A good impression of the variance of the level of milk in the lower barrel can be obtained by solving for a number of subsequent, equally long, time periods both equations (B.1) and (B.2) in a more or less discrete way: The flow out of the barrel during one time period is based on the level of the barrel in the previous time period. If the difference between both time periods is small enough than the behaviour of differential equation (B.2) can be approached sufficiently accurately. A small example hereof is given in figure B.2.

1 More information can be obtained, for example, from Dekker [4] and from magazine "Simulation".
Appendix B. Simulation techniques

Figure B.2 shows three lines in time: (1) the random flow into the lower barrel of milk which serves as an input signal for the simulation, (2) the flow out of the lower barrel of milk which is dependent on the level of milk in the barrel and (3) the total level of milk in the barrel. Note that the unit of measure for both flows is different from the unit of measure for the level of milk, which however is not shown in figure B.2.

![Graph showing flows and level over time](image)

**Figure B.2:** Example of simulation results of two barrels of milk.

Based on the results of such a simulation a quick impression is obtained for the variations in the level of milk in the lower barrel in time, which can be useful to get for instance an impression of the required size of the barrel such that it is guaranteed that the barrel will never overflow. This is a typical example of the use of simulation: Analysis of dynamic process that show a time-dependent behaviour and are too complicated to analyse by static calculations.

Two types of simulation can be discerned: continuous simulation which has been used in the previous example, and discrete simulation. Both types of simulation are used in TORCMOD and will be given some more discussion.

### B.2. Continuous simulation

The first type of simulation, *continuous* simulation, was applied in the system analysis example of figure B.2 and is also used at UCM level of detail of the proposed terminal design method. As can be seen from this example continuous simulation is characterised by a subdivision of the entire time span that is simulated into a number of equally long time windows. For each of these time windows the state of the entire modelled system is determined, based on the initial state in the previous window and on the differential equations that describe the behaviour of the system. An example of such a differential equation can be found from formula (B.2).

*Continuous simulation* is particularly fit for analysis of systems through which *flows* of some entity like milk, oil or US dollars run. A drawback of continuous simulation is that there is no distinction between the individual elements that flow through the modelled system. So, referring to the above example, continuous simulation does *not offer the possibility* to trace the course of individual particles of milk which in this case probably is not that serious. But when analysing a terminal system, and when the flow through the modelled system represent load units, than in many cases it is needed to distinct between individual load units; which can not be done by using continuous simulation. Therefore, the benefit of analysing a terminal system by means of continuous simulation is limited, as will be indicated in Chapter 5. For better analysis results another simulation technique should be used: Discrete simulation.

### B.3. Discrete simulation

Discrete simulation, in contrast to continuous simulation, *does* discern individual elements like particles of milk or load units. By consequence more detailed analysis results can be obtained than by using continuous simulation. Another consequence of using discrete simulation is the aspect of *parallel processes* of various *components* in the modelled system. This will be illustrated by an example.

---

2 Although disputable in this particular case, almost any system of barrels that is little more complex than shown in figure B.1 will be too complicated for static calculations.
Appendix B. Simulation techniques

Consider figure B.3 in which a train is being unloaded by a gantry crane. This gantry crane picks up load units from the train and places them on some machine for internal transport (not depicted). Sometimes the gantry crane needs to move along the train to pick up the next load unit, or to drive to the machine for internal transport.

Figure B.3: Example situation for discrete simulation.

From this example, that can be modelled by means of discrete simulation, a number of characteristics of a model for discrete simulation can be found:

- the model is based on a number of individual components like trains, gantry cranes and load units.
- the model is based on process descriptions that define what actions all these components are to perform, and in what sequence.

These two aspects are very important as they have a major impact on the possibilities to model process, and to analyse them after simulation. When looking in more detail at continuous and discrete simulation and number of differences, but also a number of analogies can be found. Studying these differences and similarities will increase understanding of both methods of simulation. In short, these analogies and differences are:

1. where continuous simulation used flows of some entity, discrete simulation uses individual entities or components;
2. where continuous simulation used differential equations to describe the dynamic course of the simulation, discrete simulation uses process descriptions that may be different for all components;
3. where continuous simulation used a fixed time step to solve differential equations, discrete simulation allows any part of any process description to start or end at any time, abstaining from any subdivision of time into discrete steps.
4. where continuous simulation used differential equations to model "communication" between "components", discrete simulation uses detailed communication, possibly in the form of some standard structure for messages.
5. where continuous simulation used characteristics per flow, discrete simulation uses characteristics or attributes per individual component.

A more detailed elaboration of the example shown in figure B.3 will increase insight in the consequences of these differences.

Consider figure B.4 in which two time axes are shown, for two different processes: One time axis for the process of the train that is being unloaded, and one time axis for the gantry crane that is used for unloading. From this figure a strong interrelation between these two processes can be found, which is shown by means of two notifications that are passed between the train and the gantry crane. So, there is a kind of communication between both components, i.e. the train and the gantry crane. This need for communication will become more evident when the process descriptions for both types of components are given (or at least, a part of these process descriptions).

Figure B.4: Example of parallel processes for discrete simulation.

The process description for the train could be, for instance:

1. enter the track for handling;

---

2 like the temperature of a flow of milk.
2. notify gantry crane (communication →);
3. wait until unloaded (communication ←);
4. leave the track for handling.

while the process description for the gantry crane could be:

1. wait until a train is present and ready for unloading (communication ←);
2. while not all load units have been unloaded, do:
   a. request a machine for internal transport, if required (communication →);
   b. pick up the next load unit to unload;
   c. wait until a machine for internal transport is present (communication ←);
   d. place load unit on machine for internal transport;
   e. notify machine for internal transport that is can leave (communication →);
3. notify train that it is unloaded (communication →).

In these two process descriptions two directions for communication were shown: the receiving direction (which was marked as ←) and the sending direction (which was marked as →). To guarantee a smooth communication between all components in the simulation model some standard protocol must be designed and used. In case of TORCMOD a structure that is based on requests and notifications is adapted, as described in Van Zijnderveld [38] -- which is only one out of a wide field of varieties.

The great advantage of discrete simulation is the detailed information that can be obtained after simulation on the behaviour of the modelled system. This is mainly caused by the fact that individual components are discerned to which individual data can be attached like the requested time for unloading in the previous example, or the total time these load units have resided in the system -- a parameter that can not be determined on an individual basis if continuous simulation were used.

**Mixing continuous and discrete simulation**

In some cases an interesting mix of continuous and discrete simulation is used. An example hereof is a milk factory that is controlled by an operator that has to perform several tasks. As this operator can not perform more than one task at the time and has to walk from one control panel to the other, a very interesting interaction of continuous and discrete processes is found.

Even more, such a mix of continuous and discrete simulation allows to study the effects of certain decision rules and decision criteria that are used by the operator on, for instance, the number of micro-organisms that is found per litre of produced milk.
Appendix C. Optimisation techniques

C.1. Introduction

In a number of situations there is a need to find the best strategy out of an great multitude of possible scenarios. A typical example of such a situation is found in Hillier [9], page 192. This example reads:

(quote)
"The Metro District is an agency that administers the distribution of water in a certain large geographic region. The region is fairly arid so the District must purchase and bring in water from outside the region. The sources of this imported water are the Colombo, Sacron, and Calorie Rivers. The District then resells the water to use in its region. Its main customers are the water departments of the cities of Berdoo, Los Devils, San Go and Holyglass. It is possible to supply any of these cities with water brought in from any of the three rivers, with one exception that no provision has been made to supply Holyglass with Calorie River water. However, because of the geographic layouts of the viaducts and the cities in the region, the cost to the District of supplying water depends upon both the source of the water and the city being supplied. The variable cost per acre foot of water (in dollars) for each combination of river and city is given in table C.1. Despite these variations, the price per acre foot charged by the District is independent of the source of the water and is the same for all cities."

<table>
<thead>
<tr>
<th>River</th>
<th>City</th>
<th>Berdoo</th>
<th>Los Devils</th>
<th>San Go</th>
<th>Holyglass</th>
<th>Supply</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colombo River</td>
<td>16</td>
<td>13</td>
<td>22</td>
<td>17</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Sacron River</td>
<td>14</td>
<td>13</td>
<td>19</td>
<td>15</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Calorie River</td>
<td>19</td>
<td>20</td>
<td>23</td>
<td>--</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Min. needed</td>
<td>30</td>
<td>70</td>
<td>0</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Requested</td>
<td>50</td>
<td>70</td>
<td>30</td>
<td>(\infty)</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

*Table C.1: Water resources data for Metro Water District (Hillier, [9]).*

"The management of the District is now faced with the problem of how to allocate the available water during the upcoming summer season. Using units of 1 million acre feet, the amounts available from the three rivers are given in the right-hand column of table C.1. The district is committed to providing a certain minimum amount to meet the essential needs of each city (with the exception of San Go, which has an independent source of water), as shown in the *min. needed* row of the table. The *requested* row indicates that Los Devils desires no more than the minimum amount, but that Berdoo would like to buy as much as 20 more, San Go would go up to 30 more, and Holyglass will take as much as it can get"

"Management wishes to allocate all the available water from the three rivers to the four cities in such a way as to at least meet the essential needs of each city while minimising the total cost to the District".

(end quote)

This problem shows what type of problems is to be solved by optimisation techniques as described in this Appendix: Problems in which a limited amount of resources (i.e. water) is to be assigned to a number of targets (i.e. cities), taking into account restrictions on the available resources (i.e. supply per river) and on the bounds for assignment (i.e. requests per city).

A number of optimisation techniques is found in the field of operations research, each technique being developed for a special type of problem to solve. Only a small part of these available techniques will be described in this Appendix; for a description some more techniques refer to Domschke [5], Glover [8] and Hillier [9]. This section describes some of these optimisation techniques, which will be divided into two major groups: Linear programming and network methods. Three types of network methods are discussed: the transportation problem, the transshipment problem and generalised networks.
Appendix C. Optimisation techniques

Although very interesting and useful to apply for a wide field of real-life situations, optimisation techniques appear to be somewhat underestimated by most companies nowadays. Therefore some effort should be made to make these techniques more widely known and understood.

C.2. Linear programming

Linear programming is a technique that is used to find the best solution for problems as described in the above example. This best solution is recognised by the value of a special function that is defined for the problem to solve, which is called the goal function \( f(x) \) of the problem. The variable \( x \) is a vector that contains all variables that play a role in the decision and whose value needs to be determined. In terms of the above example, \( x \) contains the amounts of water allocated from one river to one city.

There are two ways to find the best solution for the goal function \( f(x) \): Minimisation which is written as \( \text{Min } f(x) \) and maximisation which is written as \( \text{Max } f(x) \). In case of minimisation a solution \( x \) is considered the optimal solution if the goal function has reached the minimum value it can ever reach. Therefore, minimisation is normally used in those cases where \( f(x) \) expresses the costs that are involved with a particular solution. In case of maximisation all is reversed: A solution \( x \) is considered the optimal solution if the goal function has reached the maximum value it can ever reach, and in those cases \( f(x) \) normally expresses the benefits that are involved with a solution \( x \).

Next to this goal function a number of constraints can be specified which limit the value range of some of the decision variables in \( x \). These constraints play an important role in linear programming as they have a great influence on the range of values that the goal function can reach. This will be clarified with a small example.

Suppose that we own a factory that produces two types of cookies, based on oat and wheat. The more expensive type of cookie, which is called "Princess", consists for 80% of wheat and for 20% of oat. Each sold "Princess" cookie gives us a revenue of US$ 0.06. The less expensive type of cookie, which is called "Conquistador", consists for only 60% of wheat and for 40% of oat. Each sold "Conquistador" cookie gives us a revenue of US$ 0.04. The total stock of the factory is sufficient for only 10,000 pure wheat cookies and 7,500 pure oat cookies. Now we want to know the optimal mix of "Princess" and "Conquistador" cookies to produce, or in other terms, the number of "Princess" and "Conquistador" cookies that we should produce to get the highest profit available, given the limitations of our stock. This is where we need linear programming.

In this case, our decision variables \( x \) of the linear programming problem could be defined as:

\[
\begin{align*}
    x_1 &= \text{number of produced "Princess" cookies} \\
    x_2 &= \text{number of produced "Conquistador" cookies}
\end{align*}
\]

The goal function, expressing the total profit as a linear function of the decision variables \( x \) would be:

\[
\text{Max } f(x) = 0.06 \cdot x_1 + 0.04 \cdot x_2 \quad \text{(C.1)}
\]

and the constraints, expressing the limited stock of wheat and oat, would be:

\[
\begin{align*}
    0.80 \cdot x_1 + 0.60 \cdot x_2 &\leq 10,000 \\
    0.20 \cdot x_1 + 0.40 \cdot x_2 &\leq 7,500
\end{align*}
\]

This problem can be solved by any linear programming package, and the resulting values of \( x_1 \) and \( x_2 \) are the optimal values of cookies to produce.

The advantage of linear programming is that it can be used to model almost any problem. The drawback is, that in some cases it is relatively slow in finding the optimal solution of a problem, compared to specialised algorithms that use special structures of a special type of problem. The two methods discussed next are such specialised algorithms.

- **Search space**
  The constraints for any optimisation problem limit the so-called search space, which is the multi-dimensional space in which the valid values for the decision variables lies. In case of \( n \) decision variables, the dimension of this search space equals \( n \), too.
Appendix C. Optimisation techniques

- **Infeasibility**
  In some cases the used optimisation model is not able to find a solution for all decision variables in the available search space; in such cases the search space is too limited and the formulated problem is called infeasible. In order to find a solution the problem must be re-formulated and be passed to the optimisation model again.

  The reason for infeasibility can normally be detected by inspecting the so-called dual variables of the optimisation problem. As this aspect goes beyond the scope of this Appendix it is referred to Hillier [9] for more detail on the subject.

- **Integer programming (IP)**
  Depending on the type of the decision variables a number of special optimisation problems is discerned. Most of the times the decision variables are all real numbers, which means that they can get any value that lies within the specified constraints. Sometimes the decision variables are all integer numbers, which means that they can get natural values only, thus limiting the number of possible solutions. This type of problems are called integer problems and are solved by integer programming. Normally, these problems are solved by specialised optimisation methods like branch and bound which will not be described here. Refer to Hillier [9].

- **Mixed integer programming (MIP)**
  Another type of problems is characterised by a mix of real-type and of integer-type decision variables. For these problems some specialised solution methods exist too (refer to Hillier [9]), but these methods in general require more calculation time than normal real-type linear programming because of discrete boundaries to search the values for the decision variables in, i.e. of the search space.

### C.3. Network methods

Although linear programming can be used to model almost any problem, it is not optimal for all problems. This is caused by a very special structure that can be found in a number of real-life problems that allows to use a more efficient, and therefore faster, algorithm to solve the problem. A wide variety of these specialised algorithms is available, but only one group of them will be discussed in this section: The group of network methods. These network methods exploit the special network structure of a problem.

A network, as defined in the field of operations research, consists of two elements only: Nodes and arcs\(^1\). An example of these two element types is given in figure C.1. Note that flows may run in either direction over an arc. A special type of arc is the so-called directed arc over which flows may run in one direction only, as is shown in figure C.1.

![Figure C.1: Definition of nodes and arcs.](image)

These flows run from one node to the other over (directed) arcs, and may represent any real-life property like the amounts of assigned water to cities in the above problem, or flows of money in a financial problem, or even flows of operators in an operator-job assignment problem. These examples of different types of flows readily show the wide field of application of network problems.

is a set of nodes that are connected to each other by arcs or directed arcs. Refer to figure C.1. Flows through the network move from one node to another node over another node. These flows can represent anything; for instance, flows can represent flows of water (as in the above problem), flows of money (in a purchase problem), flows of operators (in a operator assignment problem) or load units (in a terminal modelling problem). The difference between normal arcs and directed arcs is the direction of the flow: Normal arcs allow flows in both directions, while directed arcs allow flows in only one direction.

\(^1\)Note the re-definition of arcs that is given shortly!
Appendix C. Optimisation techniques

An example of a small network, consisting of five nodes and six (directed) arcs is given in figure C.2 in which flows run from node A to both nodes D and E. Next from this general network structure, figure C.2 also provides some information on the flows that are to run over the network. It specifies, for instance, that in total 20 units of flow (which could be either US dollars or operators) are available at node A. This is called the supply at node A. All these 20 supplied units of flow are to be removed from node A and are to be transported either to node D or to node E. At nodes D and E on the other hand, a number of units of flow is demanded instead of supplied, and therefore these nodes are called demand nodes instead of supply nodes. Note that the sum of the demands equals the sum of the supplies; in this way the network is in equilibrium which is a requirement to solve the problem.

Figure C.2 also shows that the number of units of flow that may run over arc C-E is restricted to only 3. Therefore, arc C-E is said to have a limited or maximum capacity of 3 units of flow.

Finally, figure C.2 shows that costs are involved by assigning flows to arcs A-B and A-C that equal $c_{A,B}$ and $c_{A,C}$ respectively. These are costs per unit of flow, and are used in the goal function of the entire network. This goal function of a network problem defines what the optimal solution is; in most cases this is the solution that results in the lowest (minimal) costs while still satisfying all demands on supplies, demands and maximum flows. In some cases the goal function is to find the minimum total costs of the network rather than the maximum costs.

Redefinition

In this thesis normal arcs as shown in figure C.1 are not used; all arcs in this thesis are of the directed type, and therefore the prefix "directed" is omitted in the remainder of this thesis. So, whenever it is spoken of "arcs" then this should be replaced by "directed arcs".

C.3.1. Transportation problem

The transportation problem is used to solve direct assignment of flows from a number of supply-nodes to a number of demand-nodes, where for each assignment the costs per unit of flow are known. This will be illustrated with a problem taken from Hillier [9], page 184.

Consider a situation in which cans of peas are to be transported from a number of canneries (which are the supply-nodes) to a number of warehouses (which are the demand-nodes) by means of direct trucking between these nodes. From a geographical map it can be found over what distance a truck must drive from one cannery to one warehouse. These distances are taken then as the costs of each assignment of a truckload of peas from one cannery to one warehouse. The aim of the optimisation is to find the optimal assignment of trucks to direct transports between canneries and warehouses, while in the meantime it must be guaranteed that (1) all peas that are produced at each cannery are assigned to some warehouse, and that (2) all demands for peas at all warehouses are satisfied.

In that case the problem could look as shown in table C.2 which is taken from Hillier [9]. This table shows:

- for each cannery the total production of peas in numbers of truckloads;
- for each warehouse the demand of peas in numbers of truckloads;
- for each direct transport between cannery and warehouse the costs per truckload.

These costs are mainly dependent on the distance between each pair of cannery and warehouse.

---

2 In some cases a minimum capacity may be used, too.
3 In reality all other arcs in the network will induce costs, too.
The goal function of this problem is to find the minimum total costs of all assigned flows of peas. The result of the optimisation must respect the required allocation of peas to warehouses and the total production of peas at canneries.

<table>
<thead>
<tr>
<th></th>
<th>Warehouse</th>
<th></th>
<th></th>
<th></th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cannery</td>
<td>1</td>
<td>464</td>
<td>513</td>
<td>654</td>
<td>867</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>352</td>
<td>416</td>
<td>690</td>
<td>791</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>995</td>
<td>682</td>
<td>388</td>
<td>685</td>
</tr>
<tr>
<td>Allocation</td>
<td></td>
<td>80</td>
<td>65</td>
<td>70</td>
<td>85</td>
</tr>
</tbody>
</table>

Table C.2: Shipping costs ($) per truckload (Hillier, [9]).

This problem could also be shown as a network in which flows three nodes model the canneries (supply-nodes) and four other nodes model the warehouses (demand-nodes). Each cannery is connected to each warehouse by means of a separate arc. To each arc costs are assigned that are equal to the ones shown in table C.2.

C.3.2. Transhipment problem

A little more complicated than the transportation problem is the transhipment problem. While the transportation problem only modelled direct transportation between supply- and demand-nodes, the transhipment problem also models transport between supply- and demand-nodes over an intermediate node. These intermediate nodes are transhipment nodes. This is shown in figure C.4.

A problem can be formulated as a transhipment problem if there is some intermediate action at which flows are first gathered and are then split again into separate flows. Any costs that are induced at the transhipment nodes themselves rather than on the transportation arcs to and from these nodes can be modelled by adding these costs to all arcs that either run to, or run from those collection nodes.

If the example given in table C.2 was to be expanded by including some distribution centres for peas at strategic locations between the canneries and warehouses, then it could be found which numbers of truckloads are to be transported to and from each distribution centre, to obtain minimal total costs.

The special network structure shown in figure C.4 is used by the algorithm of the transhipment problem to increase the speed of solving the optimisation problem, compared to linear programming. Refer to Hillier [9] for more details on this special structure. A major drawback of the transhipment problem is that it can only be used for quite simple problems. One of the reasons for this is that it can handle only simple network-balance constraints.

C.3.3. Generalised networks

The last specialised variant of linear programming to discuss are the generalised networks. These generalised networks can be considered as an improved version of the transhipment problem. The main difference between generalised networks and the transhipment problem is the availability of multipliers. Refer to Glover [8]. These multipliers are used for example to convert the products that flow through the network from one type into another type, or to convert the unit of the products that flow through the network. These multipliers are indicated by triangles on the arcs.

An example of such a multiplier is given in figure C.5. The nodes in this figure represent a bank account at a specific month, and the mid arc represents the transference of money that is not spent in month I and becomes available for use at the very start of month II. The multiplier represents the interest of 4% a month.

Figure C.4: Example of a transhipment network.

Figure C.5: Example of a multiplier.
Appendix C. Optimisation techniques

Note especially that this network structure does not allow to draw money from the bank at the half of the month, nor to deposit money at another moment than at the start of the month. Although this problem could be fixed by splitting the arc of figure C.5 into two separate, subsequent arcs that represent the first and second half of the month respectively, still some discreteness in time exists. This is a common problem of optimisation models, and cannot easily be solved. This aspect will also turn out to be important in Chapter 7 when discussing the equipment optimisation model named "equipment hopping".

Some generalised network packages also allow the use of side-constraints, which are all constraints other than the normal network balancing constraints. Refer to Glover [8] for more details. The more of these side-constraints are used, the more the network problem turns into a normal linear programming problem. An example of such a side-constraint is given in figure C.6. It shows a small network that forces the flow on arc 1-2 to be 1.5 times the flow on arc 1-3. This kind of constraint can not be included in a network model other than by using side-constraints. These side-constraints will appear to play an important role in the just mentioned equipment optimisation model as described in Chapter 7.

Figure C.6: Example of a side-constraint.
Appendix D. List of symbols

This Appendix gives an overview of all variables used in the model described in this thesis. Each variable name is coded such that its meaning can be derived quickly. The initial letters of each variable name indicate the type of model element the variable is related to; the meaning of these initial letters can be found at each section of the variables list. The final letter of each variable name indicates the type of variable according to the coding shown in table D.1. The remaining letters give some more specific information on the variable, which will be clear from the respective description. The indices that may be used for variables are shown in table D.2 along with their meaning. The units that are used for the variables are shown in table D.3.

- **Reference variables**
  One special type of variables are reference variables. As mentioned in table D.1 these variables end on the letter r and contain a reference to a variable of another type. Consider, for example, an equipment pool p that is of equipment type t. In that case equipment pool p refers to equipment type t. In other terms, when the type of equipment pool p is stored at the variable ep_tr_p, then the value of would ep_tr_p simply equal t. Another feature of these variables is, that properties of the referred variable are easily accessed in following way:
  \[ \text{et_ce(ep_tr_p)} = \text{carrying capacity of the equipment type of equipment pool p}, \]
  provided that the variable et_ce specifies the carrying capacity of a specific equipment pool.

- **Temporary variables**
  At some points in this thesis variables are printed in boldface and italics, like the variable dx_reloc. Any variable that is printed thus is a temporary variable that is used during calculations only and is not a normal model parameter. Therefore, these temporary variables are defined in their context only; in some cases the same temporary variable name may appear in different contexts. In those cases the temporary variables have different meanings and may not be confused.

<table>
<thead>
<tr>
<th>Final letter</th>
<th>Abbreviation</th>
<th>Variable type</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>capacity, or costs</td>
<td>Capacity in number of generalised load units, or Costs in monetary units</td>
</tr>
<tr>
<td>f</td>
<td>factor</td>
<td>Factor without dimension</td>
</tr>
<tr>
<td>n</td>
<td>number</td>
<td>Number of load units or machines of an equipment pool</td>
</tr>
<tr>
<td>r</td>
<td>reference nr.</td>
<td>Reference number (ref. of equipment pool to equipment type)</td>
</tr>
<tr>
<td>p</td>
<td>preference</td>
<td>Preference assigned to some action related to REMBRANDT</td>
</tr>
<tr>
<td>o</td>
<td>option</td>
<td>Option that can either be selected or not (boolean-type; yes-no)</td>
</tr>
<tr>
<td>t</td>
<td>time</td>
<td>Time of an occurrence or duration of an action</td>
</tr>
<tr>
<td>v</td>
<td>velocity</td>
<td>Velocity of a movement</td>
</tr>
</tbody>
</table>

Table D.1: **Explanation of the final letters of variable names.**

<table>
<thead>
<tr>
<th>Index</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td>Location</td>
</tr>
<tr>
<td>o</td>
<td>Origin of a pool before relocating</td>
</tr>
<tr>
<td>od</td>
<td>Origin of a trailer-type pool in case of double relocation of pulling equipment</td>
</tr>
<tr>
<td>p</td>
<td>Equipment pool</td>
</tr>
<tr>
<td>r</td>
<td>Equipment routing</td>
</tr>
<tr>
<td>s</td>
<td>Section on a certain equipment routing</td>
</tr>
<tr>
<td>w</td>
<td>Time window</td>
</tr>
</tbody>
</table>

Table D.2: **Explanation of indices of variables.**
### Appendix D. List of symbols

<table>
<thead>
<tr>
<th>Unit</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>Time</td>
</tr>
<tr>
<td>d</td>
<td>Distance</td>
</tr>
<tr>
<td>gen-LU</td>
<td>Generalised load units</td>
</tr>
<tr>
<td>LU</td>
<td>Load units</td>
</tr>
<tr>
<td>MU</td>
<td>Monetary unit</td>
</tr>
<tr>
<td>y/n</td>
<td>Yes/No</td>
</tr>
<tr>
<td>-</td>
<td>Dimensionless</td>
</tr>
</tbody>
</table>

#### Table D.3: Explanation of units for the variables.

<table>
<thead>
<tr>
<th>Variable type</th>
<th>Variable name</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>ep = Equipment pool</td>
<td>ep_tr</td>
<td>equipment type reference</td>
<td>[-]</td>
</tr>
<tr>
<td></td>
<td>ep_sp</td>
<td>preference to use a pool at a section as stored at the equipment routings; may be a fixed preference</td>
<td>[-]</td>
</tr>
<tr>
<td></td>
<td>ep_ac</td>
<td>costs involved with the assignment of a pool to transport at a section per unit of flow</td>
<td>[MU]</td>
</tr>
<tr>
<td></td>
<td>ep_cbto</td>
<td>option to couple to trailer-type equipment before transhipment; only valid for pulling equipment</td>
<td>[y/n]</td>
</tr>
<tr>
<td></td>
<td>ep_wt</td>
<td>working time of a pool in a time window</td>
<td>[t]</td>
</tr>
<tr>
<td></td>
<td>ep_wtf</td>
<td>working time factor for an equipment pool in a time window</td>
<td>[-]</td>
</tr>
<tr>
<td></td>
<td>ep_tt</td>
<td>transport time factor for an equipment pool on a section from a certain origin</td>
<td>[t]</td>
</tr>
<tr>
<td>epa = Equipment pool assignment</td>
<td>epa_n</td>
<td>number of machines used for transport on a section of an equipment routing in a time window</td>
<td></td>
</tr>
<tr>
<td>et = Equipment type</td>
<td>et_cc</td>
<td>carrying capacity</td>
<td>[gen-LU]</td>
</tr>
<tr>
<td></td>
<td>et_at</td>
<td>time to attach to a load unit</td>
<td>[t]</td>
</tr>
<tr>
<td></td>
<td>et_ut</td>
<td>time to get up a load unit</td>
<td>[t]</td>
</tr>
<tr>
<td></td>
<td>et_dt</td>
<td>time to put down a load unit</td>
<td>[t]</td>
</tr>
<tr>
<td></td>
<td>et_rt</td>
<td>time to release a load unit</td>
<td>[t]</td>
</tr>
<tr>
<td></td>
<td>et_uat</td>
<td>acceleration time when unloaded (or empty)</td>
<td>[t]</td>
</tr>
<tr>
<td></td>
<td>et_lat</td>
<td>acceleration time when loaded</td>
<td>[t]</td>
</tr>
<tr>
<td></td>
<td>et_udt</td>
<td>deceleration time when unloaded (or empty)</td>
<td>[t]</td>
</tr>
<tr>
<td></td>
<td>et_ldt</td>
<td>deceleration time when loaded</td>
<td>[t]</td>
</tr>
<tr>
<td></td>
<td>et_ptr</td>
<td>pulling-type reference (if required)</td>
<td>[-]</td>
</tr>
<tr>
<td></td>
<td>et_uct</td>
<td>uncouple time from any non-self propelled machine</td>
<td>[t]</td>
</tr>
<tr>
<td></td>
<td>et_cpt</td>
<td>couple time to any non-self propelled machine</td>
<td>[t]</td>
</tr>
<tr>
<td></td>
<td>et_uv</td>
<td>driving speed when unloaded (or empty)</td>
<td>[d/t]</td>
</tr>
<tr>
<td></td>
<td>et_lv</td>
<td>driving speed when loaded</td>
<td>[d/t]</td>
</tr>
<tr>
<td>fp = Flow pattern</td>
<td>fp_p</td>
<td>preference for presence at location l of a load unit running over routing r</td>
<td>[-]</td>
</tr>
<tr>
<td>lc = Location</td>
<td>lc_sn</td>
<td>number of load units that is stacked at location l for equipment routing r at the end of time window w</td>
<td>[LU]</td>
</tr>
<tr>
<td></td>
<td>lc_lse</td>
<td>limited stack capacity at this location ?</td>
<td>[-]</td>
</tr>
<tr>
<td></td>
<td>lc_lscn</td>
<td>limited stack capacity</td>
<td>[LU]</td>
</tr>
</tbody>
</table>

Page 254
| \( \text{lu} = \text{Load unit} \) | \( \text{index: window, [pool, section], [location]} \) | \( \text{lu}_\text{tn} \) | number of load units that are transported on a section of an equipment routing by an equipment pool | [LU] |
| \( \text{lu}_\text{dltu} \) | number of dummy-load units that are transported on a section of an equipment routing by an equipment pool | [LU] |
| \( \text{lu}_\text{sn} \) | number of load units that are stored at a location (and therefore is not transported) | [LU] |
| \( w = \text{Time window} \) | \( \text{index: window} \) | \( w_\text{st} \) | start time of time window | [t] |
| \( w_\text{et} \) | end time of time window | [t] |
| \( w_\text{dt} \) | duration of time window (equals end minus start) | [t] |

Table D.4: List of symbols.
Appendix D. List of symbols
References


References


29. SVV II: "Structuurschema verkeer en vervoer; Deel D; De regeringsbeslissing; Verkeer en vervoer in een duurzame samenleving", 1990.


Abbreviations and terminology

Abbreviations

AHP  Analytic Hierarchy Process
     Multi criteria decision technique as defined by Saaty.

AGV  Automated Guided Vehicle
     Equipment type; driver-less machine for internal transport that normally cannot tranship.

ASC  Automated Stacking Crane
     Equipment type; automated version of the RMG.

CATD Computer Aided Terminal Design
     Method to design a (rail-) terminal using a computer model.

D-loc  Diesel locomotive
     Diesel powered locomotive; is mainly used locally at a rail service centre.

DDM  Detailed Discrete Model
     Highest level of detail for the described terminal design aiding tool.
     Refer to UCM and SDM.

E-loc  Electrical locomotive
     Electrical powered locomotive; is mainly used at the main rail network and only sometimes for
     local train movements at a rail service centre.

ETA  Expected Time of Arrival
     Time at which a train of vessel is expected to arrive at some location.

ETD  Expected Time of Departure
     Time at which a train of vessel is expected to leave from some location.

FGC  Fixed Gantry Crane
     Equipment type; gantry crane that can not drive.

MCA  Multi Criteria Analysis
     Technique to combine scores on various criteria for all alternatives into one overall score. Refer
     to Human assessment and Choice randomness.

MCDA  Multi Criteria Decision Aid
     French group of multi criteria analysis techniques. This group focuses on binary outranking
     relations.

MCDM  Multi Criteria Decision Making
     American group of multi criteria analysis techniques. This group focuses on utility functions to
     combine all scores on all criteria to an overall score per alternative.

MLU  Multiple Load Unit (equipment type)
     Group of equipment types that is able to carry more than one load unit at the time during internal
     transport.

MTS  Multi Trailer System
     Equipment type; set of about five coupled trailers that is pulled by a specialised terminal truck.

UCM  Unrestricted Continuous Model
     Lowest level of detail for the described terminal design aiding tool.
     Refer to SDM and DDM.

RTG  Rubber Tyred Gantry
     Equipment type; gantry crane that is able to drive on rubber tyres.
Abbreviations and terminology

RS  Reach Stacker
    Equipment type; freely movable machine for internal transport that can tranship.

RMG  Rail Mounted Gantry
    Equipment type; gantry crane that is able to drive on special rail tracks.

SC  Straddle Carrier
    Equipment type; freely movable machine for internal transport that can tranship and moves as an arc over containers.

SDM  Simplified Discrete Model
    Mid level of detail for the described terminal design aiding tool.
    Refer to UCM and DDM.

TT  Terminal Trailer
    Equipment type; trailer that is specifically fit for internal transport.

SMART  Simple Multi-Attribute Rating Technique
    Technique for MCA in combination with human assessment that uses direct rating.

OR  Operations Research
    Group of various techniques to find an optimal solution for a generally complex problem on e.g. scheduling, assignment or planning.

TORCMOD  Transhipment Of Rail Containers modelling tool
    Implementation of the proposed computer terminal design method as described in this thesis, running under Windows™.

Terminology

Arc
    Connection between two nodes in a network.

Broken internal transport
    Internal transport by one equipment pool over two subsequent sections that are separated by a section break which does not represent any kind of stacking, consisting of two separate transport movements. Broken internal transport is considered an invalid situation and is always converted into closed internal transport, ignoring the section break.

Bundle
    Group of functionally related and geographically very near rail tracks that need to be considered as a unity.

Closed internal transport
    Internal transport by one equipment pool over two subsequent sections that are separated by a section break which does not represent any kind of stacking, merged into one single transport movement. Closed internal transport can be created (1) by installing equal equipment pools at both sides of an exchange point, or (2) by expanding a private network over an (implicit) exchange point.

Closed task
    Connection between two locations at the system’s boundary that represents a route for load units flowing through the modelled system

Contra schedule
    Abstract counterpart of load schedules that do specify information on timing of transport across the system’s boundary, but do not specify detailed information on numbers and types of load units to cross the system’s boundary.

Equipment break
    Discontinuity in internal transport over an equipment routing. An equipment break can be introduced by an internal stack, an exchange point or by the first of last location of the equipment routing.
Equipment hopping
Optimisation method to find (1) the number of machines to install at a terminal complex, and (2) the planned usage of equipment during simulation of the modelled terminal complex.

Handle bundle
Bundle at which load units can be transhipped onto and from trains.

Internal transport
All transport from one location at the modelled terminal yard to another location at the modelled terminal yard by means of transport equipment that is owned by the modelled terminal complex and all stacking movements, all equipment relocation movements, and all planning of all these movements.

King pin
Pin that is created automatically for each location that has a private network, and is used as the "root" of that private network.

Light engine
Locomotive that is not coupled to a train.

Load schedule
Schemes for arrival and departure of load units at the system's boundary.

Load volume
Batch of load units included in a load schedule.

Local internal stacking
Stacking at locations other than internal stacks.

Local internal transport
Internal transport within a location rather than between locations.

Local section
Section within a location, related to that location's private network.

Machine
Piece of equipment.

Matching load units to load schedules
Process to find for each individual load unit that is created according to a load schedule at one side of the modelled terminal complex the corresponding contra schedule at the other side of the modelled terminal complex.

Node
Location in a network between from which arcs originate, or at which arcs terminate.

Normal internal stacking
Stacking at internal stacks, in contrast to local internal stacking.

Normal internal transport
Internal transport between two locations, in contrast to local internal transport.

Normal section
Section between two subsequent locations, related to public networks or to expanded private networks.

Pin
Node of an equipment network which either can be freely positioned, or can be pinned to an UCM-location or a rail point.

Private network
Network that is part of a location and is created automatically.

Public network
Network that is not part of a location and is created by the user of the terminal design aiding tool.
Abbreviations and terminology

Section for equipment
Indivisible way for internal transport for which any of the equipment pools installed for that section can be used.

Section for railways
Part of rail infrastructure that is isolated from the rest of rail infrastructure by means of welds with or without signals.

Yard bundle
Bundle at which load units can not be transhipped onto and from trains.
Index

A
Additional points. See Locations
AGV, 259
AHP, 259
Alternative, 225
Analytic Hierarchy Process
absolute assessment, 229
Arcs. See Networks
ASC, 259
Assessment. See Human assessment

B
Bedienverfahren, 173, 182, 185, 191
Bounding elements. See Locations

C
Choice, 225
Clients, 18, 23, 26, 44, 168, 186
Closed tasks, 25, 77, 97, 260
Contra schedules, 25, 46, 80, 104, 260
examples, 67
intensity, 25, 27, 51, 57
matching, 53, 261
regular, 25, 26, 27, 48, 54
related to specific closed tasks, 49
timing, 60
Criterion, 225

D
DDM, 72, 259

E
Equipment, 3, 9, 13, 71
insertion point, 75
machine, 71, 73, 261
network, 75, 76, 78
blocking, 80, 102
link, 76
overlapping, 77, 80
pin, 76, 261
king pin, 77, 79, 261
pinned pin, 76, 78
private, 77, 79, 261
public, 77, 79, 261
service range, 72, 76
pool, 75, 78
capacity, 75
fixed valuation, 75, 80
operator, 73, 75, 76
overlapping, 77, 80
working hours, 72, 75
relocation, 87, 102
routing, 29, 45, 47, 75, 77, 170, 215
break, 78, 86, 97, 260
duration, 29, 80
preferences, 48, 72, 77
section, 78, 86, 93, 262
local, 78, 87, 261
normal, 78, 261
supply node, 119, 142, 143
type, 72
multiple load unit, 74, 124
Equipment optimisation, 71, 88, 261
knowledge horizon, 63, 90, 94, 96
maximal resolution, 93
optimisation model, 71, 194
inter-window equipment hopping, 95, 150
minimum-ratio threshold, 153
INTRA-window equipment hopping, 95, 96
optimistic plan effect, 94, 95
planning horizon, 91, 93, 104
planning interval, 88, 91, 95, 157
planning phases, 93
routing transport time, 92, 170
section transport time, 92, 93, 161
time windows, 90, 93, 103
final window, 104
initial window, 104
micro window, 93
Examples
cookies factory problem, 248
crane selection problem, 226
elephant feeding problem, 238
housing problem, 230
milk barrel problem, 243
pool assignment problem, 237
stack block selection problem, 225, 229
water allocation problem, 247
Exchange points. See Locations

F
FGC, 259
Fließverfahren, 173, 182, 184, 186, 193
Flow data, 23, 25
closed tasks, 25, 77, 97, 260
contra schedules, 25, 46, 80, 104, 260
examples, 67
intensity, 25, 27, 51, 57
matching, 53, 261
regular, 25, 27, 48, 54
related to specific closed tasks, 49

1References to explanation of abbreviations are shown in boldface.
Index

Index

timing, 60
flow patterns, 45, 48, 71, 81, 88, 104, 105, 170, 214
flat, 82
mirrored, 82
load schedules, 25, 42, 104, 178, 194, 261
timing, 60
load volumes, 25, 44, 97, 189, 194, 261

G
Generalised load units, 49, 74, 164
Generalised networks, 251

H
Human assessment, 79, 225, 226
methods
absolute assessment, 226, 229
interval assessment, 226, 229
relative assessment, 226
REMBRANDT scale, 79, 81, 168, 229, 230, 236
risk behaviour, 237
Human selection. See Random selection

I
Internal stacks. See Locations
Internal transport, 71, 261
at bounding elements, 82
at internal stacks, 84
at train handle areas, 83
handle private network, 83, 166
main private network, 83, 166
broken, 86, 99, 112, 260
closed, 86, 100, 104, 113, 260
in DDM level of detail, 32
in SDM level of detail, 29
in UCM level of detail, 24
local, 82, 98, 99, 100, 101, 261

L
Linear programming, 248
Load schedules, 25, 42, 104, 178, 194, 261
timing, 60
Load units, 18, 23, 26
design class, 18, 168
height class, 18
length class, 18
special class, 18, 168
weight class, 18
width class, 18
Load volumes, 25, 44, 97, 189, 194, 261
Local internal transport. See Internal transport
Local stacking. See Stacking
Locations
additional points, 24, 78, 97
bounding elements, 24, 97
exchange points, 24, 78, 92
internal stacks, 24, 78, 97
train handle areas, 24, 97
Logarithmic regression, 234, 235, 236

M
MCA, 259. See Multi criteria analysis
MCDA, 231
MCDM, 231
MLU, 259. See Equipment-Type-Multiple load unit
Model, 20, 71
user, 20, 71
MTS, 259
Multi criteria analysis, 225, 230
alternative, 225
choice, 225
criterion, 225
methods
American, 231
AHP, 233
AHP, multiplicative, 168, 233, 234
SMART, 233
French, 231
ELECTRE I, 232
interactive, 231
proposal, 225
score, 225
utility function, 231
Multipliers, 251

N
Networks, 249, 250
arc, 260
arcs, 249
costs, 250
demand, 250
directed arcs, 249
flows, 249
goal function, 248, 250
maximum capacity, 250
minimum capacity, 250
nodes, 249, 261
supply, 250

O
Operations research, 247
Optimisation
dual variables, 249
generalised networks, 247
infeasibility, 155, 249
integer programming, 249
linear programming, 247
mixed integer programming, 249
networks, 247
search space, 155, 248
transhipment problem, 247, 251
transportation problem, 247, 250

page 264
P
Parameters. See Variables
Paths. See Physical paths. See Physical paths
Perron-Frobenius eigenvector, 228, 234
Physical paths, 24
Proposal, 225

R
Rail. See also Trains
infrastructure
  block safety systems, 31, 201
  rail path, 202, 262
  rail section, 202
  section, 201
bundles, 172, 197, 260
  handle, 172, 198, 261
  selection, 197, 205
  service level, 29, 36, 37, 198
  yard, 262
claiming, 202
  ad hoc, 203
  integral, 204
  planned, 203
  stepwise, 204
points, 198
routes, 201
  route manager, 201
  sawing, 201
  selection, 201
tracks, 200
operations
  brake tests, 174, 182
  great, 183
  small, 183
  coupling, 174
handling, 172, 184
  grazing, 185, 186, 190
  random, 186
in DDM level of detail, 31
in SDM level of detail, 28
light engine, 261
light engines, 172, 182
loading
  concurrent, 185, 186, 187, 188
  separate, 185, 186, 187
loc storage, 175, 183
relocation, 173, 187
Random seed, 34
Random selection, 55, 225, 237
exponential growth, 238
exponential growth factor, 238, 239
linear, 238
quadratic, 56, 238
skip level, 56, 238, 239
zero-based option, 238, 239
  entrance smoothing, 240, 241
  re-scaling, 240
RMG, 260
RS, 260
RTG, 259

S
Saaty eigenvector method, 228, 234
SC, 260
Score, 225
SDM, 71, 72, 76, 260
Side constraints, 95
Side-constraints, 252
SMART, 260
Stacking
  at bounding elements, 85
  at internal stacks, 85
  at train bundles, 85
  at train handle areas, 85
    handle private network, 83, 166
    main private network, 83, 166
  in SDM level of detail, 31
  in UCM level of detail, 24
  initial stack size, 104, 169
  limited stacking condition, 85, 101, 164, 165
  local, 84, 98, 99, 100, 101, 261
Standard load units. See Load units
System, 20, 71
  boundary, 19, 23, 24, 71, 78

T
Terminal complex, 18
Timing of load units, 60, 193
  by contra schedules
    intensity, 63, 65, 67
    regular, 61
  by load schedules
    at bounding elements, 60
    at train handle areas, 61, 193
  by trains, 193
TORC MOD, 19, 260
Train handle areas. See Locations
Trains, 171. See also Rail
  concepts, 175, 192
    D-loc mode, 183
    E-loc mode, 182
  handling mode, 184
    absolute positions on-train, 186, 188
    condensed, 188, 190
    equally spread, 189, 190
    concurrent loading, 186
    freedom in load schemes, 186, 187
    priority to loading, 188
    priority to unloading, 188
    use nearest free location, 187
    relative moment of handling, 186, 191
      grazing, 191
      random, 191
    relative positions on-train, 186, 189
random sorting, 190
  sorting on clients, 190
  sorting on destination, 190
  sorting on origin, 190
separate loading, 186
shuttle types, 175, 178
  yard mode, 181
length, 191, 196
locomotives, 182, 183
  automatic creation, 208
  controls, 208
  creation threshold, 208
  D-and E-loc priority, 208
  initial presence, 208
  loc manager, 208
  selection method, 208
phantom train, 204
safety margin, 193
services, 191
  numbers of load units, 193
  system entry track, 192, 200
  system exit track, 192, 200
  timing, 191
shuttle types, 175, 178
  pure, 176, 178
    single RSC selection, 178
  split, 177, 181, 190
  travel, 177, 179
    inter-RSC relocation, 180
    load mode, 181
    RSC sequence, 179
speeds, 183
TT, 260

U
UCM, 76, 259
  physical paths, 24
User, 20, 71

V
Variables, 253
  indices, 253
  referenced, 253
  temporary, 253
  units, 254