

Optimising load planning and container routing in intermodal rail transport

Citation for published version (APA):

Heggen, H. C. J. (2019). *Optimising load planning and container routing in intermodal rail transport*. [Doctoral Thesis, Maastricht University, Universiteit Hasselt]. Maastricht University. <https://doi.org/10.26481/dis.20190619hh>

Document status and date:

Published: 01/01/2019

DOI:

[10.26481/dis.20190619hh](https://doi.org/10.26481/dis.20190619hh)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Optimising load planning and container routing in intermodal rail transport

DISSERTATION

to obtain the degree of Doctor of Business Economics at Hasselt University, on the authority of the Rector Prof. dr. Luc De Schepper, and the degree of Doctor at Maastricht University, on the authority of the Rector Magnificus, Prof. dr. Rianne M. Letschert

in accordance with the decision of the Board of Deans,

to be defended in public
on Wednesday 19 June 2019 at 14:00 hours in Maastricht

by

Hilde Cecile Johanna Heggen

Supervisors

Prof. dr. A. Caris (Hasselt University)

Prof. dr. ir. C.P.M. van Hoesel (Maastricht University)

Co-supervisors

Prof. dr. K. Braekers (Hasselt University)

Dr. A. Berger (Maastricht University)

Assessment Committee

Prof. dr. A.J. Vermeulen (Chairman, Maastricht University)

Prof. dr. K. Ramaekers (Hasselt University)

Prof. dr. S. Limbourg (Université de Liège)

Prof. dr. T. Wauters (KULeuven – Campus Gent)

This thesis is part of the PhD thesis series of the Beta Research School for Operations Management and Logistics (onderzoeksschool-beta.nl) in which the following universities cooperate: Eindhoven University of Technology, Maastricht University, University of Twente, VU Amsterdam, Wageningen University and Research, and KU Leuven.



Acknowledgements

In April 2015, I started my PhD journey thanks to the persuasion and perseverance of some of my current and former colleagues. I am so glad that you convinced me to start this journey, as it turned out that I truly enjoyed the experience. Throughout the past four years, I have grown in so many ways, both personally and professionally. This would never have been possible without the guidance and support of many people.

First and foremost, I would like to thank my supervisors at Hasselt University, Prof. dr. An Caris and Prof. dr. Kris Braekers. Thank you for being my perfect supervising duo. You were invaluable to me and to the result of this PhD thesis. An, thank you for convincing me (and not giving up in the process) to take the step towards starting a PhD. You inspired me to start this PhD journey, and I did not regret that decision for a single day. You asked the right questions and always kept focus on the practical and academic relevance of my research. Thanks to your positive feedback, I left every meeting with confidence. Kris, you never run out of ideas. Thank you for all the constructive feedback and valuable insights, which every time increased the quality of my papers. Furthermore, I would like to thank my supervisors at Maastricht University, Prof. dr. ir. Stan van Hoesel and dr. Andre Berger, for their advice and suggestions during my stays in Maastricht. I also want to thank Prof. dr. Katrien Ramaekers and dr. Tjark Vredevelde for their regular feedback and relevant questions as my commission members, as well as the members of my assessment committee and jury, whose valuable comments and questions contributed to this dissertation. I would like to thank Luc Driessen, An Bonneux and Gianni Martella of Move Intermodal for providing me with interesting research ideas with strong practical relevance, for the valuable comments, feedback and insights.

I would like to express special gratitude to the research group Logistics for creating the optimal environment to conduct research of academic and practical relevance in an unbelievably encouraging atmosphere with the best colleagues I could ever imagine. Thanks to every person who crossed my path through the past years of my PhD, I cannot remember a single day I did not have enough energy to start my working day. I cannot even start thanking you individually, as each of you made my journey worthwhile in your own way.

Finally, I am sincerely grateful to my family for their unconditional support. Thank you, mom and dad, for supporting me in every decision I ever made, for making fresh orange juice, for picking up my online shopping packages and for providing me with everything I needed to pursue my goals. And lastly but most importantly, my fiancé Christian, my problem solver. You help me overcome all difficulties I encounter. Thank you for your patience (even when I lost mine), and for putting things in perspective when my stress levels were too high. I look forward to finishing our home together. I am so fortunate with you by my side, thank you for making me happy.

Hilde Heggen

June 2019

Abstract

With the rising focus towards more sustainable transport systems, intermodal transport is a promising alternative for unimodal road transport. However, it also presents a number of challenges for intermodal operators and planners who are in charge of executing transport planning in an increasingly dynamic environment. As multiple transport modes and decision makers are involved, intermodal transport is more complex. Moreover, the level of integration of different transport modes and decision levels influences the attractiveness of intermodal transport. Innovative and integrated transport systems reflecting real-life problems should be developed in order to provide adequate decision support tools for intermodal planners.

This dissertation aims at offering intermodal planning support in order to minimise total transport costs and maximise service capacity utilisation, which in turn decreases costs of the transport system. Two decision support tools are proposed by means of fast planning algorithms which include real-life characteristics. These concepts are inherent to the synchromodal vision in order to encourage a modal shift away from unimodal road transport. After a literature review on intermodal routing and vehicle routing in intermodal transport, two problems usually considered separately, an integrated intermodal routing problem is presented. By including information of local vehicle routes in the assignment of transport requests to long-haul routes through a service network, better-informed decisions can be obtained. Insights are presented on how such integrated approach can be used in practice by means of a real-life case. With transport requests assigned to intermodal routes through an intermodal rail service network, they should be assigned to specific locations on an intermodal train, which is the train load planning problem. It includes detailed loading restrictions, as omitting relevant loading constraints could result in infeasible solutions. Both an exact and a heuristic solution approach with multiple objectives and additional real-life loading constraints are presented, resulting in multiple feasible train load plans. Routes might still change for some transport requests and information about future orders should be accounted for. By providing multiple plans, planners can select the most appropriate load plan at a specific moment in time. The problem is applied to a real-life case to demonstrate advantages for practitioners.

Contents

List of Tables	v
List of Figures	vii
I Introduction	1
1 Introduction and problem statement	3
1.1 Intermodal freight transport	3
1.2 Decision support	5
1.3 Research objectives	6
1.4 Thesis outline	8
II Intermodal container routing	11
2 Integrating local drayage and intermodal long-haul routing: related literature and problem formulation	15
2.1 Introduction	15
2.2 Related literature	18
2.2.1 Long-haul service network design and intermodal routing . . .	18
2.2.2 Pre- and end-haulage transport	23
2.2.3 Research opportunities	25
2.3 Problem description	27
2.4 Problem formulation	29
2.4.1 The sequential approach	33
2.4.2 The integrated intermodal container routing problem	39
2.5 Conclusions	42

3	A large neighbourhood search heuristic for the integrated intermodal routing problem	45
3.1	Introduction	45
3.2	Large neighbourhood search heuristic	46
3.2.1	General structure of the LNS	48
3.2.2	Integrated operators	50
3.2.3	Preliminary checks for inserting a node	53
3.2.4	Feasibility of new routes	53
3.3	Comparison of the sequential and integrated approach	57
3.3.1	Generated instances	57
3.3.2	Heuristic parameters	60
3.3.3	Experimental results	68
3.4	Case study: tactical service network design decisions	73
3.4.1	The impact of changes in the service network: removing long-haul services with small capacity	74
3.4.2	The impact of congestion around terminals	76
3.4.3	Managerial insights	79
3.5	Conclusions and future research	80
III	Train load planning	83
4	Factors influencing the train load planning process: literature review and classification	87
4.1	Introduction	87
4.2	Classification of train load planning problems	90
4.2.1	Objectives	90
4.2.2	Constraints	93
4.2.3	Planning environment	97
4.2.4	Related problems	99
4.3	Conclusions and research opportunities	100
5	A multi-objective train load planning problem with real-life characteristics: formulation and exact solution approach	103
5.1	Introduction	103
5.2	Problem definition	105
5.2.1	Introduction to multi-objective concepts	105
5.2.2	Multi-objective model formulation	106

5.2.3	Adding practical considerations	110
5.3	A multi-objective exact solution approach	112
5.4	Problem instances	115
5.5	Computational experiments for the ϵ -constraint method	117
5.6	Conclusions	120
6	A heuristic solution approach for multi-objective train load planning	123
6.1	Introduction	123
6.2	Multi-directional local search heuristic	125
6.2.1	General structure of the MDLS	126
6.2.2	Constructive phase	127
6.2.3	Local search phase	127
6.2.4	Quality indicators	131
6.3	Parameter tuning	132
6.3.1	Instance classes	133
6.3.2	Parameters of the MDLS	133
6.3.3	Iterated racing procedure (<i>irace</i>)	135
6.4	Parameter sensitivity analysis	138
6.4.1	Interaction between $n_{it(MDLS)}$ and $n_{it(LS_k)}$	139
6.4.2	Sensitivity of w_k , A_{LS_k} and $DP_{wgndestin}$	143
6.4.3	Practical implications	147
6.4.4	Conclusions	151
6.5	Performance of the multi-directional local search heuristic	151
6.6	Trade-off analysis	155
6.7	Conclusions and future research	158
IV	Conclusions	159
7	Final conclusions and future outlook	161
7.1	Final conclusions	161
7.2	Future research opportunities	164
A	Detailed results for the exact ϵ-constraint method	167
B	Average proportional deviation from I_{Hr} for solutions obtained using the MDLS-heuristic	173

C Detailed tuning results for the large neighbourhood search heuristic for the integrated intermodal routing problem	175
Bibliography	179
Nederlandstalige samenvatting	193
Valorisation	195
Biography and publications	197

List of Tables

2.1	Tactical service network design problems.	21
2.2	Operational intermodal routing problems.	22
2.3	Intermodal drayage problems.	24
3.1	Input parameters for intermodal routing.	60
3.2	Tuning instances for intermodal routing.	61
3.3	LNS algorithm parameters for tuning.	61
3.4	Results of a two-way paired-samples t-test for differences in the objective function (TC_Obj) for excluding operators.	65
3.5	LNS parameter setting for testing.	68
3.6	Test instances for intermodal routing.	68
3.7	Detailed results for the removal of a long-haul connection relative to the base case, cases a and b	75
4.1	Objectives for the train load planning problem.	91
4.2	Constraints for the train load planning problem.	96
4.3	Static and dynamic train load planning problems.	97
5.1	Input data: problem types for train load planning.	115
5.2	Computational results for the ϵ -constraint method.	118
5.3	Analysis of extreme points for the ϵ -constraint method.	119
6.1	MDLS algorithm parameters for tuning.	134
6.2	<i>Irace</i> parameters.	137
6.3	Best configurations obtained by <i>irace</i>	137
6.4	Recommended parameter values based on <i>irace</i> and sensitivity analysis.	147
6.5	Best parameter setting for the multi-directional local search heuristic for the train load planning problem.	152
6.6	Comparison of average results for the exact ϵ -constraint method and the MDLS-heuristic.	154

A.1	Detailed results for the ϵ -constraint approach.	167
B.1	Average proportional deviation from I_{Hr} (in %).	174

List of Figures

1.1	Thesis summary.	9
2.1	The sequential and the integrated approach - example.	17
2.2	Thesis summary - Chapter 2.	19
2.3	Intermodal drayage: street turns.	25
2.4	Two-region intermodal network.	28
3.1	Thesis summary - Chapter 3.	46
3.2	Shift the start of a route on a single day with $F*_{service}$, including travel to and service of node j on the current day.	55
3.3	Shift the start of a route on a single day with $F*_{noservice}$, including travel to node j on the current day, and service of j on a following day.	56
3.4	Terminals in two service regions.	58
3.5	Visualisation of customer pickup and delivery locations in two regions: random or clustered around terminals.	59
3.6	LNS parameter tuning: average solution quality for the sequential and integrated approach.	62
3.7	LNS parameter tuning: average computation time for the sequential and integrated approach.	63
3.8	LNS parameter tuning: contribution of the operators.	66
3.9	LNS parameter tuning: evolution of the solution quality with the number of iterations.	67
3.10	LNS results for the integrated and sequential approach.	69
3.11	Average difference in the number of load units planned with departure and arrival in the current planning week.	71
3.12	Average absolute difference in the cost per drayage task.	72
3.13	Average relative difference in the cost per drayage task.	72
3.14	Difference in terminal throughput share due to congestion for random customer locations.	77

3.15	Difference in terminal throughput share due to congestion for clustered customer locations.	79
4.1	Thesis summary - Chapter 4.	89
4.2	Possible configurations (adapted from Bruns and Knust (2012)). . . .	95
5.1	Thesis summary - Chapter 5.	104
5.2	Example of a bi-objective maximisation problem.	106
6.1	Thesis summary - Chapter 6.	124
6.2	Quality indicators (adapted from Parragh et al. (2009)).	132
6.3	Solution quality based on $n_{it(MDLS)}$ and $n_{it(LS_k)}$	140
6.4	Average solution quality per instance size.	141
6.5	Average computation times based on $n_{it(MDLS)}$ and $n_{it(LS_k)}$	142
6.6	Sensitivity of w_k on the solution quality.	144
6.7	Sensitivity of $A_{LSLength}$ on the solution quality.	144
6.8	Sensitivity of $A_{LSUrgency}$ on the solution quality.	145
6.9	Sensitivity of A_{LSDP} on the solution quality.	146
6.10	Sensitivity of $DP_{wgndestin}$ on the solution quality.	146
6.11	Average solution quality per category of instances for different combinations of $A_{LSLength}$ and $n_{it(MDLS)}$	149
6.12	Average computation times per category of instances for different combinations of $A_{LSLength}$ and $n_{it(MDLS)}$	150
6.13	Average solution quality and time per category of instances for the best parameter setting.	153
6.14	Trade-offs for a realistic instance with 35% critical load units and a uniform weight distribution.	156
6.15	Trade-offs for a realistic instance with 20% critical load units and a uniform weight distribution.	157
7.1	Thesis summary - Chapter 7.	162
C.1	LNS parameter tuning: experimental results for the sequential approach, for each combination of parameters and all instance classes. .	176
C.2	LNS parameter tuning: experimental results for the integrated approach, for each combination of parameters and all instance classes. .	177

Part I

Introduction

Introduction and problem statement

1.1 Intermodal freight transport

Forecasts indicate that freight transport by road is expected to increase by 40% by 2030 and even by over 80% by 2050 compared to 2005. Despite the fact that road transport is less sustainable and environmentally friendly, it has the largest share in the modal split with 76.4%, followed by 17.4% and 6.2% for rail and waterways respectively (Eurostat, 2016). This expected continued growth in freight transport and its associated environmental impact present one of the major challenges in the transport sector (European Commission, 2011a; ITF, 2015). In this context, the European Commission (2011b) aims at shifting 30% of road freight transport over 300 kilometres to other transport modes by 2030, and more than 50% by 2050, to increase the sustainability of our transport system. In order to stimulate this modal shift, the European Commission (2011b) launched a roadmap to remove regulatory barriers towards an integrated European transport area. Despite this clear European strategy, the modal split has not significantly changed since 2008 (European Commission, 2016). Within this aim, the European platform ALICE has been set up to develop a comprehensive strategy for research, innovation and market deployment of logistics and supply chain management in Europe, and advise the European Commission (Alliance for Logistics Innovation through Collaboration in Europe, 2014).

A promising opportunity to achieve this modal shift may be intermodal transport, which is strongly supported by the European Commission (2009, 2011b). It is defined as the movement of goods in the same load unit through a combination of transport

modes in a single transport chain without handling the goods themselves, with most of the route travelled by rail, inland waterway or ocean-going vessel and with the shortest possible initial and final journeys by road (Bontekoning et al., 2004). It is aimed at integrating the various transport modes to provide a door-to-door delivery (Reis et al., 2013). Intermodal transport increases sustainability and relieves congested highways as the long distance is covered by rail or barge, which is especially an important aspect in the centre of Europe (Boysen et al., 2013).

Intermodal rail transport involves the combination of rail and road transport. Bontekoning et al. (2004) indicate that this type of intermodal transport is attractive because shippers in most countries have access to rail, whereas fewer have access to sea and inland shipping. They also describe the path usually followed by load units in the intermodal rail transport chain. Load units are transported by truck for the pre-haulage from their origin to the terminal of departure and for the end-haulage between the destination terminal and their final destination, which is also known as drayage or pickup and delivery. For the long haul, load units are bundled and shipped via rail. In between, terminals serve as transshipment points between road and rail transport.

Intermodal rail transport has several advantages over unimodal road transport. Although road transport is flexible, fast and requires low capital investments, fuel costs and congestion on the roads are rising and thus the attractiveness of this transport mode decreases. Reis et al. (2013) state that transport is responsible for approximately 25% of European global CO₂-emission. This may be reduced by a modal shift from road to rail transport, as rail transport is more green and sustainable compared to road transport. Rail-road transport combines the accessibility of road transport with economies of scale of rail transport. Furthermore, bundling load units on trains over long distances also reduces CO₂-emissions and thus releases pressure on the environment (European Commission, 2009).

However, the combination of rail and road transport is subjected to a different cost pattern compared to unimodal road transport. For intermodal rail transport to become advantageous compared to unimodal road transport, a number of additional costs need to be reduced. Drayage accounts for 25% to 40% of the total transport costs, while the distance travelled by truck is relatively small (Bontekoning et al., 2004). A reduction in pre- and end-haulage costs can be obtained by reducing the distance to the closest rail terminal and organising truck routes more efficiently. A second option is to reduce transshipment costs by striving for fast, easy and reliable transshipment. Another possibility is to make the train transport cheaper by increasing the efficiency of operations in organisations. Together with these costs, the

time required for intermodal rail transport must be reduced, as rail transport is currently slower (Reis et al., 2013), which can for example be obtained by more frequent services.

Furthermore, in order to achieve a sustainable transport system, the full cost of transport should be accounted for. This means that external costs should be internalised. Internal costs consist of direct costs such as fuel costs, wages and equipment costs. External costs include costs of congestion, noise, emissions, accidents and air pollution. Despite the fact that vehicles on the road are becoming greener (e.g., low emission vehicles), external costs of congestion and, related to that, costs of pollution increase. Internalising these external costs may encourage a shift towards intermodal transport (Macharis et al., 2010; Pinto et al., 2018).

Ambra et al. (2019) emphasise the need for innovative integrated support systems with a focus on cost efficiency in order for intermodal transport to increase its competitiveness. One innovative, recent solution concerns synchromodality. It is an extension of intermodal transport, which is often perceived as too static. Synchromodal transport emphasises the dynamics and flexibility which should be incorporated in decisions on routing and rerouting load units using multiple transport modes. In a synchromodal view, ideally, load units are routed through an interconnected network of hubs depending on the network capacity, where the sender is not concerned about the route of its packages (Ambra et al., 2019). Synchromodal transport requires an integrated view on the transport chain as a whole. Multiple complimentary transport modes and connections should be available, such that the most appropriate route can be selected based on the customer requirements and available transport options at a specific moment in time (Behdani et al., 2016). This requires flexible and real-time planning, which poses a number of challenges for planners. In this context, adequate decision support and fast planning algorithms with real-life characteristics are needed to support this synchromodal vision.

1.2 Decision support

Caris et al. (2008) divide the current literature on intermodal freight transport into three planning levels based on the time horizon of decision impact. Strategic, tactical and operational decision problems can be distinguished. Each planning level can be divided into four categories based on the type of decision maker involved. Drayage operators are concerned with the planning of trucks performing the transport between the terminal and the shipper or receiver. Terminal operators are responsible for the

transshipment of load units. Network operators are occupied with the infrastructure planning and organise the rail or barge transport. Finally, intermodal operators use the intermodal infrastructure and services. They select the most appropriate routes throughout the available intermodal network.

Caris et al. (2013) stress the fact that intermodal transport has a higher degree of complexity than unimodal transport, because more transport modes and consequently more decision makers are involved. The degree of complexity increases even further in the most recent synchromodal concept. Synchromodality takes intermodal transport one step further; as with intermodal transport, two or more transport modes are combined, however, synchromodality allows for flexible mode selection based on the operational circumstances and customer requirements (SteadieSeifi et al., 2014). This implies that integration, synchronisation, and alignment of processes and operations are crucial factors to guarantee a stable and efficient service to the customer. In other words, instead of optimising their processes individually, operators should integrate their activities and coordinate the entire process. For instance, ICT innovations can support the exchange of real-time information and integration to support more efficient decision making.

One key aspect of synchromodality concerns real-time switching, due to which assignment decisions are proactive, taking decisions based on available data feeds, and more responsive to internal and external perturbations compared to a more classical approach. This leads to more complex operational decision processes requiring more flexibility. Therefore, existing planning and transport management systems should be adapted accordingly. Furthermore, unexpected events may have a high impact on the chain performance due to interdependencies in the transport connections and therefore, should be managed carefully (Behdani et al., 2016). Changes to the initial planning can only be performed with knowledge on the current availability of capacity for the various transport modes at hand. Consequently, the planning performance of a capacity management system is highly dependent on the availability of adequate and efficient planning algorithms. With amodal booking, the logistics service provider, instead of the shipper, is allowed to choose the modality. A resulting flexible mode and service selection further increases the planner's responsibilities.

1.3 Research objectives

Innovative transport systems should be developed in order to stimulate a modal shift towards intermodal transport with synchronised services. Such transport systems

require an integrated view on the transport network, and can be facilitated by informed and flexible planning (Alliance for Logistics Innovation through Collaboration in Europe, 2014). The number of research papers focusing on synchromodal transport and its prerequisites is rising. However, a gap between academic research and practice can be observed, because real-life characteristics of intermodal transport companies are ignored. Therefore, the focus of this thesis is on decision support for human planners at the operational decision level during the planning process of intermodal rail transport.

In order to stimulate intermodal transport, on the one hand, additional costs should be reduced, while on the other hand fast planning algorithms with real-life problem characteristics should be available to accommodate decisions in a complex intermodal planning environment. The central research objective of this thesis is to *offer intermodal planning support in order to minimise total transport costs and maximise service capacity utilisation*, which in turn results in decreasing costs of the transport system.

This central research objective is dealt with by addressing two intermodal planning problems from the viewpoint of an intermodal operator: intermodal routing and train load planning (TLP). The *first contribution* of this thesis is the introduction of an integrated intermodal routing problem. After a transport order is received, planners must assign each load unit to an intermodal long-haul service to maximise the overall network capacity utilisation, and local drayage routes must be established to transport load units between load and unload locations and the transshipment terminals for long-haul transport. Usually, these decisions are made in a sequential way. As the level of integration of different transport modes and decision levels influences the attractiveness of intermodal transport, an integrated intermodal routing problem is proposed. By taking local truck routing and long-haul transport planning decisions simultaneously, the aim is to reduce total transport costs. In a *second contribution*, the intermodal routing model is applied to a real-life intermodal network to support the analysis of the impact of tactical service network design decisions made by an intermodal operator. Insights will be presented on how to best utilise a given service network, in order to reduce total transport costs and synchronise available services to the expected demand of orders. The approach aims at reducing the number of road kilometres and increases bundling opportunities by maximising the long-haul capacity utilisation. In this way, it contributes to the modal shift towards intermodal transport and a more sustainable transport system.

At a more operational level, train load planning is concerned with the assignment of load units to specific locations on intermodal trains, accounting for real-life loading

restrictions in order to maximise the on-train capacity utilisation. As including these real-life aspects are key to deciding which load units will be assigned to which location, the *third contribution* is the introduction of an exact and heuristic algorithm with real-life train load problem characteristics. It can be used as a decision support tool to provide human planners with a number of plans from which they can choose the best one for a specific moment in time. Early-stage planning assistance during the booking process may improve the overall capacity utilisation of the available rail services. Terminal operations required to execute the train load plan are usually optimised once the load units to be loaded are fixed.

The viewpoint considered in this dissertation is the one of an intermodal operator which owns own rail services and purchases slots on rail services of other parties. Train load planning only has to be performed for rail services owned by the company. For intermodal routing, both types of services are considered. This provides flexibility for using the proposed model both in the case of an intermodal operator which leases slots on trains of other parties, as well as service providers which offer rail services. Both problems studied in this dissertation occur at the operational planning level. While the intermodal routing problem usually occurs in weekly planning cycles, train load planning takes place at the very detailed, almost real-time, planning level. Besides, while intermodal routing is an operational problem, the solution method can be used to analyse the impact of small changes related to tactical service network decisions on the operational costs.

Decisions on these two planning problems, intermodal routing and train load planning, influence the throughput of the intermodal transport system and aim at minimising overall transport costs while accounting for customer service requirements. Fast decision support by means of planning algorithms are provided for improved, more efficient planning. Results contribute to a better understanding of the way in which intermodal operators can maximise their service network's transport capacity based on the expected demand of transport orders, and lead to an increased transport capacity utilisation and a minimisation of total transport costs.

1.4 Thesis outline

Figure 1.1 presents the general thesis overview. After an introductory chapter (Part I), Chapters 2 and 3 (Part II) focus on integrated, operational decision support for routing transport orders using trucks and multiple available long-haul services. Chapters 4 to 6 (Part III) relate to operational decision support with a shorter planning hori-

zon, during the train load planning phase, in which load units are assigned to specific locations on an intermodal train. Conclusions of both subjects studied are presented in a final Part IV.

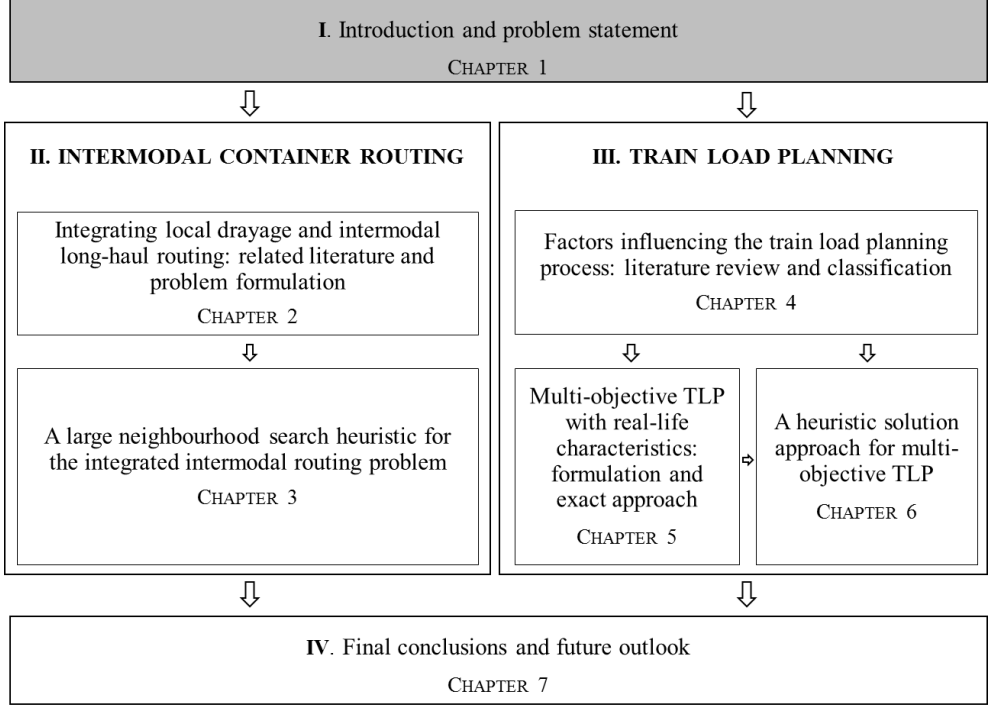


Figure 1.1: Thesis summary.

In Chapter 2, a new integrated intermodal routing problem is proposed, combining drayage and long-haul rail network flow planning decisions. In practice, requests are assigned first to long-haul routes, after which truck routes are composed. However, an integrated approach for both decisions is appropriate. While customer pickup and delivery locations of full-container requests are known, the departure and arrival terminals through which each request travels depend on the selected long-haul service. Consequently, the delivery location for pickup tasks and pickup location for delivery tasks are unknown in advance. Therefore, changes in long-haul routing decisions imply different combinations of pickup and delivery locations and thus impact total trucking costs. In this dissertation, an integrated intermodal routing problem, including a multi-day scheduling of trucks, is proposed for the first time. Mathematical formulations of a sequential and integrated approach with real-life problem character-

istics are presented. Chapter 3 proposes a large neighbourhood search (LNS) heuristic algorithm to solve the problem. The building blocks consist of operators dedicated to the integrated nature of the problem. The heuristic algorithm is used to compare both approaches and to analyse the impact of decisions related to the service network design on total transport costs. This differs from current research because the impact of costs at the operational level is usually ignored when decisions at the tactical level on the design of a service network are made.

Chapter 4 introduces the train load planning problem and its relevance for intermodal transport. It presents a classification and a state of the art of factors influencing train load planning, and results in an identification of opportunities for future research. Based on the conclusions of this chapter, real-life characteristics of the train load planning problem are introduced into a problem formulation in Chapter 5. The train load planning problem is considered from a multi-objective perspective, enabling planners to account for various objectives during the train planning phase within the context of intermodal rail transport. For the first time, a multi-objective approach for train load planning with real-life characteristics is proposed. The adaptive ϵ -constraint method is used to solve the problem exactly and computational results are presented. In Chapter 6, a multi-objective heuristic for the train load planning problem is proposed. A multi-directional local search (MDLS) heuristic is developed to solve problems of realistic size. Its performance is demonstrated by comparing results with optimal results obtained using the exact ϵ -constraint method. The algorithm provides solutions of good quality for real-world problems in short computation times.

Finally, Chapter 7 presents the main conclusions of this thesis and opportunities for future research.

Part II

Intermodal container routing

— *Introduction to Part II: Intermodal container routing* —

Part II consists of Chapters 2 and 3, in which a new, integrated intermodal routing problem is presented and solved using metaheuristics. The problem aims at providing integrated decision support for routing containers through a service network in the context of intermodal transport. It is situated at a operational level, using a weekly planning cycle. The focus is on the integration of decisions on local drayage within service regions of terminals and long-haul routing decisions in which transport orders are assigned to long-haul services.

Integrating local drayage and intermodal long-haul routing: related literature and problem formulation

2.1 Introduction

Intermodal transport involves multiple actors and decision makers, resulting in a more complex planning environment compared to unimodal road transport. Synchronomodal transport further increases this complexity as it requires informed and flexible planning (Alliance for Logistics Innovation through Collaboration in Europe, 2014). As the level of integration of different transport modes and decision levels influences the attractiveness of intermodal transport, adequate decision support systems for human planners at the operational decision level should be developed. Fast planning algorithms with real-life problem characteristics should be available to accommodate decisions in a complex intermodal setting.

Intermodal logistics service providers must decide on the planning of drayage and main-haul transport. *Tactical* decisions are related to the selection of intermodal transport services and the allocation of existing resources (Caris et al., 2008). The

This chapter is based on the following paper: Heggen, H., Molenbruch, Y., Caris, A., Braekers, K., 2019. Intermodal Container Routing: Integrating Long-Haul Routing and Local Drayage Decisions. *Sustainability*, 11, 1634.

service network design determines the routes on which services are offered and the characteristics of each service (Caris et al., 2013). Characteristics of regular long-haul services include the route, intermediary stops, frequency, vehicle and convoy type, capacity and speed. Most of these characteristics are unchanged for a given scheduling period (Crainic and Kim, 2007). Often, weekly planning cycles are fixed for at least six months. At the *operational* decision level, policies have to be determined for routing containers throughout the given intermodal network and services (Caris et al., 2013). Intermodal routing or route selection is concerned with the selection of routes or itineraries for individual containers over a given intermodal network (Caris et al., 2013; Li et al., 2015), related to a planning horizon of one or several days. The problem also includes real-time itinerary replanning and asset management (SteadieSeifi et al., 2014). Next, during train load planning, containers are assigned to specific locations on intermodal trains while including detailed capacity considerations related to the wagon and slot length and weight restrictions, as discussed in Part III (Chapters 4 to 6). The fact that these problems and decision levels are interrelated is recently stressed by several authors (e.g., van Riessen et al. (2016); Zhang and Pel (2016)).

In current literature, decisions with respect to the long-haul transport and vehicle routing of trucks in a service region are usually considered independently. The distance or travel time between each customer location and the available terminals is taken into account when solving the intermodal routing problem. The truck cost is approximated by the direct distance between a customer and a terminal location. Pre- and end-haulage is arranged in a second phase by combining pickup and delivery operations between customer locations and terminals into routes. In this sequential approach, the combination of pickup and delivery tasks in truck routes (i.e., the real total drayage cost) is ignored initially. However, the integration of vehicle routing problems into the intermodal terminal selection and container routing throughout the network may provide important cost savings and a better utilisation of the available transport capacity.

In a sequential approach, the long-haul service for each request is selected in a previous phase, and results in a single drayage task in each region. In an integrated approach, the long-haul service is not fixed in advance, which results in multiple feasible drayage tasks per region, of which one per region has to be selected simultaneously with the decision on the long-haul service. Figure 2.1 shows a possible truck cost reduction for the request indicated in black symbols by using an integrated approach. The vehicle routing problem in each region consist of a full-truckload problem in which trucks wait for containers to be fully loaded or unloaded before travelling to any other location. The selected long-haul service is indicated by the bold line

between both regions. Whereas a sequential approach would select the long-haul service connecting both upper terminals, an integrated approach might select the lower service which does not have the lowest direct trucking distance, but results in the lowest total cost.

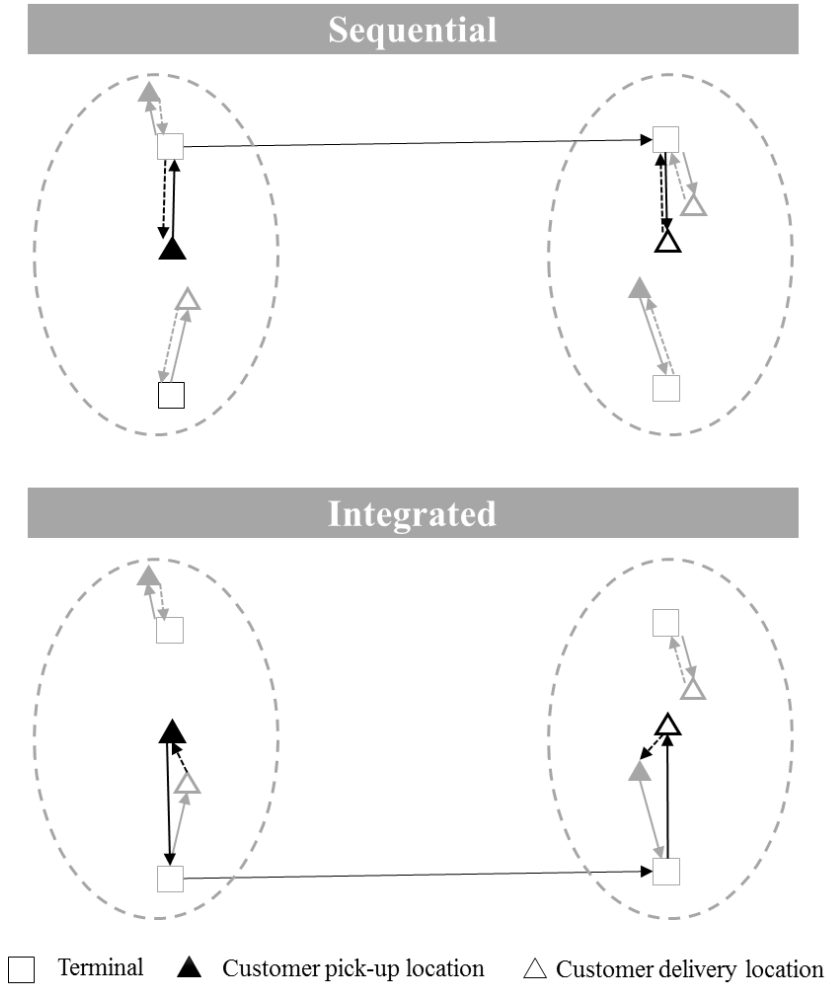


Figure 2.1: The sequential and the integrated approach - example.

In this chapter and the next, a model formulation and algorithm will be developed to compare the sequential and integrated approach, and quantify these potential savings. The aim is to provide an operational decision support tool for transport planners of an intermodal operator. Moreover, the planning tool is able to include the impact

of operational costs in tactical decisions when adapting the offered transport service network. A heterogeneous fleet of trucks and load units, multi-day truck scheduling and long-haul service capacity limits are considered. The focus is on rail transport for the long-haul, as most countries in the EU have access to an extensive rail network, whereas waterways are not always available. The proposed integrated approach contributes to a more sustainable transport system by reducing the distances travelled using road transport. Moreover, by maximising the long-haul capacity utilisation, bundling opportunities can be fully deployed.

In this chapter (Figure 2.2), literature on both intermodal routing for the long-haul and intermodal vehicle routing problems is discussed, and the focus of the remainder of this chapter is presented (Section 2.2). The considered integrated planning problem is based on a real-life case study from the viewpoint of the transport planning department of an intermodal service operator with a network of long-haul services consisting of either own trains, purchased services which are managed by external parties, or a combination of both. A more detailed problem description is provided in Section 2.3. Formulations for both subproblems in the sequential approach, as well as an integrated intermodal routing problem with real-life problem characteristics in which truck routing and long-haul routing decisions are made simultaneously, are proposed in Section 2.4.

2.2 Related literature

The integrated intermodal routing problem relates to two problem settings in current literature which are usually solved independently: on the one hand routing containers (or load units) throughout the long-haul network, including service network design problems at the tactical level and intermodal routing at the operational level, and on the other hand vehicle routing problems in the context of drayage operations around terminals. Recent advances in literature on both types of problems are discussed in Section 2.2.1 and 2.2.2 respectively. Section 2.2.3 presents the closest related research for integrating both problems and the focus of the remainder of this chapter.

2.2.1 Long-haul service network design and intermodal routing

Tables 2.1 and 2.2 provide an overview of (intermodal) service network design and intermodal routing problems considering rail, road and/or water transport modes. A distinction is made between tactical problems aimed at the selection of a service

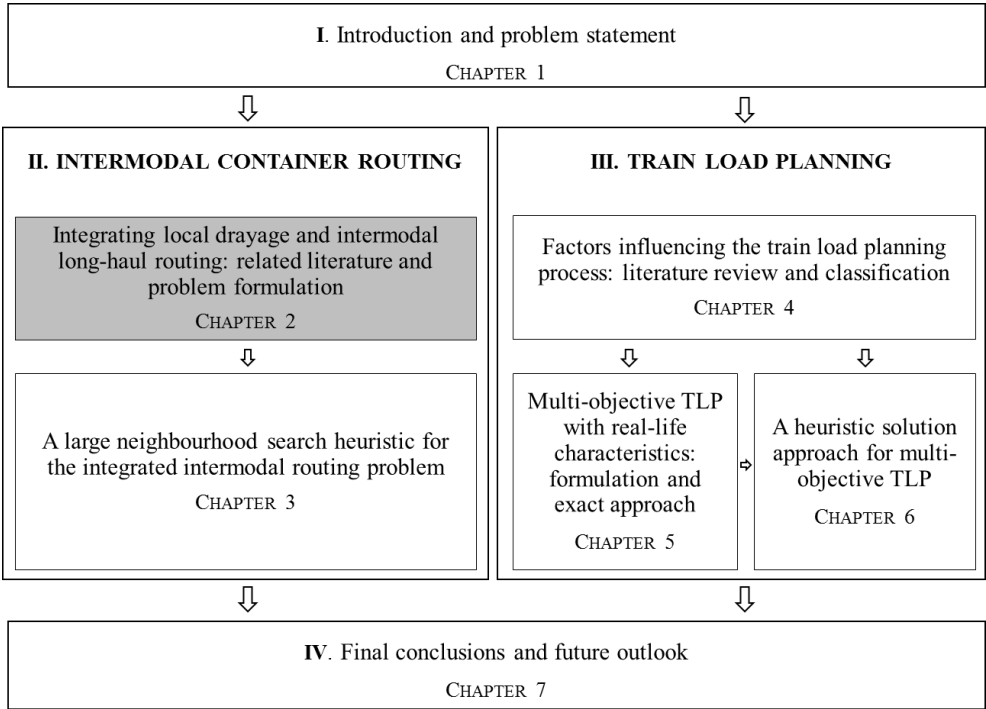


Figure 2.2: Thesis summary - Chapter 2.

network (Table 2.1), and operational problems which determine the routing of individual shipments (Table 2.2). The left-hand side of the tables displays different problem characteristics, whereas the right-hand side indicates information with respect to the objectives, where problems not solved using a weighted objective function are indicated as ‘multi-objective’.

Decision levels Current research indicates that operational and tactical problems are interrelated. Operational intermodal routing problems may include real-time planning and replanning (Grasman, 2006; Bock, 2010; Li et al., 2015; van Riessen et al., 2016; Di Febbraro et al., 2016) and asset management (and repositioning) decisions (Erera et al., 2005; Grasman, 2006; Assadipour et al., 2015; Li et al., 2015; Jiang et al., 2017). They sometimes include fixed costs for selecting services, which relates to the tactical service network design. This can be explained by the fact that routes selected at the tactical decision level may no longer be feasible if disruptions occur at the operational level (Grasman, 2006). Tactical problems usually assume that delivery time windows can be violated, whereas at the operational phase, time windows are usually constraints. However, some operational problems allow late delivery at a penalty cost (Bock, 2010; Verma et al., 2012; Assadipour et al., 2015; van Riessen et al., 2016). This shows that different decision levels are linked. Although some papers include more than one decision level, most research focuses on a single problem and decision level. This research combines two operational problems. The integrated problem is used to study the influence of changes in tactical decisions on operational transport costs.

Service characteristics and objectives A mix of different types of transport services may be used, such as flexible or fixed, and own or externally managed services (e.g., Moccia et al. (2011)). Flexible services can be activated ad hoc in accordance with the requirements, such as the decision to operate trucks or purchase slots on trains during peak periods, while fixed services usually provide a fixed capacity between a given departure and arrival location with a given schedule. Operators need to determine the capacity of own means and negotiate slot contracts with third parties based on estimated future demand. While services operated by other parties are generally purchased at a fixed cost per container, own services may be subject to economies of scale (e.g., Chang (2008); Caramia and Guerriero (2009); Moccia et al. (2011)).

For each service offered, current literature generally includes high-level capacity restrictions such as a maximum ‘flow’ or ‘quantity’ (these cases are indicated with (1)

Table 2.1: Tactical service network design problems.

	Decision level	Problem characteristics										Objectives				
		Capacity limit	Fixed costs for services	Itinerary replanning	Asset management	Strict time windows	Departure times/frequency	Time-dependent travel times	Economies of scale	Multi-objective	Min flow cost	Min travel time/distance	Min transit time variability	Min hazmat risk exposure	Green objectives	
Andersen and Christiansen (2009)	T	1	x	—	x	—	x	—	—	—	x	—	x	—	—	
Pedersen et al. (2009)	T	1	x	—	x	—	—	—	—	—	x	—	—	—	—	
Andersen et al. (2009a)	T	1	x	—	x	—	x	—	—	—	x	—	—	—	—	
Andersen et al. (2009b)	T	1	x	—	x	—	x	—	—	—	x	—	—	—	—	
Bai et al. (2010)	T	1	x	—	x	—	—	—	—	—	x	—	—	—	—	
Vu et al. (2013)	T	1	x	—	x	—	—	—	—	—	x	—	—	—	—	
Zhu et al. (2014)	T	#wgns	x	—	x	—	x	—	—	—	x	x	—	—	—	
Chouman and Crainic (2015)	T	1	x	—	x	—	—	—	—	—	x	—	—	—	—	
van Riessen et al. (2015)	T	2,3	x	—	—	—	x	—	—	—	x	—	—	—	—	
Baykasoğlu and Subulan (2016)	T	1	x	—	—	—	—	—	—	x	x	x	—	—	x	
Qu et al. (2016)	T	1	x	—	—	—	—	—	—	—	x	x	—	—	x	
Demir et al. (2016)	T	3	x	x	—	—	x	x	—	—	x	x	—	—	x	
Rudi et al. (2016)	T	1,2	—	—	—	—	—	—	x	—	x	x	—	—	x	
Li et al. (2017)	T	1	x	—	x	—	x	—	—	—	x	—	—	—	—	
SteadieSeifi et al. (2017)	T/O	1	x	—	x	x	x	—	—	—	x	x	—	—	—	

O = Operational

T = Tactical

1 = limit on the number of containers

2 = weight limit

3 = length limit

Table 2.2: Operational intermodal routing problems.

	Decision level	Problem characteristics								Objectives				
		Capacity limit	Fixed costs for services	Itinerary replanning	Asset management	Strict time windows	Departure times/frequency	Time-dependent travel times	Economies of scale	Multi-objective	Min flow cost	Min travel time/distance	Min transit time variability	Min hazmat risk exposure
Barnhart and Ratliff (1993)	O	—	—	—	—	x*	—	—	—	—	x	—	—	—
Boardman et al. (1997)	O	—	—	—	—	—	—	—	—	—	x	x	—	—
Ziliaskopoulos and Wardell (2000)	O	—	—	—	—	—	x	x	—	—	—	x	—	—
Erera et al. (2005)	O	1	—	—	x	x*	x	—	—	—	x	—	—	—
Grasman (2006)	O/T	—	x	x	x	x*	—	—	—	x	x	x	—	—
Chang (2008)	O	1	x	—	—	x	x	—	x	x	x	x	—	—
Caramia and Guerriero (2009)	O/T	2	x	—	—	x	x	—	x	x	x	x	—	—
Bock (2010)	O	2, 3	—	x	—	—	—	—	—	—	x	—	—	—
Moccia et al. (2011)	O	1, 2, 3	x	—	—	x	x	—	x	—	x	—	—	—
Yang et al. (2011)	O	—	x	—	—	—	—	—	—	x	x	x	x	—
Gromicho et al. (2011)	O	—	—	—	—	x	x	—	—	—	x	—	—	—
Cho et al. (2012)	O	—	—	—	—	—	x	—	—	x	x	x	—	—
Verma et al. (2012)	O	1	x	—	—	—	—	—	—	x	x	—	—	x
Assadipour et al. (2015)	O	1	x	—	x	—	—	—	—	x	x	—	—	x
Li et al. (2015)	O	3	—	x	x	—	x	x	—	—	x	x	—	—
Resat and Turkay (2015)	O	1	—	—	—	—	x	x	—	x	x	x	—	—
Zhang and Pel (2016)	O	3	—	—	—	—	x	—	—	—	x	x	—	—
Behdani et al. (2016)	O	1	—	—	—	x	x	—	—	—	x	x	—	—
van Riessen et al. (2016)	O	2, 3	—	x	—	—	x	—	—	—	x	—	—	—
Di Febbraro et al. (2016)	O	1; #wgns	—	x	—	x	x	—	—	—	x	x	—	—
Jiang et al. (2017)	O	1	x	—	x	—	—	—	—	—	x	—	—	—

O = Operational

T = Tactical

1 = limit on the number of containers

2 = weight limit

3 = length limit

x*: only delivery time

in Table 2.1, i.e., a limited number of containers that can be loaded). Some papers include a total allowed maximum length in twenty foot equivalent units (TEU) or weight per service. From contacts with the transport sector, we learned that rail transport contracts with external parties include agreements which stipulate how much transport capacity is purchased. Slot agreements include detailed information on the number of slots, their (length) dimensions and the allowed weight for each selected slot. For own services, wagons are leased for a longer period based on expected transport requests. These aspects related to the determination of the mix of slot types should be determined at the medium term (i.e., tactical decision level). At the operational level, multiple types of capacity limits are included by Bock (2010); Moccia et al. (2011); van Riessen et al. (2016) and Di Febbraro et al. (2016). At the tactical level, van Riessen et al. (2015) and Rudi et al. (2016) consider a maximum weight and length, and number of containers and weight respectively.

Pre- and end-haulage costs are part of the total route costs, and are usually considered as direct links between shipper or receiver nodes and terminal nodes in the transport network. Combinations of pickup and delivery tasks during drayage operations are not accounted for, but occur in practice.

2.2.2 Pre- and end-haulage transport

Intermodal local drayage operations are concerned with the pickup and delivery of full-truckload inbound and outbound containers in the service region of an intermodal terminal within time windows at terminals and customer locations (Caris and Janssens, 2009). The aim is to find efficient truck routes between intermodal terminals and customer locations, where truck typically have a capacity of a single full container. For example, Figure 2.3 shows that the total transport distance and cost can be reduced allowing direct transport between customer locations (i.e., street turns), in comparison to transport between terminals and customers (where street turns are not considered), by reducing empty transport distances.

An overview of intermodal drayage problems and their characteristics in the literature is provided in Table 2.3. For a state of the art on general vehicle routing problems the reader is referred to Toth and Vigo (2014) and Braekers et al. (2016b).

The vehicle fleet is often considered homogeneous; only two papers discuss a heterogeneous fleet (Dotoli and Epicoco, 2016; Pérez Rivera and Mes, 2017). Multiple vehicle depots can be included (Zhang et al., 2009, 2010; Nossack and Pesch, 2013; Sterzik and Kopfer, 2013; Reinhardt et al., 2016; Shiri and Huynh, 2016). Moreover, to cope with imbalances between demand and supply of containers at different loca-

Table 2.3: Intermodal drayage problems.

Authors	Terminals	Vehicle depots	Heterogeneous truck fleet	Empty container repositioning	TW customer	TW terminals
Wang and Regan (2002)	1	1	—	—	x*	—
Francis et al. (2007)	M	1	—	x	x	x
Imai et al. (2007)	1	1	—	—	—	—
Caris and Janssens (2009)	1	1	—	—	x	—
Zhang et al. (2009)	1	M	—	x	x	x
Caris and Janssens (2010)	1	1	—	—	x	—
Zhang et al. (2010)	M	M	—	x	x	x
Braekers et al. (2013)	M	1	—	x	x	x
Escudero et al. (2013)	1	1	—	—	—	x
Nossack and Pesch (2013)	M	M	—	x	x	x
Sterzik and Kopfer (2013)	M	M	—	x	x	x
Dotoli et al. (2015)	1	1	—	—	—	—
Dotoli and Epicoco (2016)	1	1	x	—	x	—
Reinhardt et al. (2016)	M	M	—	x	x	—
Shiri and Huynh (2016)	1	M	—	x	x	x
Pérez Rivera and Mes (2017)	M	1	x	x	x	x

* = only pickup time window

M = multiple

HE = heterogeneous

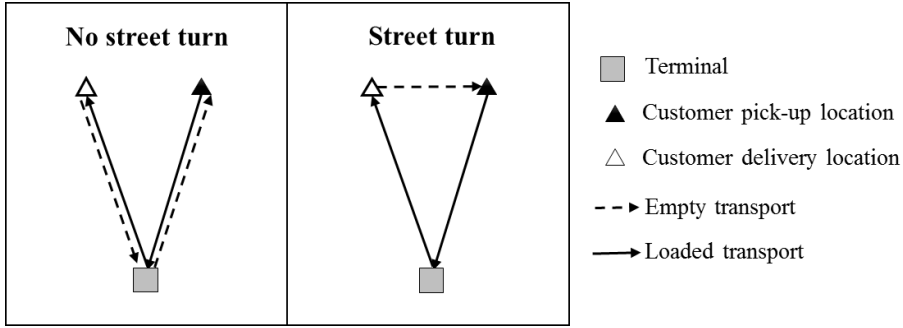


Figure 2.3: Intermodal drayage: street turns.

tions, the allocation of empty containers can be modelled (Francis et al., 2007; Zhang et al., 2009, 2010; Braekers et al., 2013; Nossack and Pesch, 2013; Sterzik and Kopfer, 2013; Reinhardt et al., 2016; Shiri and Huynh, 2016; Pérez Rivera and Mes, 2017). Time windows at intermodal terminals reflect either opening hours of terminals or ultimate arrival times for a long-haul service selected in advance. In the latter case, the long-haul service network and routing characteristics are implicitly included and assumed to be given.

Although a number of papers include multiple terminals (Francis et al., 2007; Zhang et al., 2010; Sterzik and Kopfer, 2013; Nossack and Pesch, 2013; Braekers et al., 2013; Shiri and Huynh, 2016; Pérez Rivera and Mes, 2017), it is generally assumed that the pickup location (i.e., arrival terminal) of inbound full containers and delivery location (i.e., departure terminal) of outbound full containers are known and fixed. This implies that long-haul routing decisions are assumed to be given.

2.2.3 Research opportunities

Integrating the decision on which pickup tasks and delivery tasks to combine in truck routes when making intermodal routing decisions may lead to reduced trucking costs and possibly different long-haul itineraries for individual containers. The integration of both interdependent problems, the determination of long-haul routes and drayage routes, into an integrated intermodal container routing problem may result in important cost savings. However, research efforts which simultaneously consider both decisions are still limited.

In the literature on dial-a-ride problems, Posada et al. (2017) integrate transfers to and from timetabled public transport into demand-responsive passenger transport.

Similarly, decisions on which transfer terminals to use (and thus which public transport line) are made simultaneously with vehicle routing decisions. Public transport offers a high frequency and many intermediate stops compared to freight transport by rail, and the public transport capacity is generally very high. These characteristics imply an inherent difference in the problem setting. The closest related research on freight transport consists of a number of recent research efforts. Medina et al. (2018) synchronise the service network design and vehicle routing problem by assigning each customer location to a single terminal. Dragomir et al. (2018) formulate a less-than-truckload (LTL) pickup and delivery problem with interregional lanes for long-haul vehicles between multiple regions, using multiple depots transport modes. No solution method is provided. Soriano et al. (2018) solve a two-region LTL multi-depot pickup and delivery problem. However, long-haul schedules and capacity limits for the main-haul, multi-day truck scheduling and heterogeneous vehicles are not considered. Wolfinger et al. (2018) also consider transport over longer distances using multiple transport modes. They include a limited daily active time, but no minimum overnight's rest. Further, as schedules or capacity restrictions for the long-haul are ignored, trade-offs between drayage costs and long-haul routing cannot be examined. Current research on integrating short- and long-haul flows in intermodal freight transport is still very recent and scarce. The fact that long-haul transport capacity is usually limited and operates according to service schedules is generally ignored. No research exists in which realistic truck driving hours with respect to the daily active time as well as a minimum amount of rest are included.

Moccia et al. (2011) focus on the decision of which mix of fixed and flexible (requested ad hoc) services should be activated from other operators, given the fact that certain train services are purchased with train weight and length limits. Their problem is clearly related. However, truck routing in a service region and detailed capacity considerations are not included. With respect to intermodal drayage, only one paper (Pérez Rivera and Mes, 2017) includes a flexible terminal assignment cost in a vehicle routing problem within the service region of an intermodal terminal. However, no long-haul service characteristics are incorporated. To the best of our knowledge, no paper considers both real-life rail service and drayage characteristics and handles intermodal routing and drayage in an integrated way.

With respect to capacity levels, clear differences exist between literature on intermodal routing and train load planning. Current research on train load planning includes detailed capacity requirements (i.e., size and weight) when containers are already assigned to specific rail routes. On the other hand, problems aimed at routing containers throughout a service network generally consider a very high-level view

on capacity (e.g., a number of containers). Furthermore, in practice, at the tactical decision level, the intermodal operator determines slots to be purchased on externally managed trains and wagon lease agreements of own services. This relates to the physical dimensions of transport requests at the operational level (e.g., 30ft, 45ft). When demand for transport requests is high, capacity limits of slots of each dimension on external services or the total available length or weight for own services may influence the selected long-haul route for individual requests (e.g., if not enough capacity is available of a specific dimension at one service, another should be selected). Within this regard, a heterogeneous fleet of trucks is also employed for containers of varying dimensions. These more detailed capacity requirements related to the dimensions of slots, wagons and trucks must be integrated into container routing decisions.

In the remainder of this chapter, the intermodal local drayage problem with a heterogeneous truck fleet is combined with a long-haul routing problem with detailed capacity considerations into an integrated intermodal routing problem. The vehicle routing problem includes multiple depots located at terminal locations and a flexible departure and arrival terminal allocation, which implies flexible delivery and pickup locations in each region.

2.3 Problem description

Direct freight transport between (large-volume) terminals is usually operated based on weekly planning cycles with fixed, regular service schedules and capacity for each connection (Crainic and Kim, 2007), which can be established if high volumes are available to be transported between two regions. These schedules include cut-off times, indicating the latest moment at which freight must be available at the departure terminal to meet the scheduled departure, and release times, at which freight becomes available at the destination terminal.

The planning department of an intermodal service provider must decide on the tactical service network design as well as daily operational routing of individual containers. The service network indicates which and how many terminals are operated in a service region, and connects these terminals in different service regions by long-haul services. The transport of individual requests (cfr. freight containers) from their customer origin to their customer destination relates to two decision. On the one hand, intermodal routing decisions should be made. For each container, the departure and arrival terminals have to be decided upon, taking into account the long-haul connections, schedules and capacities in the network. On the other hand, drayage operations

should be arranged accordingly. Inbound containers are picked up by trucks at intermodal terminals while outbound containers must be delivered at terminals. This involves solving a vehicle routing problem in each service region.

The intermodal routing problem deals with own rail services (i.e., the company owns a train and rents a railway path), as well as the fact that slots may need to be purchased on rail services managed by other parties. At the tactical level, this means that for own trains, the company should negotiate with infrastructure managers on the rented railway paths. For rail services managed by other parties, the company should decide how many slots of each container length are rented for a fixed period. Both decisions are fixed for multiple months. At the moment of renegotiating both types of transport contracts, the proposed solution method provides decision support. For example, the impact of expectations about future demand or tactical decisions on the offered services and used terminals can be evaluated.

Context Given a long-haul service network with fixed schedules, the aim is to optimally utilise the available intermodal network in order to minimise the total transport cost of the complete system, which consists of rail transport costs and truck routing costs. Within this objective, no unimodal direct trucking between customer origins and destinations is allowed, which is acceptable if the distance between two freight regions is relatively high. Moreover, within the scope of this chapter, it is assumed that empty container repositioning between terminals is determined afterwards in accordance with the demand for empty containers. An example of a network is visualised in Figure 2.4.

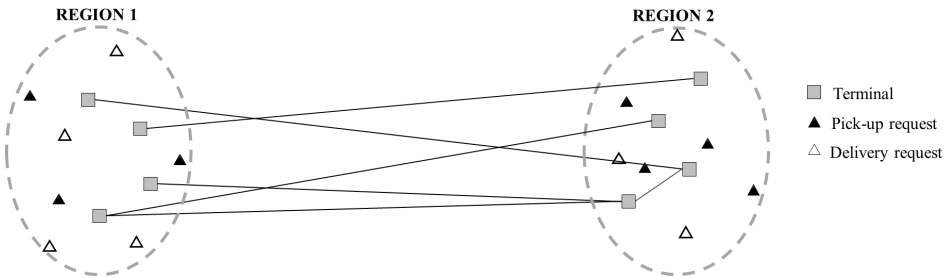


Figure 2.4: Two-region intermodal network.

Contributions To the best of our knowledge, a new, integrated intermodal routing problem is proposed and compared with a sequential approach in which first

intermodal routing decisions and next decisions on drayage operations are made. An integrated planning approach simultaneously decides on the routing of containers throughout the network and local truck routing in each service region. By considering drayage and long-haul costs simultaneously, total transport costs may be reduced. In this integrated approach, the customer pickup location for outbound containers and the customer delivery location of inbound containers within each service region are known and time windows at customer locations are fixed. A container can be transported via any available long-haul connection, implying that the departure and arrival terminal are not fixed.

Service contracts with other transport operators include detailed capacity information, such as a number of slots per length type and weight category. Moreover, not all containers can be carried by any truck, depending on the dimensions of containers and trucks. This may result in separate routes for different types of containers and trucks. Different container types related to the physical dimensions are assigned to the appropriate slot dimensions for long-haul rail services and with the routing of a heterogeneous fleet of trucks for pre- and end-haulage operations. While trucks start and end their route at a depot located at a terminal, they can be out for multiple days before returning to the depot. An overnight's minimum resting period is enforced at any location after a maximum daily active time. These operational considerations influence the routing decisions and should therefore be included.

In the next sections, operational capacity requirements are included in decisions on both routes of containers throughout a service network and intermodal drayage routes for pickups and deliveries. The aim is to analyse potential savings by using an integrated approach, and to provide insights in how to best utilise the current network for intermodal activities.

2.4 Problem formulation

An intermodal service operator considers a number of full-container requests which should be transported from their customer origin location to their customer destination location through an intermodal network with two large service regions. Within each service region, local pre- and end-haul operations should be performed. Between both service regions, long-haul transport is available. The aim is to minimise the total transport cost, which consists of both rail transport and trucking costs.

Each request has a known weight, length (or container type), and time window within which it can be transported, and must be routed throughout the intermodal

network from its customer origin location to its customer destination location. Time windows consist of an earliest pickup time at the customer origin location (i.e., release time of a request) and a latest arrival at the customer delivery location (i.e., due time of a request).

The long-haul network consists of a number of terminals in each service region. Links representing available services connect the terminals in periodical, usually weekly, planning cycles. Each rail service is characterised by a capacity limit, a departure terminal and an arrival terminal, a cutoff time before which containers using that service must arrive at the departure terminal and a time at which the containers are released at the arrival terminal, after which the containers are available for pickup by trucks for last-mile drayage. Multiple services with the same departure and arrival terminal may exist with different departure days and times. It is assumed that the available scheduled services are known and a long-haul service has to be selected for each load unit. The rail haul cost for transporting a request using a service depends on the service and the container type of each request.

Within each service region, trucks perform pickup and delivery operations. Each request requires two truck drayage tasks, one pickup task from its customer origin location to any terminal in the service region, and one delivery task from any terminal in another region to its customer destination location. In between two terminal nodes, different rail connections may be available. Truck depots are located at some of the terminals. Containers never stay at a customer location, as trucks wait for loading or unloading. A limited, heterogeneous fleet of trucks, with a capacity of a single container, is available, with as many container types as truck types. Two types of dimensions are considered, and this is assumed for each of the two service regions. This results in four sets of vehicles, each with a limited capacity. Separate truck routes for each type of container must be established. Truck costs are assumed to be proportional to the travelled distance. Trucks depart at the start of the planning horizon at a depot and return before the end of the planning horizon, consisting of multiple consecutive days, to the same depot. After performing pickup and delivery operations during a maximum daily active time, trucks should spend the night at any node in the network for a minimum overnight's rest time.

Flows in both directions between the service regions are considered to route containers in two ways between the regions, which implies that the truck routing in each of the regions includes both pickup tasks of containers at customers to deliver at terminals and delivery tasks of containers picked up at terminals to customer locations. pickup and delivery locations of containers are known while containers can be flexibly routed throughout the rail network.

This results in the following notation.

Sets and indices

$K = \{1, \dots, |K|\}$ = set of vehicles, index k

$D = \{1, \dots, |D|\}$ = number of days in the planning horizon, index d

$R = \{1, \dots, |R|\}$ = set of containers (requests), index r

$N^P = \{1, \dots, |R|\}$ = set of customer pickup locations, index g, h , one per request

$N^D = \{|R| + 1, \dots, 2|R|\}$ = set of customer delivery locations, index g, h ,
one per request, with delivery node $|R| + g$ related to pickup node $g \in N^P$

$N^T = \{2|R| + 1, \dots, 2|R| + m\}$ = set of m terminal nodes, index g, h

$N = \{1, \dots, 2|R| + m\}$ = node set with terminals, pickup and delivery nodes
 $= N^P \cup N^D \cup N^T$

$V = \{1, \dots, |V|\}$ = possible physical dimensions for containers/trucks/slots,
index v

S_{own} = set of own long-haul services

S_{ext} = set of externally owned long-haul services

S = set of long-haul services with index s between terminals,
where $S = S_{own} \cup S_{ext}$

Parameters

a_r = release time of container $r \in R$

b_r = due time of container $r \in R$

l_r = length dimension of container/truck/slot $r \in R$, with $l_r \in L$

w_r = weight of container $r \in R$

reg_g = region in which node $g \in N$ is located, $reg_g \in \{1, 2\}$

t_g = service duration at node g , i.e., (un)load time

t_{gh} = truck travel time between two nodes g and h ,
with $t_{gh} = \infty$ if $reg_g \neq reg_h$

c_{gh} = truck travel cost between two nodes g and h

L_s = maximum length allowed on service $s \in S_{own}$

W_s = maximum weight allowed on service $s \in S_{own}$

- C_{sl} = maximum number of containers of dimension l on service $s \in S_{ext}$
 λ_s^{dep} = departure terminal of service $s \in S$ with $\lambda_s^{dep} \in N^T$
 λ_s^{arr} = arrival terminal of service $s \in S$ with $\lambda_s^{arr} \in N^T$, $reg_{\lambda_s^{arr}} \neq reg_{\lambda_s^{dep}}$
 T_s^{dep} = cut-off time for load units to be at departure terminal for service $s \in S$
 T_s^{arr} = release time of load units at arrival terminal for service $s \in S$

The problem will be solved in the following two ways in order to compare an integrated approach with a sequential approach. The sequential method first assigns each request to a long-haul service. At this point, only direct truck cost from customer origin to departure terminal and from arrival terminal to customer destination are accounted for. A request can only be assigned to a service if the departure terminal can be reached in time with a direct truck route from the customer pickup location to the departure terminal, and the customer delivery location can be reached in time with a direct truck route from the arrival terminal to the customer delivery location. Furthermore, capacities of the services are accounted for. The objective is to find the least-cost assignment, where trucking costs are approximated using the direct distances between customer locations and the selected terminals. Consequently, a single pickup task - from a customer pickup location to a known departure terminal - and a single delivery task - from a known arrival terminal to a customer delivery location - are defined for each request. In a second phase, after the assigned service is fixed for each request, for each region these drayage tasks are combined into truck routes which minimise the total trucking cost. Only in this second phase, empty truck movements are accounted for. The integrated approach considers both the service selection and the truck routing simultaneously. Besides direct truck costs, costs for constructing vehicle routes, i.e., also costs for the movement of empty trucks, are accounted for in the decision on the long-haul rail service. This implies that the entire truck route costs are considered when assigning requests to services, instead of just the transport costs incurred by travel between a customer location and a terminal.

Both subproblems of the sequential approach are discussed in Section 2.4.1. The section presents a mathematical formulation for the intermodal long-haul routing problem for routing containers through a given intermodal network of long-haul services, as well as the vehicle routing problem used to minimise trucking costs for local pickups and deliveries of full containers. Finally, both problems are considered simultaneously in Section 2.4.2 in order to minimise the total operational cost of the intermodal truck-rail transport network using an integrated approach.

2.4.1 The sequential approach

2.4.1.1 The intermodal long-haul routing problem

Given the available periodic service schedules and characteristics, the aim of the intermodal routing problem is to determine the movement of requests throughout the given network in order to minimise total transport cost, consisting of the pre-haul trucking cost from a customer pickup location and the departure terminal, long-haul rail costs, and end-haul trucking from the arrival terminal to the customer delivery location. Each container r departs from a single, unique pickup node $g \in N^P$ (i.e., the customer origin location) after its release time a_r . First-mile local drayage is used for transport to a rail terminal node $h \in N^T$ in the same region. After arriving at a rail terminal in another region, the container is transported by truck to a unique delivery node $|R| + g \in N^D$, located at the customer destination location, before its due time b_r . Trucking costs are included by means of a cost which depends on the distance between a customer location and the selected terminal. Moreover, time windows and service departure times are accounted for. Our problem is formulated as follows:

Parameters

- c_{rs}^{pre} = trucking cost between customer origin location of r and departure terminal of service s
- c_{rs}^{end} = trucking cost between arrival terminal of service s and customer destination location of r
- c_{rs}^{rail} = rail cost for container r on service s
- t_{rs}^{pre} = trucking time between customer origin location of r and departure terminal of service s
- t_{rs}^{end} = trucking time between arrival terminal of service s and customer destination location of r

Decision variables

$$X_{rs} = \begin{cases} 1, & \text{if container } r \text{ uses rail service } s, \\ 0, & \text{else} \end{cases}$$

For a single planning period, which consists of a one-week planning cycle, the following problem can be formulated.

Problem P_1 :

$$\min \sum_{r \in R} \sum_{s \in S} (c_{rs}^{pre} + c_{rs}^{rail} + c_{rs}^{end}) \cdot X_{rs} \quad (2.1)$$

subject to

$$\sum_{s \in S} X_{rs} = 1 \quad \forall r \in R \quad (2.2)$$

$$\sum_{r \in R | l_r = l} X_{rs} \leq C_{sl} \quad \forall s \in S_{ext}; l \in L \quad (2.3)$$

$$\sum_{r \in R} l_r \cdot X_{rs} \leq L_s \quad \forall s \in S_{own} \quad (2.4)$$

$$\sum_{r \in R} w_r \cdot X_{rs} \leq W_s \quad \forall s \in S_{own} \quad (2.5)$$

$$a_r + t_r + t_{rs}^{pre} + t_{\lambda_s}^{dep} \leq T_s^{dep} + (1 - X_{rs}) \cdot M_{rs}^1 \quad \forall r \in R; s \in S \quad (2.6)$$

$$T_s^{arr} + t_{\lambda_s}^{arr} + t_{rs}^{end} + t_{r+R} \leq b_r + (1 - X_{rs}) \cdot M_{rs}^2 \quad \forall r \in R; s \in S \quad (2.7)$$

$$X_{rs} \in \{0, 1\} \quad \forall r \in R; s \in S \quad (2.8)$$

The objective is to minimise total transport costs (2.1). Exactly one service is to be selected for each load unit, as indicated by constraints (2.2). The total capacity between two rail nodes is limited by constraints (2.3)–(2.5). The heterogeneous slots purchased on external long-haul services are linked with containers of the same dimensions by constraints (2.3). For own long-haul services, constraints (2.4) and (2.5) impose a maximum length and weight respectively. It is assumed that a feasible assignment of containers to the specific locations on wagons can be found using a detailed train load planning algorithm as for example proposed in Chapters 4 to 6. Furthermore, each container r can only be assigned to a specific rail service if its release time a_r at the customer increased with the time needed to arrive at the considered terminal allows the container to arrive before the service departure (2.6). Each container can only leave a terminal after release at the arrival terminal of the selected service, and must arrive at its customer delivery location before its due date b_r (2.7). For these time constraints, $M_{rs}^1 = \max\{0; a_r + t_r + t_{rs}^{pre} + t_{\lambda_s}^{dep} - T_s^{dep}\}$, and $M_{rs}^2 = \max\{0; T_s^{arr} + t_{\lambda_s}^{arr} + t_{rs}^{end} + t_{r+R} - b_r\}$ respectively. Finally, the decision variables X_{rs} are binary (2.8).

2.4.1.2 Vehicle routing problem in each region

Within the scope of this problem, it is assumed that the terminal at which each container should be picked up or delivered, is known and fixed because the long-haul problem has been solved first. Total truck route costs in each of the service regions should be minimised, given pickup and delivery locations of containers. Additionally, separate truck routes are established for each vehicle type and related container di-

mensions v . This results in $2 \cdot |V|$ routing problems, one for each vehicle type in each region. The proposed formulation can solve these problems all at once.

Each container (or request) requires two drayage tasks. Each drayage task has a pickup node g and delivery node h , of which one is always a customer pickup or delivery location and the other a terminal location to drop off or pick up a container. Depots $i \in N^{Dep}$ at which trucks must start and end their routes are located at at least one terminal location per region. The resulting problem is a full-truckload pickup and delivery problem with time windows.

The problem can be formulated as an asymmetric multiple vehicle travelling salesman problem with time windows (am-TSP-TW) (Braekers et al., 2013), in which a pickup node g and its corresponding delivery node h are merged into a single node i and the distance between two nodes i differs depending on the travel direction. Using the notation of Braekers et al. (2013), each node $i \in O$ can be:

- a depot node ($i \in O^{Dep}$), O^{Dep} identical to N^{Dep}
- a pickup task ($i \in O_r^{pre}$), which consists of a customer pickup location ($g \in N^P$) and a terminal ($h \in N^T$), both located in the same service region.
- an delivery task ($i \in O_r^{end}$), which consists of a terminal ($g \in N^T$) and a customer delivery location ($h \in N^D$).

This implies the following characteristics of each node i :

- a service time $t_i = t_g + t_{gh} + t_h$, with $t_i = 0 \forall i \in O^{Dep}$.
- a cost of visiting a node $c_i = c_{gh}$, with $c_i = 0 \forall i \in O^{Dep}$.
- a node type l_i , related to the request dimensions. If the node is a pickup task ($g \in N^P$, and consequently $h \in N^T$), $l_i = l_g$, while $l_i = l_h$ if it is a delivery task ($h \in N^D$, and consequently $g \in N^T$).
- a time window $[a_i, b_i]$ within which service at node i should start. For pickup tasks ($g \in N^P$): $a_i = a_g$, $b_i = T_s^{dep} - t_i$. For delivery tasks ($h \in N^D$): $a_i = T_s^{arr}$, $b_i = b_h - t_i$.
- a rail service $\sigma_i \in S$, related to node i .

As a rail service σ_i is selected for each request in a previous phase, a single pickup and a single delivery node related to each request are known. Routes have to be established in which different nodes are linked, and each truck must start and end its route at a depot. In order to be able to limit the total maximum duration of a truck

route, each depot node $i \in N^{Dep}$ is duplicated two times the maximum number of vehicles of the considered type available at the depot. This results in d depot nodes, with $d = 2 * Vehicles * Depots$, as each route k must start at start depot $i = O_k^{Dep_s}$, and end at its related end depot $j = O_k^{Dep_e} = i + d/2$. Physical dimensions differ per truck, and the number of trucks of each dimension is limited in each region.

It would not be realistic to assume that drivers can continue their routes infinitely until the end of the planning horizon. To guarantee a limited daily active route duration and impose a minimum overnight rest, variable m_{ki}^d is added, which is a binary variable that indicates whether truck k spends the night following day d at node i . A binary variable α_i^d indicates whether the service in node i is performed at any day or time before night d ($\alpha_i^d = 1$) or after night d ($\alpha_i^d = 0$). In the formulation, the value of the variable α_i^d is only relevant if $\sum_{k \in K} m_{ki}^d = 1$. During the execution of a single drayage task (represented by the service of node i), the truck is loaded, and trucks are not allowed to take an overnight rest. As the last night of the planning horizon constitutes the end of the planning horizon, no overnight stay is included for the night after the last day of the planning horizon.

The compatibility between two nodes in a given route (i.e., for a single truck) cp_{ijk} indicates whether the truck type available for a route k is compatible with two nodes i and j . Depot nodes are characterised by the dimensions v of the truck linked to that route. Therefore, node i and j are compatible in the route of truck k ($cp_{ijk} = 1$) if their dimensions are identical ($l_i = l_j = l_k$), i.e., two identical types of requests with the same dimensions and a truck with the same dimensions, and if they are located in the same region ($i \neq j, reg_i = reg_j, reg_i = reg_k$). Finally, M is a sufficiently large number.

The following additional notation is used.

Sets:

O^{Dep_s} = set of start depots, index i, j

$O_k^{Dep_s}$ = start depot of vehicle k , index i, j

O^{Dep_e} = set of end depots, index i, j

$O_k^{Dep_e}$ = end depot of vehicle k , index i, j

O^{Dep} = set of depots, index i, j

$= \bigcup_k O_k^{Dep_s} \cup \bigcup_k O_k^{Dep_e}$, with one duplicate depot per route/vehicle k

O_r^{pre} = pickup task for request r , indices i, j

O_r^{end} = delivery task for request r , indices i, j

$$O = O^{Dep} \cup \bigcup_r O_r^{pre} \cup \bigcup_r O_r^{end}, \text{ set of nodes, indices } i, j$$

Parameters:

Q_d^a = earliest time at which trucks can depart after an overnight's rest
at the end of day d , with $1 \leq d < |D|$

Q_d^b = latest time at which trucks finish their active time before an overnight's
rest at the end of day d , with $1 \leq d < |D|$

T_{max} = maximum daily duration of a single truck route

T_{min} = minimum night's rest between routes on two consecutive days

$$cp_{ijk} = \begin{cases} 1, & \text{if nodes } i \in O, j \in O \text{ in route } k \text{ are compatible} \\ 0, & \text{else} \end{cases}$$

c_{ij}^T = trucking cost of travelling from the end of node i to the start of node j

c_i = cost of serving node i

t_i = service time at node i

σ_i = rail service related to pickup or delivery node $i \in O \setminus O^{Dep}$

t_{ij} = trucking time between nodes i and j

Decision variables

A_i = arrival time at node $i \in O$

B_i = begin time of service at node $i \in O$

D_i = time at which a vehicle leaves the destination location of task $i \in O$

BT_k^d = start time of route k on day $d + 1$, with $1 \leq d < |D|$

ET_k^d = end time of route k on day d , with $1 \leq d < |D|$

$$m_{ki}^d = \begin{cases} 1, & \text{if vehicle } k \text{ spends the night following day } d \text{ at node } i \\ 0, & \text{else} \end{cases}$$

$$\alpha_i^d = \begin{cases} 1, & \text{if service of } i \text{ is performed before the night following day } d \\ 0, & \text{else} \end{cases}$$

$$Y_{ijk} = \begin{cases} 1, & \text{if arc } (i, j) \text{ is used in route } k, \text{ with } cp_{ijk} = 1 \\ 0, & \text{else} \end{cases}$$

The am-TSP can be formulated as follows:

$$\min \sum_{k \in K} \sum_{i \in O} \sum_{j \in O} (c_{ij}^T + c_i) \cdot Y_{ijk} \quad (2.9)$$

subject to

$$Y_{ijk} \leq cp_{ijk} \quad \forall i, j \in O, k \in K \quad (2.10)$$

$$\sum_{k \in K} \sum_{j \in O} Y_{ijk} = 1 \quad \forall i \in O_r^{pre}, i \in O_r^{end}, r \in R \quad (2.11)$$

$$\sum_{j \in O} Y_{ijk} = \sum_{j \in O} Y_{jik} \quad \forall i \in O \setminus O^{dep}, k \in K \quad (2.12)$$

$$\sum_{j \in O \setminus O^{dep}} Y_{i_1jk} = \sum_{j \in O \setminus O^{dep}} Y_{ji_2k} \quad \forall k \in K, i_1 \in O_v^{Dep_s}, i_2 \in O_v^{Dep_e} \quad (2.13)$$

$$\sum_{k \in K} \sum_{j \in O} Y_{ijk} \leq 1 \quad \forall i \in O^{Dep_s} \quad (2.14)$$

$$\sum_{i \in O} m_{ki}^d = 1 \quad \forall 1 \leq d < |D|, k \in K \quad (2.15)$$

$$m_{ki}^d \leq \sum_{j \in O} Y_{ijk} \quad \forall i \in O, k \in K, 1 \leq d < |D| \quad (2.16)$$

$$a_i \leq B_i \leq b_i \quad \forall i \in O \quad (2.17)$$

$$A_i \leq B_i \quad \forall i \in O \quad (2.18)$$

$$B_i + t_i \leq D_i \quad \forall i \in O \quad (2.19)$$

$$D_i + t_{ij} \leq A_j + (1 - \sum_{k \in K} Y_{ijk}) \quad \forall i \in O, j \in O \quad (2.20)$$

$$B_i \geq Q_d^a - (1 - \sum_{k \in K} m_{ki}^d)M - \alpha_i^d \cdot M \quad \forall i \in O, 1 \leq d < |D| \quad (2.21)$$

$$B_i + t_i \leq Q_d^b + (1 - \sum_{k \in K} m_{ki}^d)M + (1 - \alpha_i^d)M \quad \forall i \in O, 1 \leq d < |D| \quad (2.22)$$

$$BT_k^d \leq D_i + (1 - m_{ki}^d)M + (1 - \alpha_i^d) \cdot M \quad \forall i \in O, k \in K, 1 \leq d < |D| \quad (2.23)$$

$$BT_k^d \leq B_i + (1 - m_{ki}^d)M + \alpha_i^d \cdot M \quad \forall i \in O, k \in K, 1 \leq d < |D| \quad (2.24)$$

$$ET_k^d \geq A_i - (1 - m_{ki}^d)M - \alpha_i^d \cdot M \quad \forall i \in O, k \in K, 1 \leq d < |D| \quad (2.25)$$

$$ET_k^d \geq B_i + t_i - (1 - m_{ki}^d)M - (1 - \alpha_i^d)M \quad \forall i \in O, k \in K, 1 \leq d < |D| \quad (2.26)$$

$$BT_k^d \geq Q_d^a \quad \forall k \in K, 1 \leq d < |D| \quad (2.27)$$

$$ET_k^d \leq Q_d^b \quad \forall k \in K, 1 \leq d < |D| \quad (2.28)$$

$$ET_k^1 - D_i \leq T_{max} \quad \forall k \in K, i \in O_k^{dep_s} \quad (2.29)$$

$$ET_k^d - BT_k^{d-1} \leq T_{max} \quad \forall k \in K, 1 < d < |D| - 1 \quad (2.30)$$

$$A_j - BT_k^{|D|-1} \leq T_{max} \quad \forall k \in K, j \in O_k^{dep_e} \quad (2.31)$$

$$BT_k^d - ET_k^d \geq T_{min} \quad \forall k \in K, 1 \leq d < |D| \quad (2.32)$$

$$BT_k^d, ET_k^d \geq 0 \quad \forall k \in K, 1 \leq d < |D| \quad (2.33)$$

$$\alpha_i^d \in \{0, 1\} \quad \forall i \in O, 1 \leq d < |D| \quad (2.34)$$

$$m_{ki}^d \in \{0, 1\} \quad \forall i \in O, k \in K, d \in D \quad (2.35)$$

$$Y_{ijk} \in \{0, 1\} \quad \forall i, j \in O, k \in K \quad (2.36)$$

The objective is to minimise total trucking costs (2.9) for the pickup and delivery of customer requests. Two nodes can only be connected if physical dimensions of nodes and vehicles are compatible (2.10). Each pickup and delivery node is visited once by a compatible vehicle (2.11). Constraints (2.12) imply at most one entering and one leaving vehicle for each pickup and delivery node. Constraints (2.13) link the start depot and end depot of each route. Only one route can depart from each start depot (2.14), which results in a single entering arc for each related end depot.

Constraints (2.15) determine for each vehicle and day in the planning horizon the node at which a night's rest is included, while this can only occur if that node has an outgoing arc (constraints (2.16)). Time windows for feasible service starts at each node must be respected (2.17)–(2.19). Constraints (2.20) set the time variables between subsequent nodes.

Constraints (2.21)–(2.26) represent three pairs of constraints of which either one or the other must hold, depending on whether a service in a node is executed at any time or day before or after spending a particular night at that node. They determine the service start within daily time windows (constraints (2.21)–(2.22)), daily route start times (constraints (2.23)–(2.24)) and daily route end times (constraints (2.25)–(2.26)). Constraints (2.27)–(2.28) imply a departure after the start of day d and a finish of the daily active time before the end of that day.

Furthermore, a limit is placed on the time a truck can operate in a single daily route (2.29)–(2.31). The minimum night's rest between arrival and departure for a given node at which a night's rest is spent is included by means of constraints (2.32). Finally, the decision variables BT_k^d and ET_k^d are defined as positive real numbers (2.33), and α_i^d , m_{ki}^d and Y_{ijk} must be binary (2.34)–(2.36).

2.4.2 The integrated intermodal container routing problem

In the integrated approach, truck routing and rail planning decisions are interdependent and made simultaneously. Now, truck cost are immediately defined as the

total cost of combining pickup and delivery tasks into truck routes as considered in the VRP. In the sequential problem, initially solely transport operations between each customer location and the assigned terminal are included in the long-haul routing problem, after which only the second subproblem (i.e., the VRP) considers total trucking distances.

For a single request r , multiple pickup tasks O_r^{pre} and delivery tasks O_r^{end} are available, related to the available services. A single, feasible pickup and delivery task have to be performed in order to transport a request from its customer origin to its destination. A new variable $Z_i = 1$ if a drayage task i should be performed as a consequence of selecting the related long-haul service. All feasible drayage tasks serve as input to the problem and are defined as follows. A drayage task i for pickup operations (i.e., a combination of a pickup location $g \in N^P$ with a rail departure terminal $h \in N^T$ related to available service s , i.e., $\lambda_s^{dep} = h$) is feasible if the pickup location and departure terminal of a service are located in the same service region ($reg_g = reg_{\lambda_s^{dep}}$), and if a timely arrival at the delivery location $g + n$ is possible (i.e., $a_g + t_g + t_{g, \lambda_s^{dep}} + t_{\lambda_s^{dep}} \leq T_s^{dep}$ and $T_s^{arr} + t_{\lambda_s^{arr}} + t_{\lambda_s^{arr}, g+n} + t_{g+n} \leq b_g$).

A drayage task i for delivery operations (i.e., a combination of a rail arrival terminal $g \in N^T$ of an available service s , i.e., $\lambda_s^{arr} = g$, with a delivery location $h \in N^D$) is feasible if the pickup location and departure terminal of a service are located in the same service region ($reg_{\lambda_s^{arr}} = reg_h$), and if a timely arrival at the delivery location h is possible (i.e., $a_{h-|R|} + t_{h-|R|} + t_{h-|R|, \lambda_s^{dep}} + t_{\lambda_s^{dep}} \leq T_s^{dep}$ and $T_s^{arr} + t_{\lambda_s^{arr}} + t_{\lambda_s^{arr}, h} + t_h \leq b_h$). The following additional notation is used:

Decision variables:

$$Z_i = \begin{cases} 1 & \text{if node } i \text{ is visited} \\ 0 & \text{else} \end{cases}$$

The problem is formulated as follows:

$$\min \sum_{r \in R} \sum_{s \in S} c_{rs}^{rail} \cdot X_{rs} + \sum_{k \in K} \sum_{i \in O} \sum_{j \in O} c_{ij}^T \cdot Y_{ijk} + \sum_{i \in O} c_i \cdot Z_i \quad (2.37)$$

subject to

Rail constraints:

$$\sum_{s \in S} X_{rs} = 1 \quad \forall r \in R \quad (2.38)$$

$$\sum_{r \in R | l_r = l} X_{rs} \leq C_{sl} \quad \forall s \in S_{ext}; l \in L \quad (2.39)$$

$$\sum_{r \in R} l_r \cdot X_{rs} \leq L_s \quad \forall s \in S_{own} \quad (2.40)$$

$$\sum_{r \in R} w_r \cdot X_{rs} \leq W_s \quad \forall s \in S_{own} \quad (2.41)$$

Trucking constraints:

$$Y_{ijk} \leq cp_{ijk} \quad \forall i, j \in O, k \in K \quad (2.42)$$

$$\sum_{k \in K} \sum_{j \in O} Y_{ijk} = Z_i \quad \forall i \in O \quad (2.43)$$

$$\sum_{j \in O} Y_{ijk} = \sum_{j \in O} Y_{jik} \quad \forall i \in O \setminus O^{dep}, k \in K \quad (2.44)$$

$$\sum_{j \in O \setminus O^{dep}} Y_{i_1jk} = \sum_{j \in O \setminus O^{dep}} Y_{ji_2K} \quad \forall k \in K, i_1 \in O_k^{Dep_s}, i_2 \in O_k^{Dep_e} \quad (2.45)$$

$$\sum_{i \in O} m_{ki}^d = 1 \quad \forall 1 \leq d < |D|, k \in K \quad (2.46)$$

$$m_{ki}^d \leq \sum_{j \in O} Y_{ijk} \quad \forall i \in O, k \in K, 1 \leq d < |D| \quad (2.47)$$

$$a_i \leq B_i \leq b_i \quad \forall i \in O \quad (2.48)$$

$$A_i \leq B_i \quad \forall i \in O \quad (2.49)$$

$$B_i + t_i \leq D_i \quad \forall i \in O \quad (2.50)$$

$$D_i + t_{ij} \leq A_j + (1 - \sum_{k \in K} Y_{ijk}) \quad \forall i \in O, j \in O \quad (2.51)$$

$$B_i \geq Q_d^a - (1 - \sum_{k \in K} m_{ki}^d)M - \alpha_i^d \cdot M \quad \forall i \in O, 1 \leq d < |D| \quad (2.52)$$

$$B_i + t_i \leq Q_d^b + (1 - \sum_{k \in K} m_{ki}^d)M + (1 - \alpha_i^d)M \quad \forall i \in O, 1 \leq d < |D| \quad (2.53)$$

$$BT_k^d \leq D_i + (1 - m_{ki}^d)M + (1 - \alpha_i^d) \cdot M \quad \forall i \in O, k \in K, 1 \leq d < |D| \quad (2.54)$$

$$BT_k^d \leq B_i + (1 - m_{ki}^d)M + \alpha_i^d \cdot M \quad \forall i \in O, k \in K, 1 \leq d < |D| \quad (2.55)$$

$$ET_k^d \geq A_i - (1 - m_{ki}^d)M - \alpha_i^d \cdot M \quad \forall i \in O, k \in K, 1 \leq d < |D| \quad (2.56)$$

$$ET_k^d \geq B_i + t_i - (1 - m_{ki}^d)M - (1 - \alpha_i^d)M \quad \forall i \in O, k \in K, 1 \leq d < |D| \quad (2.57)$$

$$BT_k^d \geq Q_d^a \quad \forall k \in K, 1 \leq d < |D| \quad (2.58)$$

$$ET_k^d \leq Q_d^b \quad \forall k \in K, 1 \leq d < |D| \quad (2.59)$$

$$ET_k^1 - D_i \leq T_{max} \quad \forall k \in K, i \in O_k^{dep_s} \quad (2.60)$$

$$ET_k^d - BT_k^{d-1} \leq T_{max} \quad \forall k \in K, 1 < d < |D| - 1 \quad (2.61)$$

$$A_j - BT_k^{|D|-1} \leq T_{max} \quad \forall k \in K, j \in O_k^{dep_e} \quad (2.62)$$

$$BT_k^d - ET_k^d \geq T_{min} \quad \forall k \in K, 1 \leq d < |D| \quad (2.63)$$

Linking constraints:

$$Z_i = X_{r\sigma_i} \quad \forall r \in R; i \in O_r^{pre}, i \in O_r^{end} \quad (2.64)$$

Domains:

$$BT_k^d, ET_k^d \geq 0 \quad \forall k \in K, 1 \leq d < |D| \quad (2.65)$$

$$\alpha_i^d \in \{0, 1\} \quad \forall i \in O, 1 \leq d < |D| \quad (2.66)$$

$$m_{ki}^d \in \{0, 1\} \quad \forall i \in O, k \in K, d \in D \quad (2.67)$$

$$X_{rs} \in \{0, 1\} \quad \forall r \in R; s \in S \quad (2.68)$$

$$Y_{ijk} \in \{0, 1\} \quad \forall i, j \in O, k \in K \quad (2.69)$$

$$Z_i \in \{0, 1\} \quad \forall i \in O \quad (2.70)$$

The objective is to minimise total trucking and rail costs (2.37) for the delivery of customer requests. Constraints (2.38)–(2.41) relate to (2.2)–(2.5) of the intermodal routing problem, where one of the feasible services for each request must be selected (with for each feasible service, one feasible pickup and one feasible delivery task). Trucking operations are modelled by means of constraints (2.42)–(2.63), related to constraints (2.10)–(2.32). As multiple pickup and delivery tasks are often available for each request r , a single pickup node $i \in N_r^{pre}$ and delivery node $i \in N_r^{end}$ should be visited, which is implicitly included. The link between truck route times and rail departures and arrivals are included in the definition of the TSP-nodes. Precedence constraints (2.64) link a single pickup and delivery task per request to its related rail service. Finally, the decision variables must be positive real numbers (2.65), or binary (2.66)–(2.70).

2.5 Conclusions

Many intermodal planning problems and decision levels are interrelated, as recently stressed in literature. In this chapter, literature related to decisions on local drayage routes and on long-haul intermodal routes is described and research gaps related to the level of integration between both problems are identified.

Both local drayage and intermodal long-haul routing problems are formulated in a sequential approach. In this approach, pickup and delivery tasks for truck routes are

fixed and known, as long-haul routing decisions have been made in advance. Next, both decisions are made simultaneously in an integrated approach. Its formulation is presented as an integrated intermodal routing problem. In that case, pickup and delivery tasks are not known in advance, as the departure and arrival terminals through which each request travels depend on the selected long-haul service. Furthermore, capacity considerations are added which are generally not dealt with in intermodal routing, such as length and weight limits. This is in line with operational problems such as the train load planning problem, which does include a very detailed level of capacity limits. The integrated planning approach, with capacity considerations for trucking and rail planning, is aimed at maximising the overall long-haul network capacity utilisation and minimising total transport costs.

For intermodal transport to be advantageous, an integrated planning should be executed efficiently in order to minimise total transport costs and maximise service capacity utilisation, which in turn also results in decreasing costs of the transport system. Costs related to the execution of first- and last-mile drayage operations are an important aspect within this regard. At the same time, customer requirements should be respected. Preliminary experiments showed that even problems with a very small number of requests (i.e., 10 to 15 requests) could not be solved to optimality in short computation times. As the underlying vehicle routing problem is NP-hard (Lenstra and Kan, 1981), solving the problem exactly for realistic instance sizes would result in very large computation times. Additionally, the scheduling, and more specifically the timing and multi-day aspects of the vehicle routing problem make the integrated intermodal routing problem hard to solve. The number of feasible drayage tasks (i.e., a pickup and a delivery task for each service to which a request can be feasibly assigned) increases substantially with the number of requests and the number of feasible services per request. Although the number of drayage tasks executed per day will be relatively low, the vehicle schedule of one day impacts the next day's schedule start due to a maximum daily active time and a minimum amount of overnight's rest. Therefore, a heuristic algorithm is developed in Chapter 3 to solve the integrated as well as the sequential planning problem in order to provide decision support for planners in such complex environment, and quantify potential savings of an integrated approach.

A large neighbourhood search heuristic for the integrated intermodal routing problem

3.1 Introduction

In Chapter 2, the integrated intermodal routing problem is introduced, which focuses on the simultaneous decision on intermodal routing and local drayage. A mathematical formulation for both the sequential and integrated approach are presented. In the current chapter (Figure 3.1), a large neighbourhood search algorithm is developed to solve the problem using both approaches. Section 3.2 proposes the structure of the large neighbourhood search heuristic, including a discussion on the multi-day scheduling procedure required to find feasible truck schedules. A feasible truck schedule has to respect a minimum overnight's rest and a maximum daily active time of trucks, due to the fact that trucks can be out for multiple days. In Section 3.3, the generated instances and the parameter tuning phase are discussed. Next, the sequential and integrated approach are compared by conducting computational experiments. Using the integrated approach, Section 3.4 presents a real-life case study and focuses on analysing the impact of tactical service network design decisions made by an intermodal service operator on the total operational costs. The main conclusions are summarised in Section 3.5.

This chapter is based on the following paper: Heggen, H., Molenbruch, Y., Caris, A., Braekers, K., 2019. Intermodal Container Routing: Integrating Long-Haul Routing and Local Drayage Decisions. *Sustainability*, 11, 1634.

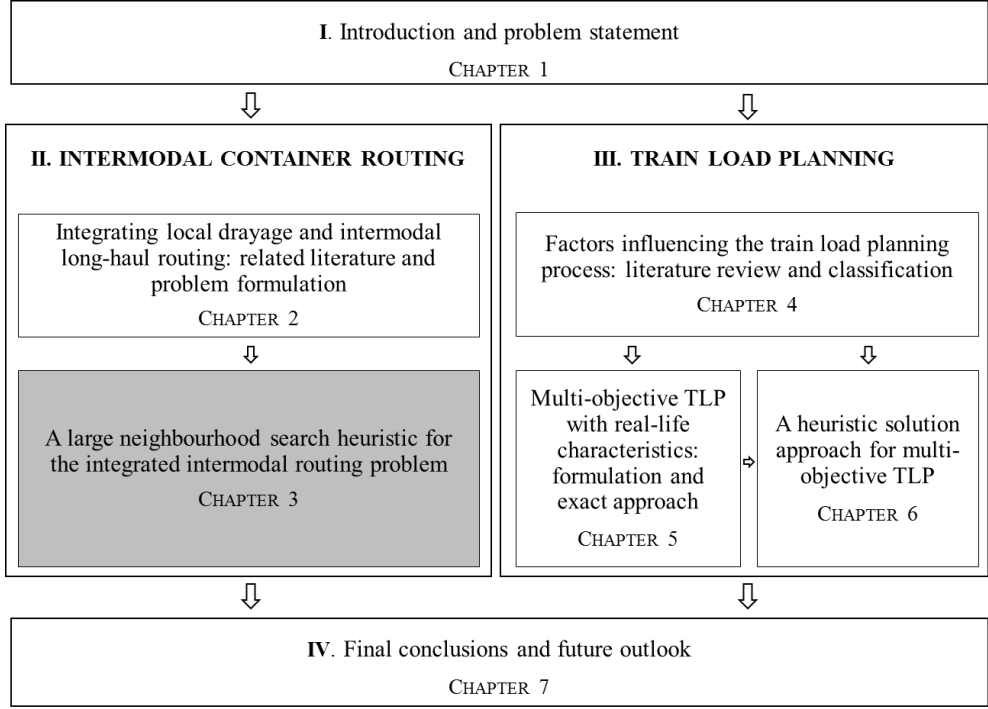


Figure 3.1: Thesis summary - Chapter 3.

3.2 Large neighbourhood search heuristic

A large neighbourhood search (LNS) (Shaw, 1998) algorithm is developed to solve the intermodal routing, in which a large number of requests is removed, and next reinserted into the solution in each single iteration. The framework has proven to lead to good results for vehicle routing problems and pickup and delivery problems (Ropke and Pisinger, 2006). Such large changes are desirable in the intermodal routing problem due to dependencies of different problem aspects, related to the integrated nature of the problem. Both a sequential and an integrated version of the heuristic algorithm are presented. The sequential approach serves as a benchmark for the improvements that may be obtained by the integrated intermodal routing problem. The heuristic building blocks are similar for both approaches, but different parts of the solution are destroyed and repaired. For each request, one pickup drayage task (from a customer pickup location to a departure terminal) and one delivery drayage task (from an arrival terminal to a customer delivery location) have to be executed. Both tasks relate to the selected long-haul service. Whereas operators in a sequential setting remove and reinsert individual, fixed drayage tasks (with predetermined terminals),

an integrated setting removes and inserts both the pickup and delivery task of any request simultaneously, as well as the selected long-haul service.

Moreover, in practice, the planning for one time period is executed, while an initial planning for future time periods is included, to be used as a starting point for the following time period. We pay attention to this dynamic nature of the problem by including information from past week's planning as well as the initial planning for the following week. If this information would be excluded, besides the fact that this is unrealistic, many pickup tasks would occur at the beginning of the current week, and many delivery tasks would need to be executed at the end of the week. Moreover, this would result in a limited choice of long-haul services, with a lower capacity utilisation for long-haul services with departure in the beginning and at the end of the week, compared to long-haul services in the middle of the week. First, we include requests which were already assigned to a long-haul service with departure in the previous week. For these requests, only the truck delivery task still has to be executed after arrival at the destination terminal. In practice, this is given input for current week's planning. Therefore, it is assumed that these requests are assigned to long-haul services in advance, and only the delivery trucking cost is relevant for this week's problem. The assigned long-haul service and related delivery task do not vary during the execution of the heuristic. The delivery task should only be assigned to a position in a truck route. Second, some requests are assigned to long-haul services which depart in the current planning cycle and arrive in the next. For this type of assignment, the long-haul service is selected by the heuristic. Only the pickup task should be included in truck routes. Delivery tasks in the next week are included by means of the two-way direct distance between the destination terminal and the customer destination, plus a high value M in order to avoid postponing requests. At the same time, this ensures that if postponement is required, a smaller end-haul trucking distance is preferred. For these requests, total assignment costs consist of costs for including the pickup task in truck routes (VRP), long-haul service costs, and direct two-way end-haul truck costs.

The general heuristic structure is presented in Section 3.2.1, after which the removal and insertion operators are described in Section 3.2.2. Preliminary feasibility checks for inserting a single node into a candidate solution are explained in Section 3.2.3. If these checks do not indicate an infeasible solution, a more complex, multi-day scheduling procedure aims at finding a feasible truck schedule for the candidate solution. The procedure is discussed in detail in Section 3.2.4. These checks are performed in this sequence, with the least time-consuming checks first, as it is hard to determine in advance whether a solution is feasible in a multi-day scheduling context

with a minimum overnight's rest and a maximum daily active duration.

3.2.1 General structure of the LNS

Algorithm 1 provides the general heuristic structure. It includes a constructive heuristic (lines 2–4) as well as a framework with a number of removal operators (lines 9–13) and insertion operators (lines 14–17).

Algorithm 1 Heuristic structure.

```

1: Set parameters:  $maxIt$ ,  $deviation$ 
2: CONSTRUCT INITIAL SOLUTION  $sol_0$ 
3:   Solve exact sequential long-haul problem
4:   Apply random order best insertion
5:  $sol_{best} = sol_0$ ;  $sol_{curr} = sol_0$ 
6: LARGE NEIGHBOURHOOD SEARCH
7: for  $it = 0$  to  $maxIt$  do
8:    $sol_{new} = sol_{curr}$ 
9:   Randomly select a DESTROY operator  $d(\cdot)$  from:
10:     Random removal
11:     Worst removal
12:     Random order removal of linked tasks or requests
13:     Related removal of tasks or requests from related long-haul services
14:   Randomly select a REPAIR operator  $r(\cdot)$  from:
15:     Random order best insertion
16:     Greedy insertion
17:     Two-regret insertion
18:    $sol_{new} \leftarrow r(d(sol_{new}))$ 
19:   if  $sol_{new}$  is complete and
20:      $sol_{new}.vrpCost + sol_{new}.directTruckCost + sol_{new}.railCost$ 
21:      $< sol_{curr}.vrpCost(1+deviation) + sol_{curr}.directTruckCost + sol_{curr}.railCost$ 
22:   then
23:      $sol_{curr} = sol_{new}$ 
24:     if  $sol_{new}.TotalCost < sol_{best}.TotalCost$  then
25:        $sol_{best} = sol_{new}$ 
26:     end if
27:   end if
28: end for

```

The constructive heuristic consists of first assigning requests to long-haul services by solving the long-haul routing problem using integer linear programming using ILOG Cplex 12.6 in C++, and next inserting for each request its pickup and delivery task - related to the known, assigned long-haul services - into truck routes. These given drayage tasks are inserted in a random order at their best feasible insertion position, i.e., with the lowest additional cost, in order to obtain a good initial solution. This corresponds to the random order best insertion operator of the sequential approach.

In each iteration, one removal operator is selected randomly in which requests are removed until a given threshold is reached with respect to the number of removed requests or tasks. Next, one insertion operator is selected randomly. Requests are reinserted in order to obtain a new, complete solution. In a sequential approach (Seq.), after requests are assigned to long-haul services in the constructive phase, minimum-cost truck routes are established with known pickup and delivery tasks. Sequential operators aim at removing or inserting single drayage tasks, instead of requests, and the assigned long-haul services remain unchanged. In an integrated approach (Int.), destroy and repair operators consider removing and reinserting requests from the solution. More information on these integrated operators is presented in Section 3.2.2.

The acceptance criterion for moves includes a function with multiple cost components (lines 19–21). The total costs of each solution consist of vehicle routing costs in both service regions ($sol_{new}.vrpCost$), two-way direct truck costs for drayage tasks to be executed in the next planning period as a result of long-haul services with arrival in the next week ($sol_{new}.directTruckCost$), and long-haul transport costs ($sol_{new}.railCost$). In comparison with the objective function, this cost function omits the artificially high value assigned to load units with an arrival in the future planning period to avoid postponement as much as possible. This value is excluded from the acceptance criterion in order to avoid having an extensive impact on the evaluation of each solution due to its relatively high value. Moreover, for accepting new solutions, a percentage deviation is only included for the vehicle routing cost ($sol_{new}.vrpCost \cdot (1 + deviation)$) in order to focus on the optimisation of vehicle routes in each service region. A small percentage allowed deviation in a cost with a higher value, such as the two-way direct truck cost, would always be beneficial, while the aim is also to reduce routing costs in each region.

3.2.2 Integrated operators

In this section, the operators used within the large neighbourhood search framework are described. Operators which are well-known and have shown to obtain good results for pickup and delivery problems (random and worst removal, random order best insertion, greedy and two-regret insertion) (Ropke and Pisinger, 2006), as well as operators designed to fit the characteristics of the problem are included (random order removal of linked tasks/requests, related removal from related long-haul services).

Integrated operators, which are dedicated to the integrated nature of the problem, are explained in this section. Sequential operators are not discussed individually, but use the same building blocks. The only difference is that in the sequential approach, improvements in truck routes are considered without changing selected long-haul services, as this decision is made in advance. To summarise, sequential operators focus on individual drayage tasks, whereas integrated operators address the entire inter-modal route of a request, i.e., the pickup and delivery task as well as the assigned long-haul service. Four removal operators and three insertion operators are included and described next. The removal procedure is repeated until the number of removed requests or drayage tasks, for the integrated and sequential approach respectively, reaches a given threshold. The process of inserting uninserted requests or drayage tasks is repeated until either all requests are inserted or no feasible insertion can be found for a request. If no feasible insertion can be found, the solution is incomplete and the iteration ends.

Random removal of requests: Randomly selected requests are removed from the solution until a given threshold is reached. This implies removal of its drayage tasks from truck routes in the pickup and delivery region, as well as removal from its assigned long-haul service.

Worst removal of requests: Requests are ordered based on the savings that could be obtained by excluding that request from the solution, i.e., if its pickup task and delivery task are omitted from truck routes and the request is removed from the previously assigned long-haul service. The request with the largest savings is removed and savings are updated.

Random order removal of linked requests: Randomly selected requests are removed from the solution. For each randomly selected request r , both the pickup and delivery task are removed, and the request is removed from its assigned long-haul service. Further, if any delivery task immediately precedes the pickup task of request r in its current truck route, the request related to that delivery task is removed too. Similarly, if any pickup task immediately succeeds the delivery task of request r , the

request related to that pickup task is removed. Obviously, in the sequential variant of this operator, if a pickup task is selected, the delivery task executed right before is removed if available. For selected delivery tasks, a single pickup task executed afterwards is removed, again if available. By removing subsequent pickup and delivery tasks, subroutes (or street turns) between two terminal visits can be removed from the solution.

Related removal of requests: This operator removes similar requests assigned to similar long-haul services, as this might increase the probability of finding new solution structures in a reinsertion phase. Selecting requests on similar long-haul services is also likely to result in available capacity on long-haul services which are more likely to be feasible for the removed requests. In each iteration, one long-haul service is selected randomly, as well as its two most related long-haul services. Obviously, a long-haul service can only be selected if at least some of the available capacity is used. The relatedness between two long-haul services s_1 and s_2 is measured by means of four concepts: the absolute difference in departure time T^{dep} , the absolute difference in arrival time T^{arr} , the distance between the locations of the departure terminals λ^{dep} and the distance between the locations of the arrival terminals λ^{arr} . Next, each value is normalised (i.e., each measure is divided by its maximum dissimilarity value $maxdissim$), and the four values are summed up. Noise for selecting long-haul services is added by multiplying the overall dissimilarity score of two long-haul services $dissim(s_1, s_2)$ with a random number between 0.8 and 1.2, in order to allow for more variation in the solution space. A higher dissimilarity score indicates a larger dissimilarity between two long-haul services. Consequently, a smaller value is preferred during the selection of similar long-haul services. The dissimilarity between two long-haul services s_1 and s_2 can be calculated as follows:

$$dissim(s_1, s_2) = dissimT^{dep}(s_1, s_2) + dissimT^{arr}(s_1, s_2) + dissimdepDist(s_1, s_2) \\ + dissimarrDist(s_1, s_2)$$

The calculation of each term is presented below:

$$dissimT^{dep}(s_1, s_2) = \frac{|T_{s_1}^{dep} - T_{s_2}^{dep}|}{maxdissimT^{dep}(s_1, s_2)}$$

$$dissimT^{arr}(s_1, s_2) = \frac{|T_{s_1}^{arr} - T_{s_2}^{arr}|}{maxdissimT^{arr}(s_1, s_2)}$$

$$dissimdepDist(s_1, s_2) = \frac{dist(\lambda_{s_1}^{dep}, \lambda_{s_2}^{dep})}{maxdissim.depDist(s_1, s_2)}$$

$$dissimarrDist(s_1, s_2) = \frac{dist(\lambda_{s_1}^{arr}, \lambda_{s_2}^{arr})}{maxdissim.arrDist(s_1, s_2)}$$

Next, from the pool of requests assigned to these selected, similar long-haul services (with low dissimilarity scores), requests with the highest relatedness are removed first until either all requests on these long-haul services are removed or the threshold of requests to remove is reached. The relatedness of two requests r_1 and r_2 is based on the average distance between pickup and delivery locations of both requests in the same service region, with $distRegion1(r_1, r_2)$ the distance between the tasks of both requests located within the same region (similar for the second region), and the average percentage overlap in time window between two requests $TWoverlap(r_1, r_2)$. Noise is included during this process in a similar way as for the dissimilarity of long-haul services. The dissimilarity between two requests $dissim(r_1, r_2)$ is calculated as follows:

$$dissim(r_1, r_2) = dissimloc(r_1, r_2) + dissimTW(r_1, r_2)$$

with the definition of each term calculated indicated below:

$$\begin{aligned} dissimloc(r_1, r_2) &= \frac{(|distRegion1(r_1, r_2)| + |distRegion2(r_1, r_2)|)/2}{maxdissimloc(r_1, r_2)} \\ TWoverlap(r_1, r_2) &= \max\{0; \min\{b_{r_1}, b_{r_2}\} - \max\{a_{r_1}, a_{r_2}\}\} \\ dissimTW(r_1, r_2) &= \frac{1 - (\frac{TWoverlap(r_1, r_2)}{b_{r_1} - a_{r_1}} + \frac{TWoverlap(r_1, r_2)}{b_{r_2} - a_{r_2}})/2}{maxdissimTW(r_1, r_2)} \end{aligned}$$

In the sequential approach, the dissimilarity between two drayage tasks is related to the average of the distance between two customer locations and the distance between two terminals. As the long-haul service is fixed in advance and cannot change, this calculation is not explicitly based on long-haul service characteristics. The time window dissimilarity is measured using $\frac{\max\{0; a_{n1} - a_{n2}\}}{maxdissimTW(n_1, n_2)}$, using the difference between the release times of two nodes n_1 and n_2 .

Random order best insertion of requests: Uninserted requests are selected randomly. For each request, the best long-haul service is selected, defined by taking an integrated perspective. The sum of the long-haul rail cost and the best corresponding pre- and end-haul cost for inserting tasks into truck routes determine the total cost of assigning a request to a long-haul service and inserting it into a route in each region.

Greedy insertion of requests: First, for all uninserted requests, the best insertion is calculated, defined by the best long-haul service, insertion of the pickup node and

insertion of delivery node, again taking an integrated perspective. Next, requests are inserted based on their insertion costs. The request with the lowest total insertion cost is inserted in the solution, after which insertion costs for the remaining requests are updated if required.

Two-regret insertion of requests: The two-regret insertion operator works similar to the greedy insertion. However, in this case both the best and second-best insertion cost are calculated. Next, the request with the highest difference between the lowest and the second-lowest total insertion cost is inserted in the solution. Again, insertion costs are updated if required. The idea is to first insert the requests which are the most difficult to insert in order to avoid that they cannot be inserted feasibly anymore if more requests are inserted.

3.2.3 Preliminary checks for inserting a node

Inserting a node n at its best insertion position, i.e., after predecessor i currently assigned to a route, can only be feasible if a number of conditions hold. These conditions focus on a single insertion. First, only insertion positions within the same service region are considered ($reg_i = reg_n$). Next, an initial check on the time windows examines whether $a_i + t_i + t_{in} \leq b_n$ between a predecessor i and n , and whether $a_n + t_n + t_{nj} \leq b_j$ for inserting n before the current successor of i (denoted with j). If also the dimensions of both nodes match ($l_i = l_n$), it is checked whether a node n can be feasibly inserted between node i and its current successor j based on a regular calculation of the earliest start ES (forward from predecessor i , with $ES_n = \max\{a_n; ES_i + t_i + t_{in}\}$) and latest start LS (backward from the current successor of i , with $LS_n = \min\{b_n; LS_j - t_{nj} - t_n\}$), respecting its time window $[a_n, b_n]$.

3.2.4 Feasibility of new routes

Only if all conditions in the preliminary check are satisfied (see Section 3.2.3) – which are required, but insufficient conditions for feasibility –, another procedure determines whether a feasible schedule can be found for a certain truck with a given sequence of nodes in a candidate solution. As trucks can be out for multiple days, it would not be realistic to assume that they are active for an entire week. Therefore, the scheduling procedure for multiple days should include a minimum overnight's rest and a maximum daily active time, as considered in the problem formulation in Chapter 2. The day at which a task should be executed is not fixed in advance (within given release and due times of a request) and, as such, is part of the decision process.

Nodes in a given sequence are added to a given day's route using the earliest feasible start as long as the daily limit on the route duration T_{max} is not violated. If not feasible, the forward time slack principles, explained in detail later on, are used to shift the route schedule within a single day. It is always better to travel to and serve a node on the current day than shifting one or both of these activities to a future day. Therefore, if a feasible route can be composed on a particular day by reducing that day's route duration, and thus postponing the start of that route, this is always executed. In this part of the procedure, the earliest possible arrival in the last node of the day is assumed to be given and does not shift. Hence, the shift does not restrict the feasibility of the next day. Only if shifting the start of the route is not feasible, travel to and/or service of the node under consideration is tried to be performed on a future day and the entire procedure is repeated for that future day. If all nodes in a route can be feasibly visited, the procedure provides a schedule which satisfies a maximum daily route duration T_{max} , a minimum night's rest T_{min} and daily time windows Q_d^b and Q_d^a between which trucking is allowed.

Within this procedure, four types of feasible scheduling options are considered for each node to be added to a route. These are explained below. The procedure evaluates a given sequence of nodes, where nodes i and j are assumed to be two subsequent nodes in the sequence. If adding a new node j does not result in any infeasibility, the next node in the sequence will be evaluated (i.e., $i = j$ and $j = \text{successor}_j$, and the procedure is repeated).

The procedure requires two modifications to the forward time slack (F) principles of Savelsbergh (1992) and Vidal et al. (2015). The forward time slack (F) indicates the time with which each node served on a particular day can be shifted later in time such that a feasible route can be composed serving all nodes before (and including) the currently considered node based on the earliest start procedure, which calculates the earliest feasible start of a task at any node in a sequence using time windows, service times and travel times. In order to calculate the forward time slack, both the earliest feasible execution date T_i and the cumulative idle time W_i are required for a given node i . A first modification, $F^*_{service}$, is explained in case 2 below. A second modification, $F^*_{noservice}$ is used in case 3.

Case 1: Travel and service on current day d without changing service start of nodes in route of current day If adding travel to and service of node j to the current day's route results in a feasible route duration (i.e., $T_{service} \leq T_{max}$, with $T_{service}$ equal to the route duration of the current day) and if it can be performed before the end of the current day Q_d^b , node j is added to the current day's route if

this does not result in any infeasibilities of the route of future nodes.

Case 2: Travel and service on current day d with removal of excess waiting time in route of current day If adding a node j with the regular earliest feasible start results in a violation of the maximum daily route duration ($T_{service} > T_{max}$), it is checked whether it is possible to postpone the service start of previous nodes in the route of the current day such that node j can be feasibly included on the current day's route within the maximum daily route duration. If feasible, this is always preferred compared to serving it on a future day.

In this case, a first modification to the regular forward time slack $F^*_{service} = \min\{F_i, W_j\}$ is performed. Using the regular time slack F could influence the start of the minimum overnight rest and consequently the start time of the next day, which might lead to an infeasible schedule on the next day. Therefore, one difference with the regular time slack principles is the fact that the service start of the currently considered node j remains unchanged when attempting to shift the start of a day's route. In order to feasibly include travel to and service at node j , the maximum daily route duration should be respected after shifting the current route ($F^*_{service} \geq T_{service} - T_{max}$), and the truck should be able to finish service before the end of that day ($ES_j + t_j \leq Q^b_d$). Figure 3.2 visualises such a feasible shift. Initially, inserting node j is infeasible with a given start of the day, in this case represented by the start of service in node i . However, by postponing the start of a route on a given day with $F^*_{service}$, it becomes feasible to travel to and serve node j , without changing its earliest feasible service start. At the same time, excessive wait time on the route of the current day is reduced.

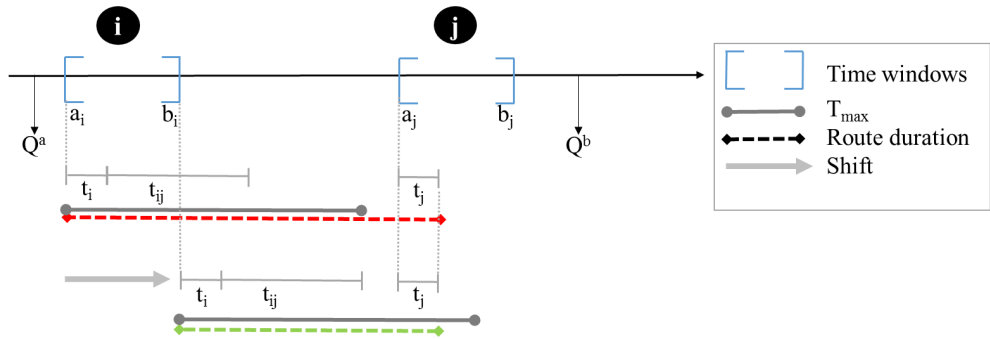


Figure 3.2: Shift the start of a route on a single day with $F^*_{service}$, including travel to and service of node j on the current day.

Case 3: Travel on current day d with removal of excess waiting time in route of current day, service on a future day If adding a node j with the current earliest feasible start results in a violation of the maximum daily route duration (i.e., case 2 is infeasible), it is checked whether it is possible to postpone the service start of previous nodes in the route of the current day such that the travel to node j can be feasibly included on the current day within the maximum daily route duration.

An important aspect within this regard is the fact that a minimum night's rest might need to be added before or after serving (i.e., travel towards a node can occur on that day, before actually serving that node on a future day). However, a truck is only allowed to rest with an empty container. Therefore, if postponing the route with $F^*_{service}$ is infeasible, a second modification to the forward time slack $F^*_{noservice} = \min\{F_{predec_i}, W_i\}$ is performed, which is equal to zero if no predecessor of i is available on the current day (i.e., no node is served before this node on the current day). It is used to check whether it is feasible to only include the travel to node j on the current day, while service occurs on any future day. Furthermore, the service start of the last node on the current day (represented by node i in Figure 3.3) is not allowed to shift, as this could influence the feasibility of the next day's route. Again, three conditions must hold: the current day should not be the last day in the planning horizon ($d < D$), the current day's route can be feasibly shifted such that it does not exceed the maximum daily active time ($F^*_{noservice} \geq T_{noservice} - T_{max}$), where $T_{noservice}$ denotes the total active time of a truck on the current day with travel to but without service of node j , and travel should be finished before the end of the current day ($T_i + t_i + t_{ij} \leq Q_d^b$).

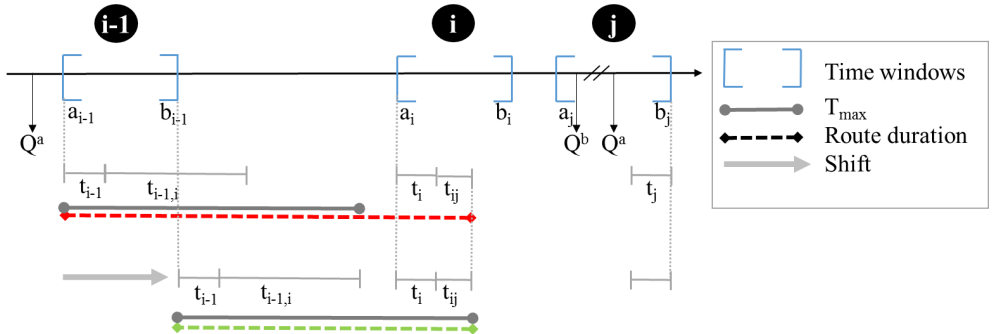


Figure 3.3: Shift the start of a route on a single day with $F^*_{noservice}$, including travel to node j on the current day, and service of j on a following day.

If a single overnight's rest is required before service at node j occurs on the following day, $\max\{ES_j; \max\{Q_d^a; T_i + t_i + t_{ij} + T_{min}\}\}$ defines the earliest start. If more than one day passes between travel to and service at node j , the service start is calculated by $\max\{Q_d^a; T_j\}$. If the service start resulting from this procedure turns out to be infeasible, the solution is infeasible.

Case 4: Travel on next day $d + 1$, service on future day $\geq d + 1$ If cases 1, 2 and 3 are not feasible, it is tested whether it would be feasible to include a night's rest after serving node i on day d . Travel to node j is performed on the next day $d + 1$ and service of node j occurs on any future day $d \geq d + 1$. If travel and service can be feasibly inserted on the same future day, after a minimum night's rest, the service start is defined by $\max\{Q_d^a + t_{ij}; ES_j\}$ if this is on the next day, or by $\max\{ES_j; \max\{Q_d^a + t_{ij}; T_i + t_i + T_{min} + t_{ij}\}\}$ any day after the next. Identical to case 3, if one or more days pass between travel and service, the service start of node j as a first activity on any day further in the future depends on the number of days that pass between travel to and service of node j .

3.3 Comparison of the sequential and integrated approach

The large neighbourhood search heuristic is used to compare results of a sequential and integrated approach. Realistic instances based on a real-life case study are used as input. An experimental design is proposed to determine an appropriate parameter setting. With these parameter values, the added value of the integrated approach is demonstrated. The experiments in this chapter are conducted on an Intel Xeon E5-2680v3 processor at 2.5 GHz of the Flemish Supercomputer Center (VSC).

3.3.1 Generated instances

A real-life case study is investigated in order to provide realistic input to the problem. In the case study, the intermodal operator offers direct rail connections between multiple terminals located within two service regions, the Benelux and Northern Italy. Trains are operated based on given periodical schedules with a cycle of one week, using both own trains and slots on trains operated by other companies. For own trains, the available loading metres and path weight restrictions limit the flows on each route, while for long-haul services managed by other parties the number of slots per length

type and their weights must be accounted for. The costs for using long-haul services are fixed once available long-haul services are known.

Long-haul services: The long-haul rail network in this case study consists of real-life long-haul services between two service regions with both forward and backward flows between the regions. This results in 84 unique long-haul services in total, of which multiple long-haul services could consist of the same connection between two terminals, with different departure times. All service characteristics, such as the capacity, departure time, travel time and arrival time of each service are based on real-life data of weekly rail services. The terminals are located at their real-life locations, four within each of the two service regions of 300 km by 250 km (Figure 3.4). One terminal per region is selected as a truck depot, and a sufficient number of trucks is assumed to be available at each depot. Trucking distances are calculated by means of the euclidean distance between two points.

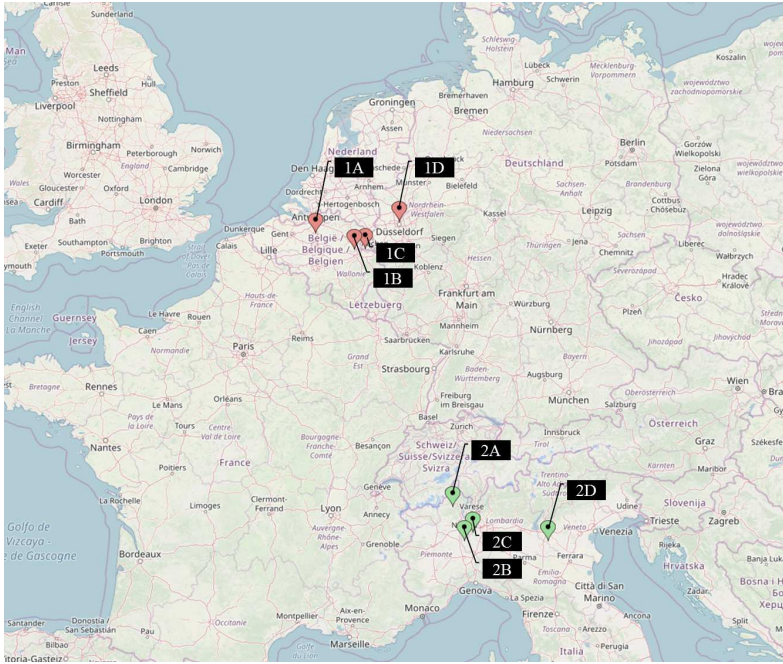
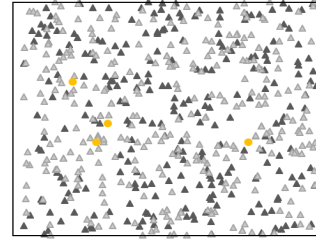
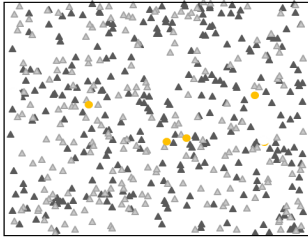


Figure 3.4: Terminals in two service regions.

Requests: Pickup and delivery locations for a request are either defined randomly or clustered within each square service region, as visualised in Figure 3.5. Clustered request locations are obtained using intermodal terminals as cluster centres based on

the procedure explained by Cordeau et al. (1997), with $\phi = 0.08$. By considering both clustered and random customer locations, the influence of customer locations can be analysed. In real-life, a scenario in between both cases occurs, and companies can relate to the scenario which is most realistic for their customer base.

a) Random customer locations



b) Clustered customer locations

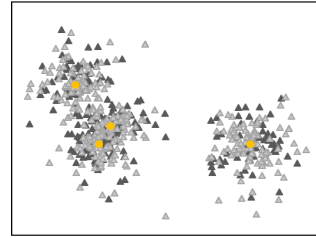
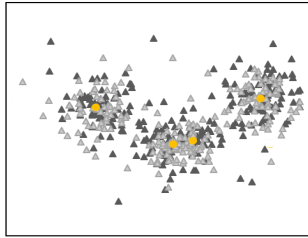


Figure 3.5: Visualisation of customer pickup and delivery locations in two regions: random or clustered around terminals.

Moreover, the demand level is varied between a relatively low and high capacity utilisation, with a maximum long-haul capacity of 470 requests for all long-haul services with departure in a single week (and arrival this week or the next). It corresponds to a scale of 1:3 in comparison with real-life problems. Additional requests are generated to account for last week's planning. Requests should be picked up and delivered at customer locations within 5 days after their release date. It is assumed that requests are released equally divided over the extended planning horizon. An equal demand of two types of container requests and trucks (30ft and 45ft) is assumed. The variation in demand characteristics results in 2x3 instance classes, with either randomly generated or clustered customer locations and the demand (in number of requests to be transported) set to either 75%, 85% or 95% of the overall long-haul service capacity. The case with 95% capacity utilisation relates the most to the circumstances a real-life intermodal operator deals with in practice.

Other input: Values for remaining input parameters are summarised in Table 3.1. Depot opening hours $[a_g; b_g], g \in N^{dep}$ are set to the length of current week's planning horizon. Service times for pickup and delivery operations at customers $t_g, g \in N^P, N^D$, and load and unload operations at terminals $t_g, g \in N^T$ are added. Service time at the depot $t_g, g \in N^{Dep}$ is assumed to be zero, as it is assumed that no operations occur at the depot. A trucking cost c_{ij} of one Euro per kilometre travelled is assumed.

While a real-life setting is used to perform the case study, the proposed solution method is able to deal with various realistic characteristics of both the long-haul service network and the local service regions. Multiple truck depots and more transport modes can be included, and the number of local service regions can be extended.

Table 3.1: Input parameters for intermodal routing.

Input parameter	Value
$[a_g; b_g], g \in N^{dep}$	[0;6*24] hours
$t_g, g \in N^P, N^D$	2 hours
$t_g, g \in N^T$	1 hour
$t_g, g \in N^{Dep}$	0 hours
T_{max}	15 hours
T_{min}	6 hours
c_{ij}	1 euro per km

3.3.2 Heuristic parameters

For the purpose of tuning the heuristic parameters, two instances are generated per instance class, resulting in 12 instances (1a-12a). An overview is provided in Table 3.2.

The goal of this research is to explore and quantify the differences between a sequential and integrated approach. Within this aim, the parameter setting should allow a fair comparison of both approaches. Therefore, for each solution approach, the best parameter setting is determined independently. This implies that the parameter settings for the sequential and the integrated approaches may differ. An experimental design is set up in order to obtain insights in the obtained results. In this way, knowledge on the impact of variations in (combinations of) parameter values as well

Table 3.2: Tuning instances for intermodal routing.

Class	Instance	Locations	Demand level
1	1a-2a	random	75%
2	3a-4a	cluster	75%
3	5a-6a	random	85%
4	7a-8a	cluster	85%
5	9a-10a	random	95%
6	11a-12a	cluster	95%

Table 3.3: LNS algorithm parameters for tuning.

Input parameter	Tested values
Number of iterations	[0;3000]
Removal of requests	5%; 10%; 15%
Acceptance threshold of new solutions <i>dev</i>	0%, 1%, 3%, 5%

as additional insights into the functioning of the heuristic can be obtained.

To determine a parameter setting, tests are executed using various parameter values for both the sequential and the integrated approach. Table 3.3 shows the tested parameter values of the large neighbourhood search heuristic.

The result is an experimental design with 12 instances as input, with three factor levels for the removal percentage when destroying part of the solution, and four factor levels for the allowed percentage deviation from the best solution in order to accept new moves. All combinations are tested. Five replications are performed for each instance, for each combination of parameter values. Each replication of a single instance is tested using both approaches.

Experimental results for the parameter tuning phase with respect to the solution quality are summarised in Figure 3.6. Computation times are visualised for both approaches in Figure 3.7 after 3000 iterations. For each instance class individually, these results are shown in more detail in Figures C.1 and C.2 in Appendix C for the sequential and integrated approach respectively. These figures indicate the solution quality and computation time in seconds for different parameter settings.

For each combination of parameter values, Figure 3.6 displays the percentage re-



Figure 3.6: LNS parameter tuning: average solution quality (reduction in total costs compared to initial solution) for the sequential (SEQ) and integrated (INT) approach after 3000 iterations;

Improvement_obj = reduction in total costs including high value for load units with long-haul service arrival in next planning cycle;

Improvement_real= reduction in real total costs.

duction in the objective function value obtained after 3000 iterations in comparison with the initial solution found by the constructive heuristic on the vertical axis, averaged over all instances. The lower on the vertical axis, the larger the cost reduction, and the higher the solution quality. The initial solution is identical in both approaches for a single instance. On the horizontal axis, the percentage of requests or tasks that is removed from the solution is indicated in the bottom. Above each removal percentage, variations in the allowed deviation for accepting new moves are reported. Results of the sequential and integrated approach are shown in grey and blue respectively. Dotted lines indicate the solution quality of the objective function, while real total costs are illustrated using solid lines. The difference between both performance measures relates to the artificial, high value for load units assigned to long-haul services with arrival in the next planning cycle. This high value is included in the objective function, but excluded from the real total cost.

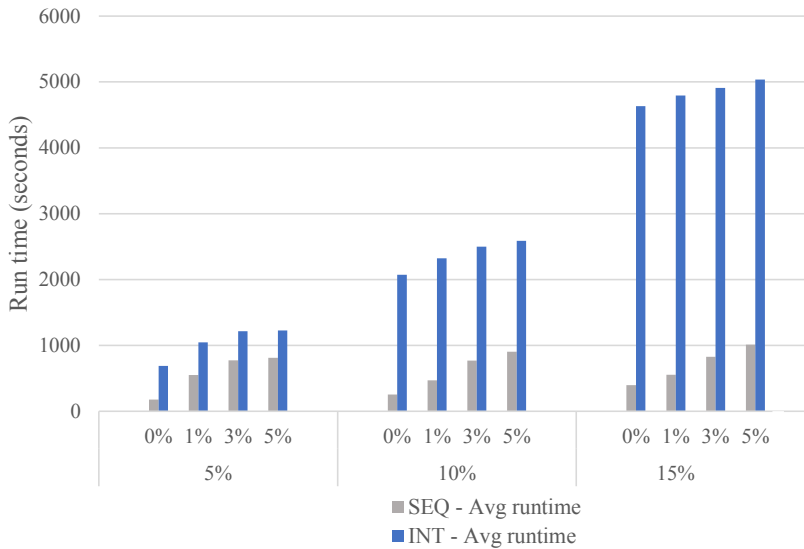


Figure 3.7: LNS parameter tuning: average computation time for the sequential (SEQ) and integrated (INT) approach.

Generally, it can be observed that, when the allowed deviation from the current best solution is too small, changes in the long-haul service assignments are not possible, and reductions in the objective value are small. On the other hand, if the allowed deviation is too big, this has a negative impact on the costs of truck routes,

and thus on the real total costs.

For each parameter setting, the integrated setting (blue lines) shows a larger percentage reduction in the objective function (dotted lines) as well as total costs (solid lines) in comparison with the sequential setting (grey lines). This is partly due to the fact that more load units can be assigned to this week's long-haul services. For the same parameter setting, total costs for drayage operations are always smaller in an integrated approach, and more drayage tasks are performed in the current planning week.

Tuning results for the sequential approach show that small improvements in the objective value are observed compared to the initial solution found by the constructive heuristic (indicated by the grey dotted lines) for all combinations of parameter settings. This demonstrates that the quality of the initial solution seems to be rather high. The real total costs (grey solid lines) consistently increase as the allowed deviation for accepting new solutions increases. This can be observed for each value of the removal percentage. A similar pattern is observed for the objective function (grey dotted lines). A slight decrease in the performance of the sequential heuristic is observed if the removal percentage is increased above 5%. This may be due to the fact that a small removal percentage already relates to a high absolute number of drayage tasks. Based on these observations, it can be concluded that total costs rise while the number of tasks to plan does not change when either the acceptance threshold or the removal percentage increases. Besides, computation times increase when removing more drayage tasks. As the graph demonstrates a good overall performance when accepting only improving solutions, and real total costs are minimised at a 5% removal rate, a 0%-acceptance threshold in combination with a removal of 5% of the drayage tasks leads on average to a minimum-cost solution for the sequential approach.

For the integrated version of the heuristic, results show larger differences in the objective and total cost values for varying parameter settings. For the objective function, it is clear that accepting improving solutions only (i.e., 0% deviation) results in the worst objective value. Whether a 1%, 3% or 5% deviation provides the best acceptance criterion, depends on the demand characteristics (see results in Appendix C). When looking at the overall average real total cost (given by the blue solid lines), it can be concluded that a 1% deviation is the best choice. Other values for the allowed deviation when accepting new solutions (3% and 5%) result in the same objective value, at a higher real total cost. With respect to the different removal percentages, a very slight reduction in the objective value is observed when increasing the removal percentage. As results are quite similar and the computation time increases substantially when increasing the removal percentage, a 5% removal is selected.

Additionally, the added value of each operator is analysed. Using the parameter settings with the best heuristic performance – indicated by the squares in Figure 3.6 –, the heuristic is run multiple times, each time excluding a single operator. Results are illustrated in Figure 3.8. Again, both the objective function (dotted lines) and the real total cost (solid lines) are displayed for both the sequential and the integrated version of the heuristic algorithm.

The base case in which all operators are included (*ALL*), is displayed on the left. Then, from left to right, results for omitted removal operators and insertion operators are illustrated. Improvements obtained in the objective function are the main criterion to determine the added value of each operator. Slight variations in the objective value can be observed, which are mainly visible in the integrated heuristic, and which are in most cases enlarged in the real total cost. For two operators, *removal_noRelated* and *insertion_noGreedy*, their omission causes a reduced objective function value, indicating a higher solution quality. The related removal is one of the new operators focused on the integrated characteristics of the problem, but decreases the solution quality. The same holds for the greedy insertion operator. As the greedy insertion operator works similar to the random order best insertion operator, it might be omitted.

Results of a paired-samples t-test of the difference in the objective function when excluding operators can be found in Table 3.4.

Table 3.4: Results of a two-way paired-samples t-test for differences in the objective function (*TC_Obj*) for excluding operators.

Excluded operator	Mean	Std. Dev.	Std. Error	95% CI		t	df	Sig.
				LB	UB			
removal_noLinked	-1912.26	6688.18	863.44	-3640.00	-184.53	-2.215	59	*0.031
removal_noRelated	1269.50	6823.14	880.86	-493.11	3032.10	1.441	59	0.155
removal_noWorst	-353.36	5851.07	755.37	-1864.85	1158.13	-0.468	59	0.642
removal_noRandom	-1151.39	9041.98	1167.31	-3487.18	1184.40	-0.986	59	0.328
insertion_no2regret	-1151.39	9041.98	1167.31	-3487.18	1184.40	-0.986	59	0.328
insertion_noGreedy	2593.27	6245.17	806.25	979.97	4206.57	3.216	59	*0.002
insertion_noRndOrdBest	-359.71	7387.33	953.70	-2268.06	1548.64	-0.377	59	0.707
omit_greedyinsert_relatedremoval	1801.87	5882.63	759.44	282.23	3321.52	2.373	59	*0.021

When looking at both scenarios with an increased objective function value, *removal_noRelated* and *insertion_noGreedy*, only the exclusion of the greedy insertion shows a significant decrease in the objective function at a 5% significance level. There-

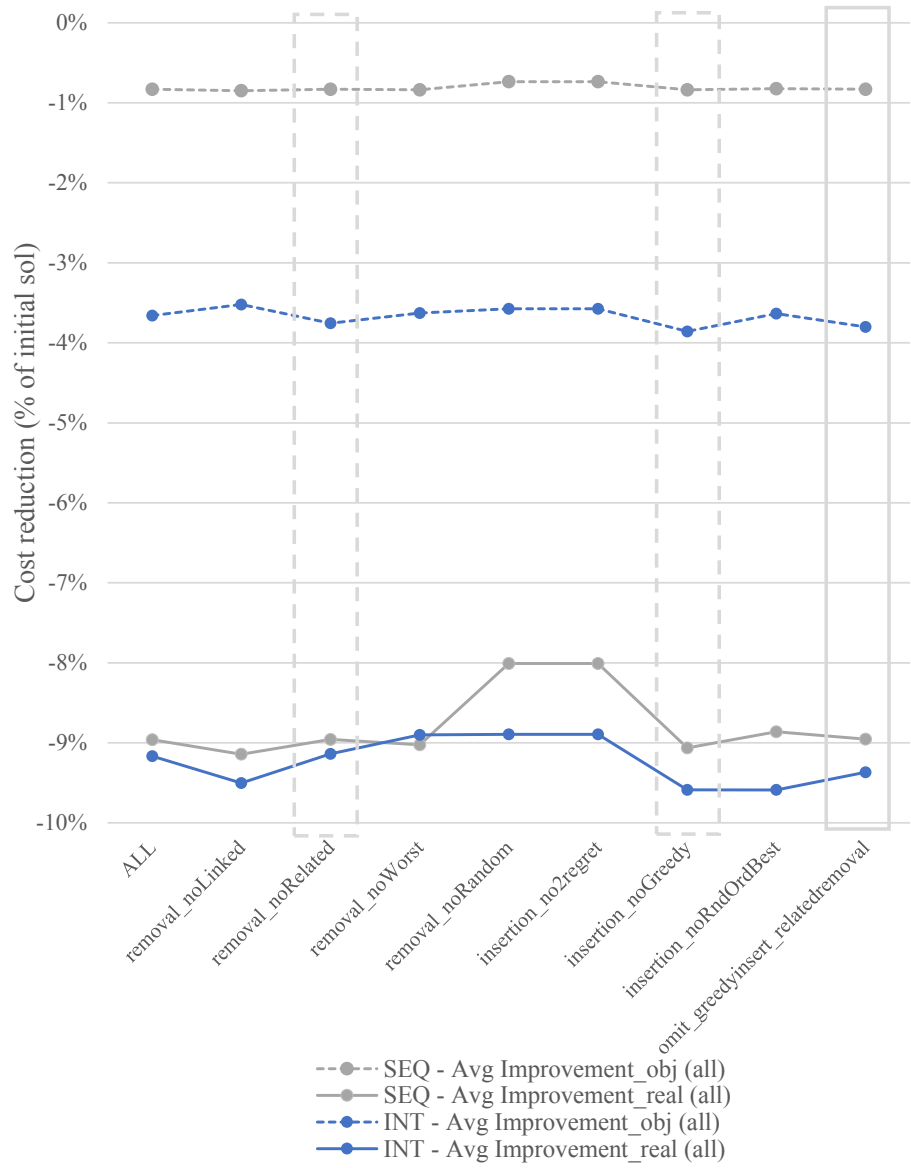


Figure 3.8: LNS parameter tuning: contribution of the operators.

fore, the case which excludes both operators (*omit_greedyinsert_relatedremoval*) is added after this first analysis of the contribution of each operator. Its performance is shown in Figure 3.8 on the right, and as a bottom line in Table 3.4. As the removal of both operators simultaneously also results in a significantly increased solution quality, the sequential and integrated heuristic without the greedy insertion and related removal operator will be used in the remainder of this chapter.

Next, the stopping criterion is analysed. Figure 3.9 shows the relationship between the solution quality and the number of executed iterations. The solution quality is measured by means of the average percentage improvement (i.e., reduction) in the objective value, and thus the artificial, high values for load units assigned to long-haul services with arrival in the next planning cycle are included. It is an average over all replications and all instances within a single instance class.

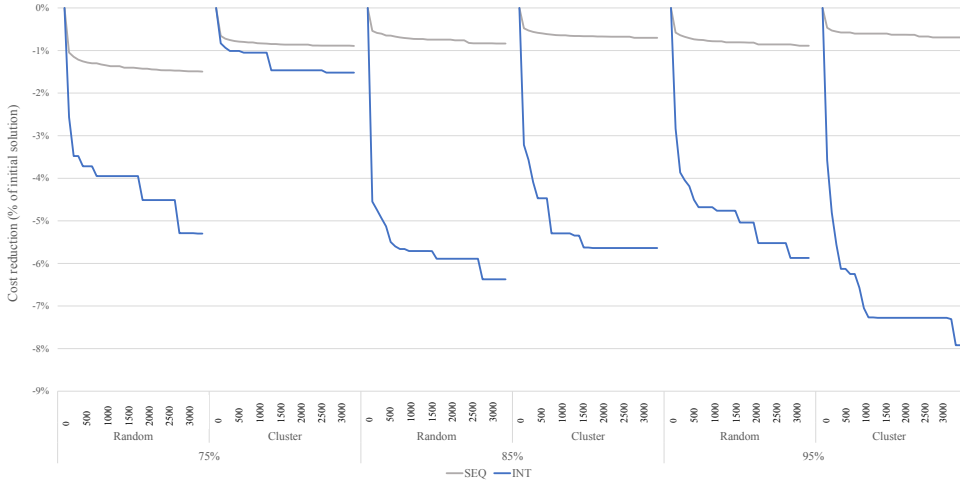


Figure 3.9: LNS parameter tuning: evolution of the solution quality with the number of iterations.

It can be observed that the largest improvements are obtained for realistic instances with a demand at 95% of the available capacity, and thus the integrated approach is most advantageous for these realistic instances. In the integrated approach, the solution quality increases consistently and flattens out around 3000 iterations. The solution quality of the sequential heuristic levels out much earlier. However, in order to maintain a fair comparison, 3000 iterations are used for both approaches. Within the aim of this study, computation time is not the most critical criterion in order to demonstrate the difference between two approaches. If the solution approach

is to be used in practice and computation time is an issue, one could easily adapt the stopping criterion to a maximum allowed computation time.

A summary of the selected parameter values for each approach can be found in Table 3.5. The greedy insertion operator and related removal operator are excluded from the final heuristic setting.

Table 3.5: LNS parameter setting for testing.

Input parameter	Sequential	Integrated
Number of iterations	3000	3000
Acceptance threshold of new solutions <i>dev</i>	0%	1%
Removal of requests	5%	5%

3.3.3 Experimental results

In order to compare the results of both approaches, experiments are performed with the defined parameter setting using new instances with identical demand characteristics. Within a single instance class, 10 instances are generated with the same combinations of input characteristics, as summarised in Table 3.6. Again, five replications are performed for each instance.

Table 3.6: Test instances for intermodal routing.

Class	Instance	Locations	Demand level
1	1-10	random	75%
2	11-20	cluster	75%
3	21-30	random	85%
4	31-40	cluster	85%
5	41-50	random	95%
6	51-60	cluster	95%

Figure 3.10 summarises the results obtained for the sequential (in grey) and integrated (in blue) intermodal routing heuristic for all demand characteristics. On the primary vertical axis on the left, the reduction in the objective value (compared to

the initial solution of the constructive heuristic) and total real cost are displayed by means of rectangles (*improvement_obj*) and triangles (*improvement_real*) respectively. The improvement in the real total costs shown in blue (*improvement_real*) are based on the real total costs incurred, i.e., the sum of long-haul service costs for all departures in the current planning week, costs of truck routes in the current planning week, and direct two-way trucking costs for delivery tasks in the future week (excluding the artificially added high value for assigning requests to long-haul services with arrival in the next week). The secondary vertical axis on the right shows the computation times in seconds, displayed by means of bars for both approaches.

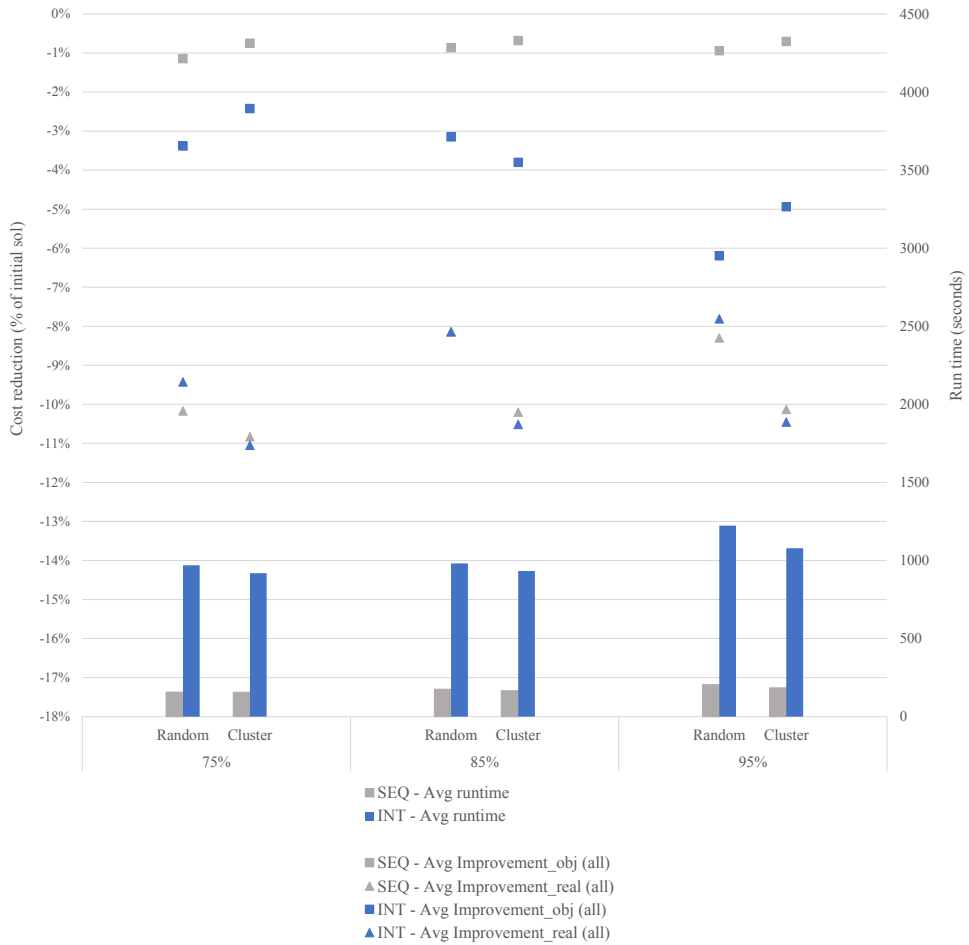


Figure 3.10: LNS results for the integrated and sequential approach.

Generally, the sequential approach leads to shorter run times, which is due to the fact that drayage tasks are considered separately. Drayage tasks in different service regions are known and planned independently, as the long-haul service is unchanged. On the other hand, with a focus on changing the route of entire requests in the integrated viewpoint, a larger computation time is required to find a feasible insertion for both the pickup and delivery drayage task when considering assigning that request to another long-haul service.

With respect to the objective function value (*Improvement_obj*), few improvements are obtained in the sequential approach, as long-haul services are already fixed and known. Cost reductions are due to a reduction in drayage costs. The objective value is clearly lower in an integrated approach, which demonstrates the advantage of the integrated approach. It is partly due to the fact that more load units can be assigned to long-haul services in the current planning week.

As illustrated by the improvement in the real total cost (*Improvement_real*), the integrated approach seems especially promising for instances with clustered demand locations. More load units can be planned in the current planning cycle at a reduced total cost. By introducing flexibility in the long-haul service, other drayage tasks can be considered which provide more opportunities to further reduce trucking costs in the service regions. A long-haul service is selected in which the drayage task is in the vicinity of other existing tasks which should be executed, hereby resulting in a reduced real total cost. For instances with random customer locations, more load units can be planned, at the cost of an increased drayage cost. The additional cost for serving more requests in the current planning week is larger for random locations. Note that costs are identical for the 85%-random-case. For a smaller demand level, the closest terminal can always be selected if enough capacity is available, and no trade-offs are required. On the contrary, with a higher capacity utilisation, a decision should be made based on the relationship between costs of truck routes and assigned long-haul services, and the closest terminal cannot always be selected. In all demand scenarios, the integrated approach postpones fewer load units to next week's long-haul services at a reduced or small additional truck cost.

Comparing both approaches, the difference in the number of load units assigned to a long-haul service with arrival and departure in the current planning cycle can be observed from Figure 3.11. Values larger than zero indicate a higher number of load units assigned to this week's long-haul services in the integrated approach compared to the sequential approach.

Clearly, as already suggested based on the general results, for all demand scenarios, the integrated approach is able to process more requests in the current planning cycle.

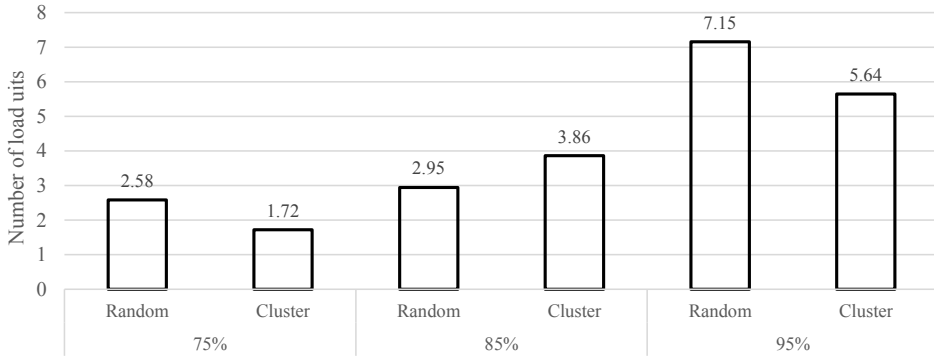


Figure 3.11: Average difference in the number of load units planned with departure and arrival in the current planning week.

These differences can be explained by the fact that no feasible insertion to long-haul services with departure and arrival in the current planning week is obtained for some requests in the construction of the initial truck routes. Whereas in a sequential viewpoint, no change to the assigned long-haul service is considered, the integrated approach is able to find an assignment to a long-haul service with arrival in the current planning week.

Figures 3.12 and 3.13 display the average absolute and percentage difference in the cost per drayage task between the sequential and the integrated approach. A positive value indicates a reduced cost for the integrated approach. The measures are calculated under the assumption of one Euro per euclidean distance unit. For long-haul services with departure in the past planning period and arrival in the current period, only costs related to the delivery drayage task are included. For long-haul services with departure and arrival in the current period, two drayage tasks – both the pickup and delivery task – relate to this week's truck routing costs. Long-haul services with departure in the current period and arrival in the next also result in costs related to both drayage tasks. While the pickup task is included in this week's truck routes, costs for the delivery task relate to its two-way direct distance.

Results show a lower trucking cost per drayage task for instances with clustered customer locations in an integrated approach compared to a sequential approach. The higher the capacity utilisation, the larger the savings. On the one hand, a cost reduction is obtained by providing more efficient truck routes. On the other hand, this effect is reinforced because of the fact that more load units are planned on a long-haul

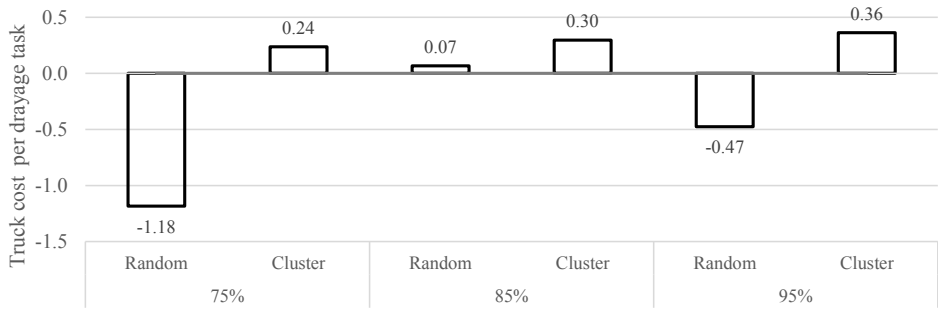


Figure 3.12: Average absolute difference in the cost per drayage task. Positive values indicate a lower cost in the integrated approach.

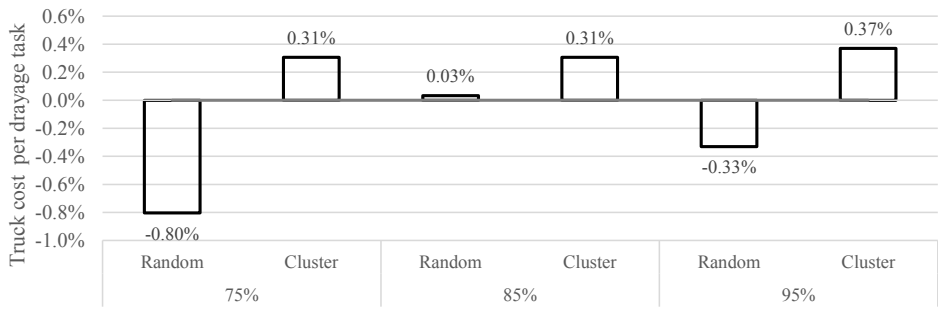


Figure 3.13: Average relative difference in the cost per drayage task. Positive values indicate a lower cost in the integrated approach.

service which departs and arrives in the current planning cycle. This further decreases costs as this higher number of load units can also be planned more efficiently.

For demand with random customer locations, it becomes harder to combine drayage tasks in existing truck routes, as distances between the locations are relatively larger. This is also reflected in the results by means of increased costs. Some customer locations might be relatively far away from some terminals, which makes the terminal selection a fixed choice. In these cases, the sequential approach is able to focus on improving truck routes with the given drayage tasks, whereas the integrated approach does not explicitly deal with optimising the truck routes for given drayage tasks. Therefore, the sequential approach might be able to include tasks which are harder to insert into truck routes in a more efficient way.

Trade-offs might exist between increasing the capacity utilisation in the current planning week and reducing trucking costs. For some load units, it could be possible to include these in this week's planning. However, in order to plan these orders, a higher trucking cost must be incurred, especially if the demand is spread randomly across the service region.

To conclude, results of the integrated approach demonstrate that the largest savings are obtained for clustered instances with demand characteristics closest to real-life cases (i.e., high capacity utilisation). In all cases, fewer load units are postponed, consistent with an improved capacity utilisation. For these instances with high demand, trucking costs per drayage tasks are reduced in the case of clustered instances, but not for random customer locations. Furthermore, in an integrated approach, more information is used to make better-informed decisions and increase this week's capacity utilisation. This might indirectly reduce transport costs, as more load units can be transported with the same long-haul capacity in a single planning cycle.

3.4 Case study: tactical service network design decisions

Using the integrated approach for real-life data, insights can be obtained into how operational transport costs are influenced by decisions concerning the service network design at the tactical level. This section serves as an exploratory study on the impact of small changes in the service network on operational costs.

In the medium term, decisions on the service network design are usually determined based on expected aggregate demand flows, without knowledge of their impact on the operational planning. Besides, the decision on the offered service network

may also be of strategic importance for reasons of competitiveness. This results in sunk costs for operating long-haul services (e.g., wagon leases, costs of purchased slots) in the short term, in addition to variable transport costs. The analyses in this section aim to demonstrate an additional way in which the integrated intermodal routing problem can be used in practice. Intermodal rail service operators can use this modelling approach to optimise the utilisation of their service network and adapt their services to their customers' expectations by changing some of their decisions in the future, for example when renegotiating transport contracts of slots on long-haul services.

In this section, the effect of two possible changes in the service network is studied in order to analyse their impact on the total operational transport costs of the system. Section 3.4.1 examines the possible advantages of dividing the overall long-haul rail capacity over a smaller number of long-haul services with a higher capacity per departure. Section 3.4.2 analyses the impact of congestion at and around terminals. The impact of these small realistic adaptations to the current service network, related to the long-haul network design and selection of long-haul rail services offered, on operational transport costs can be evaluated to provide decision support in an intermodal transport planning environment. Finally, Section 3.4.3 concludes with managerial insights on the service network design.

3.4.1 The impact of changes in the service network: removing long-haul services with small capacity

Intermodal service operators decide which services they offer in order to minimise costs and satisfy expected service levels of current or future customers at a tactical level. The selected service network relates to characteristics of the transport requests, such as the customer locations. For example, the distance between customer locations and their closest terminal should not be too large.

Various reasons may exist for service network decisions on (not) opening certain long-haul service connections. Economies of scale can be obtained if less services are operated, resulting in a lower fixed cost for opening and maintaining these services, determined at the tactical decision level. However, if more terminals are operated from, this may reduce vehicle routing costs at the operational level. Therefore, the impact of routing transport requests using a smaller number of long-haul services is evaluated. Within this context, long-haul service connections with a smaller capacity are excluded and its capacity is added to long-haul services which already own a larger share of the capacity. Two cases are studied in which each time one connection

between two terminals is removed in comparison to the base scenario. Case *a* omits the connection between terminal $1C$ and $2D$ – which is the sole connection of $2D$ –, while case *b* excludes the connection between terminal $1D$ and $2C$, which is the only existing long-haul service from terminal $1D$. The most realistic instance classes with a high capacity utilisation of 95% are used for these analyses. The capacity of the excluded long-haul service is added to the other remaining services to ensure that enough capacity is available for all requests. It is added to the remaining long-haul services in proportion to the current capacity share of each long-haul service.

Table 3.7 displays the average results for both alternative scenarios, *a* and *b*, in comparison with the base scenario which consists of the current service network of an intermodal operator. Trucking costs are directly related to the euclidean transport distances.

Table 3.7: Detailed results for the removal of a long-haul connection relative to the base case, cases *a* and *b*.

		base	a	b
Random	TC_obj	1,315,115.29	1,316,977.03	1,330,566.26
	TC_real	144,115.29	144,921.47	148,366.26
	vrpCost	108,755.37	109,088.22	111,905.62
	directTruckCost	21,381.42	21,959.91	22,486.64
	Nbr LU d1-a1	257.10	253.57	255.78
Cluster	TC_obj	1,304,305.69	1,314,235.72	1,328,823.97
	TC_real	137,305.69	144,423.22	147,712.86
	vrpCost	103,223.35	108,734.29	111,374.22
	directTruckCost	20,145.84	21,832.68	22,380.31
	Nbr LU d1-a1	256.30	253.52	255.67

Generally, a slight increase in the number of load units with long-haul service arrival in the next week is observed in all cases, which partly explains the increase in the total cost objective function (*TC_obj*) in all cases. This also results in an increased two-way direct truck cost (*directTruckCost*). Due to a lower overall frequency of service obtained by omitting a long-haul service, a small number of requests cannot be feasibly assigned to a long-haul service arriving in the current planning week. Consequently, more load units arrive at their arrival terminal in the next planning

cycle. Furthermore, the availability of fewer long-haul services results in fewer routing possibilities for each request, and higher terminal-to-customer distances, and as such, increases vehicle routing costs (*vrpCost*).

Operating a smaller number of services results in increased operational, variable transport costs. However, at the medium term, fixed costs for operating these services may be reduced. These costs are not included in the results. Case *b* involves a higher increase in total costs, as the only connection to terminal *2D* (which is not centrally located) is excluded. In case *a*, two routing alternatives are available to the terminals in Italy, while the only connection of terminal *1D* is removed. However, another terminal is closely located to terminal *1D*, which explains the smaller increase in total costs. When comparing instances with respect to their customer locations, the increase in transport costs is clearly larger for clustered customer locations. For requests with random customer locations, the number of alternative, feasible long-haul services is often higher, resulting in more options and thus flexibility for routing trucks, and lower additional trucking costs.

To conclude, multiple reasons exist for (not) opening long-haul service connections between intermodal terminals. Operating to and from multiple terminals could result in reduced trucking costs at the operational level and more flexibility in the assignment of requests to long-haul services, and as such might increase customer service levels. If the locations of these terminals are spread across the service area, operations can be performed close to both customers and terminals. However, the more terminals are used for transport, the higher the expected fixed costs related to operating long-haul services between terminals. These costs are not included in the analyses. For a given demand volume, costs of a higher long-haul service frequency with a smaller capacity per departure – which may lead to a higher service level offered to customers – should be weighed against the costs of a low departure frequency and a higher capacity per departure. The latter is likely to reduce flexibility due to the fact that fewer long-haul services are available and may result in increased trucking costs.

3.4.2 The impact of congestion around terminals

In practice, some terminals are much more congested compared to other terminals and service operators might tend to avoid such terminals. In this alternative case, the influence of congestion is analysed by means of a penalty time and cost for visiting the direct service area of a single busy terminal, in this case terminal *1A*. The true costs of congestion may consist of truck driver wage costs, costs for not being in time – such as planning time and additional costs for finding alternative routing options

– and customer dissatisfaction. Therefore, an additional cost of 30 Euro per hour of congestion is included for all trips to and from the congested terminal. As the amount of congestion is uncertain and unknown in advance, scenarios are included with the additional time varying from one hour to three hours, resulting in three congestion scenarios. The goal is to show how the integrated approach can be used on given input data to analyse the impact of congestion on planning decisions.

Experiments are conducted using a representative instance with respect to realistic input data. In real-life, customer locations are somewhere between random and clustered. Therefore, both classes are presented, such that practitioners can select the case which fits their customer base the most. In order to explore the full impact, experiments are performed using an instance with 85% utilisation to ensure that enough capacity is available at other long-haul connections. In this way, flexibility is allowed in order to show possible routing scenarios, which can be used as the basis for renegotiating transport contracts. These analyses may show at which connections capacity should be increased in order to avoid congested terminals.

Figure 3.14 displays the obtained results for an instance with random customer locations. Possibly impacted terminals with a direct connection to terminals 1A are terminal 2A and 2C. The throughput share of each terminal within each service region is indicated, calculated based on the sum of inbound and outbound requests for each terminal. The base case represents the case without congestion.

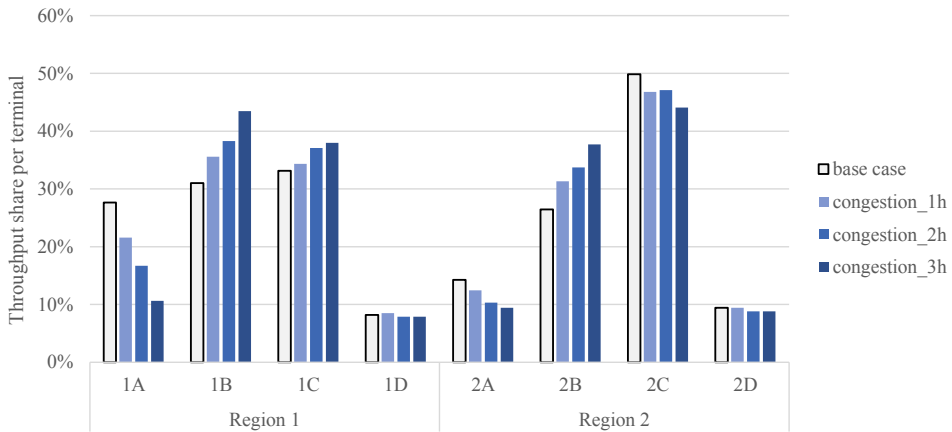


Figure 3.14: Difference in terminal throughput share due to congestion for random customer locations.

A structural decrease can be observed in the throughput of terminal 1A. The terminal usage decreases further when congestion time increases. Due to the random customer locations, an assignment to terminal 1A might be required for some requests in order to be feasibly assigned and arrive in time at the customer destination location. Impacted terminals with a connection to terminal 1A in service region 2 are terminals 2A and 2C.

In the Benelux region, terminals 1B and 1C are alternative routing options. They are centrally located in the service region and present feasible alternatives for many requests. The throughput of terminal 1B clearly increases the most as it is directly connected to one of the impacted terminals in the second region. If requests originally assigned to terminal 1A should be rerouted, terminals 1B and 1C are the closest ones with similar connections. They pose plausible alternatives for feasibly routing load units through the intermodal service network. In northern Italy, the relative throughput clearly increases in terminal 2B, which is also centrally located with multiple routing options. While terminal 2C is also characterised by a centralised location, the throughput decreases due to its connection with the congested terminal. This effect outweighs the possible advantages of its central location. Furthermore, due to its direct link with the congested terminal, the share of terminal 2A also slightly decreases. A steady throughput is observed for both terminals 1D and 2D. They are not centrally located, they are less likely to present a feasible alternative for requests which were initially transported via terminal 1A and both terminals each only have a long-haul service connection to a single other terminal.

Figure 3.15 visualises the throughput changes due to congestion for clustered customer locations.

Results for clustered customer locations show a similar pattern with respect to the impact of a single congested terminal. In the Benelux, the relative increase in the usage of terminal 1B is larger. Due to the clustered request customer locations and the centralised location, this terminal might provide an interesting, feasible alternative for more requests impacted due to congestion. For terminal 1C, the impact is smaller, as in the base case, its share is already relatively high. In the second service region, the decrease in the importance of terminal 2C is even larger.

The above analysis provides insights in alternative routing decisions in order to avoid congested terminals. An additional cost was included in order to account for the cost of congestion in order to show how routing decisions might change.

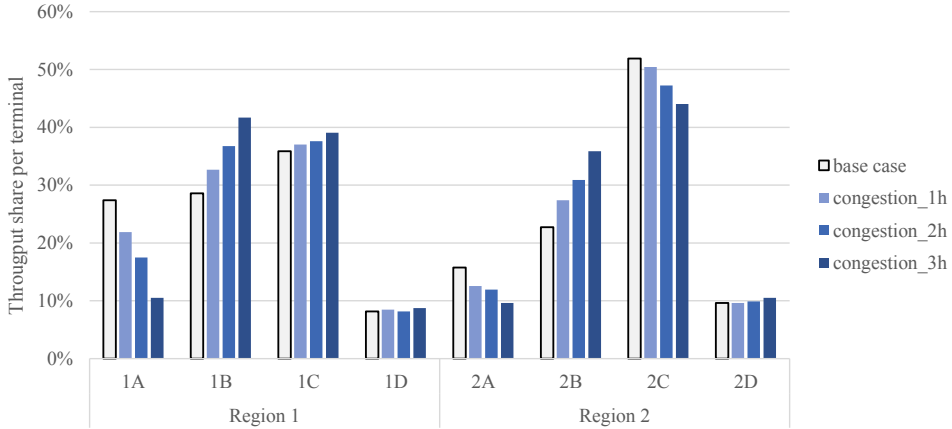


Figure 3.15: Difference in terminal throughput share due to congestion for clustered customer locations.

3.4.3 Managerial insights

Intermodal operators offer a service network through which transport requests are routed for the long-haul transport. Tactical decisions on the design of the service network are made in the long term using expectations about future demand, and transport contracts for operating long-haul services or purchased slots on long-haul services of other operators are fixed for multiple months. At the operational level, transport requests are routed through this given service network offered by the intermodal operator.

Usually, transport requests are assigned to the long-haul service with the smallest total terminal-to-customer distance. However, this might not always prove to be the best decision as not all information is accounted for. Vehicle routing costs are ignored, and initial assignments might not always be feasible, for example due to customer-depot distances or trucking and driving constraints. By taking an integrated perspective, infeasible long-haul routing decisions based on incomplete information can be avoided. By including information on feasible vehicle and long-haul routes and accounting for capacity restrictions on long-haul services and a heterogeneous truck and container fleet, the integrated approach aims at providing better-informed decisions in comparison with the traditional, sequential method. For example, when setting pricing decisions for customers based on more complete information, a more accurate price might be charged corresponding closer to real costs incurred. Further-

more, ideally, the service network should evolve in time in order to adapt to expected customer demand characteristics. The proposed integrated intermodal routing problem allows to evaluate the impact of variations in expected customer demand. Before accepting new customers, the integrated method could analyse the possible impact on transport costs and routing decisions. When renegotiating transport contracts on purchased slots, such evaluation can be used to propose adaptations to the service network accordingly.

3.5 Conclusions and future research

In this chapter, potential savings and advantages of an integrated approach for intermodal routing are examined. The proposed integrated intermodal routing heuristic encourages a seamless integration of decisions on truck routing and long-haul intermodal service selection. A reduction of drayage costs and an improvement of the capacity utilisation stimulate the use of intermodal transport. While this study focuses on a two-region truck-rail network, the solution approach can deal with other transport modes and multiple regions.

It is demonstrated that cost savings can be obtained by the integrated approach in which decisions on truck routing and long-haul container routing are made simultaneously. Although cost savings are not obtained for instances with customer locations placed randomly in the service regions, more load units can be transported, indicating a better use of the available long-haul capacity. Besides, for clustered instances, savings in trucking costs are obtained. Moreover, results indicate that a trade-off exists between increasing the capacity utilisation in the current planning week and reducing vehicle routing costs. An additional intensification phase in the large neighbourhood search heuristic for integrated intermodal routing could further improve the transport cost savings. In that case, the obtained results in this chapter could be a lower bound on the possible advantages of the integrated approach. This can be studied in future research.

An integrated perspective allows to make better-informed decisions, hereby avoiding infeasible solutions. Moreover, computation times amount to a maximum of 1000 seconds, which is certainly acceptable for the design of the planning of an entire week.

Chapters 2 and 3 developed an integrated viewpoint on local drayage decisions and the selection of long-haul service connections. The approach aims to foster a synchromodal viewpoint on the transport chains by integrating decisions on different transport modes. The approach can be adapted to a day-to-day or even real-time

planning horizon. Future work may focus on a dynamic version of the problem. Whereas the current approach serves the purpose of providing insights in the way the current service network design impacts total transport costs, a more operational approach considering planning on a daily basis could support the day-to-day planning process even further. In line with a synchromodal view, such an approach could also be able to deal with disruptions in real-time and shift transport modes and routes whenever necessary.

In the current research setting, a single intermodal operator offers multiple services between two service regions of a given corridor. If more services and flows can be considered for routing a request for the long-haul, this may yield further cost reductions. Within this regard, one interesting future research direction consists of taking a cooperative approach. If different decision makers would be willing to collaborate, this might further improve the efficiency of the transport system as a whole.

In this chapter, a large neighbourhood search heuristic was proposed to solve the integrated intermodal routing problem. For both the sequential and integrated solution approach, the same initial solution is constructed based on a sequential version of the problem. Alternatively, the final solution of the sequential approach could be used as initial solution to the integrated approach. Future research may investigate the influence of this adaptation on the heuristic performance of the integrated approach. Besides, the integrated problem can be decomposed into the two problems considered separately in the sequential approach. The weekly vehicle routing problems might also be decomposed into daily routing subproblems. Moreover, the number of variables increases strongly when the size of the problem instances increases. Because of these problem characteristics, the use of column generation to solve the problem presents a valuable opportunity for future research.

Finally, at the tactical level, the service network and its characteristics are variable. In order to cope with occasional capacity shortages, overflow can be assigned to one-time purchased slots at an additional cost. By optimally managing the own services and services offered by other operators, the total costs of operating this transport system could be minimised even further.

Part III

Train load planning

— *Introduction to Part III: Train load planning* —

Part II (Chapters 2 and 3) focused on intermodal routing. In this problem context, load units are assigned to services on a weekly basis, accounting for restrictions on the total service network capacity. In Part III (Chapters 4 to 6) the train load planning problem is studied. Compared to the intermodal routing problem, train load planning occurs on a daily basis. Load units are assigned to specific locations on intermodal trains using detailed, short-term information. Furthermore, some flexibility is allowed in the decision process. For example, one could decide on which load units will be postponed and which should certainly be assigned to ensure an optimal capacity utilisation based on real-time information and satisfy axle weight constraints of wagons. Trucks carrying specific load units might not arrive in time, urgent transport orders arrive into the system and the impact of cancellations of other long-haul transport services or information about future transport orders should be included.

Factors influencing the train load planning process: literature review and classification

4.1 Introduction

By significantly raising the efficiency and capacity of rail transport over long distances, intermodal rail freight transport can be encouraged (Boysen, 2014). One factor determining the railway system capacity concerns the load capacity per train (Boysen, 2012). However, on-train capacity utilisation has not received much research attention in comparison to route and network capacity (Woodburn, 2015), discussed in Chapters 2 and 3. For an overview of literature on this topic, the reader is referred to Crainic (2000) and Chapter 2. Improved capacity utilisation per train can increase rail freight volumes, and consequently the overall railway system capacity utilisation, without adding expensive network capacity. Therefore, it is important to determine which load units will be loaded, and on which location on the train, to maximise the train's loading degree and minimise costs per load unit. This is the train load planning problem.

This chapter presents a literature classification and review of train load planning problems (Figure 4.1). The train load planning problem is situated at the operational decision level and is related to two types of decision makers, the intermodal operator

This chapter is an updated version of the following paper:

Heggen, H., Braekers, K., Caris, A., 2016. Optimizing Train Load Planning: Review and Decision Support for Train Planners. *Lecture Notes in Computer Science*, 9855, pp. 193-208.

and the terminal operator. The intermodal operator organises transport and often deals with a large variety of wagon and load unit types. This increases the complexity of train load planning (Macharis and Bontekoning, 2004). Given a number of outbound load units, available wagons and their corresponding characteristics, the intermodal operator's train planner typically has to determine which load units will be loaded on which train, and establish a feasible train load plan. The second decision maker involved is the terminal operator, who is responsible for transshipment from one mode to another. Given the load units to be loaded as provided by the intermodal operator, he focuses on the efficient allocation of resources at the terminal, for example to minimise handling costs. The transshipment process increases the chain lead time and total transport cost. Hence, it needs to be executed fast and efficiently (Macharis and Bontekoning, 2004). Transshipments can be made by means of a gantry crane, which is able to transport load units very fast, but is very capital intensive. Alternatively, mobile instruments such as reach stackers can be used. They are slower and require driving lanes, but they are cheaper and more flexible (Reis et al., 2013). The train load plan serves as an input for the planning of both types of handling equipment. Given the assignment of load units to a location on an intermodal train, the sequence of load unit moves per crane or mobile instrument can be decided (Boysen et al., 2010). The areas cranes operate in can be either fixed or free. In the latter case cranes are allowed to move freely along the yard. When more than one crane is available, interference should be avoided.

Intermodal transport companies often perform the assignment of load units to a specific location on an intermodal train manually. Automation of this task can assist train planners in their decision-making process. It may result in large cost savings by increasing the number of load units assigned and consequently, decreasing the cost per load unit, and by avoiding costs for the removal of a wagon and the transfer of its cargo due to excess axle loads at certain measure points. Moreover, substantial time savings can be gained in the planning process. Finally, it can have a significant influence on the time and energy spent on handling load units (Corry and Kozan, 2008).

A lot of research has already been conducted on crane and storage planning in container terminals. An overview of literature on container terminals is provided by Steenken et al. (2004), and is updated by Stahlbock and Voß (2008). Further, Boysen et al. (2013) provide an overview of processes at railway yards. Train loading is part of the land-side operations in a container terminal, but has not been discussed extensively so far (Corry and Kozan, 2008). Whereas most literature focuses on a more general view on capacity at the level of network capacity, the train load

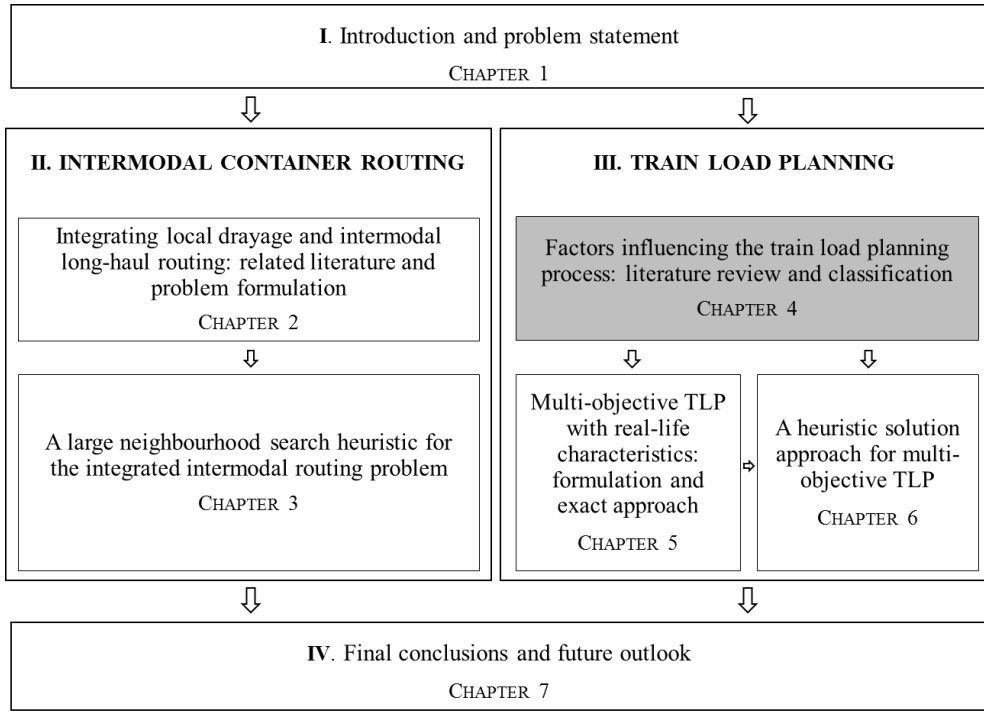


Figure 4.1: Thesis summary - Chapter 4.

planning problem considers a more detailed definition of train capacity utilisation. The aim of train load planning is to find an assignment of load units to the available locations or slots on a train. The objective can be based on either capacity utilisation or handling operations at the intermodal terminal. At the same time, axle payloads, wagon weights and total train weights are restricted, a balance should exist between the payloads on the adjacent axles of each wagon, and possibly other operational constraints are applicable.

This chapter is structured as follows. Section 4.2 presents literature on train load planning. Current research is categorised using three groups of problem characteristics: objectives, constraints and the planning environment in which train load planning is carried out. Other problems closely related to train load planning are also discussed. Conclusions and research opportunities are presented in Section 4.3.

4.2 Classification of train load planning problems

A major contribution to the development of the train loading problem has been provided by Corry and Kozan (2008), who developed a realistic model which can be extended to various specific environments. Furthermore, Bruns and Knust (2012) are the first - and currently still the only ones - to adopt continuous weight restrictions in a train load planning problem. Both works (Bruns and Knust, 2012; Corry and Kozan, 2008) laid the groundwork for realistic train load planning problems. Recently, the optimisation of train load planning has been integrated with optimisation of other operational decisions in an intermodal seaport terminal. This integration is first introduced by Ambrosino et al. (2011), who simultaneously optimise crane and storage planning. The overview in this section discusses both types of container loading problems, namely problems focusing on pure train load planning and problems integrating train load planning with crane and storage planning.

The composition and revision of train load plans is influenced by three groups of factors. Models can be distinguished by their defined objectives (Section 4.2.1). Furthermore, train load planning is affected by characteristics inherent to the train components and operational constraints related to the specific environment in which train loading takes place (Section 4.2.2). Section 4.2.3 introduces the influence of the planning environment on the load plan.

4.2.1 Objectives

Objectives of train load planning problems in current literature can be divided into four categories. An overview is presented in Table 4.1.

A first measure for the intermodal operator to define its performance is the train utilisation or loading degree. It can be expressed in number of load units, total weight or length of the load units. The urgency for load units to arrive at their destination can also be accounted for. This objective can be implemented in two ways. On the one hand, the number of load units assigned can be maximised, where a large number of load units serve as input. This is sometimes achieved by penalising containers left behind or adding a profit for assigning a load unit. On the other hand, a fixed number of load units can be used as input, which all must fit on the train using the least possible number of wagons, or minimising train length.

In a second category of objectives, the terminal operator aims at minimising costs of handling operations at the terminal during execution of the load plan. Corry and Kozan (2008) divide handling into three components: double handling, carry travel

Table 4.1: Objectives for the train load planning problem.

	Train load planning										Integration with crane & storage planning						
	Feo and González-Velarde (1995)	Corry and Kozan (2006a)	Corry and Kozan (2004)	Corry and Kozan (2006b)	Corry and Kozan (2008)	Aggoun et al. (2011)	Bruns and Knust (2012)	Bruns et al. (2014)	Upadhyay et al. (2017)	Mantovani et al. (2018)	Ambrosino et al. (2011)	Ambrosino et al. (2013)	Ambrosino and Siri (2014)	Anghinolfi and Paolucci (2014)	Anghinolfi et al. (2014)	Ambrosino and Siri (2015)	Ambrosino et al. (2016)
Utilisation																	
Max nbr of load units							•	•	•	•	•	•	•	•	•	•	•
Max nbr of urgent load units									•		•	•	•	•	•	•	•
Max load unit length							•	•							•		
Max load unit weight							•	•							•		
Min nbr of wagons	•			•	•	•			•								
Handling costs																	
Min changes in pin position			•		•		•	•							•		•
Min transport cost		•	•		•		•	•						•			
Min double handling		•	•		•												
Opt weight distribution			•	•	•					•							
Unproductive moves																	
Min rehandles in storage area											•	•	•	•	•	•	•
Min backward empty moves												•	•			•	

and pin changes. Double handling occurs when a load unit is not directly transferred from the truck to a wagon, which means that excess handling occurs (Corry and Kozan, 2006a). It is only considered in cases in which loading and unloading are executed simultaneously. In these cases uncertainty exists about the occurrence of double handling for each load unit, because truck arrivals and move sequences of handling material are uncertain. As inbound load units are unloaded, slots become vacant. Consequently, the probability of double handling for outbound load units which are assigned to these slots but did not arrive yet by truck becomes zero. Another definition of double handling may be unloading a loaded load unit. Carry travel or handling equipment travel corresponds to the transport cost from the storage position to the assigned location on the train (Bruns and Knust, 2012). Uncertainty about this performance indicator exists if not all load units arrive before loading starts. Furthermore, each load unit is fixed on a wagon by means of four pins, which must be aligned with castings located at certain points on load units (Corry and Kozan, 2008). Changing the pin positions of a wagon is labour-intensive and occurs when the pin positions from the inbound wagons do not satisfy the pins needed to lock an outbound load unit on the wagon. In that case, set-up costs occur.

A third performance indicator is related to the weight distribution of the train (Corry and Kozan, 2006a). A good weight distribution may reduce the wear of the brakes. The weight is optimally distributed if the distance from the centre-of-mass to the front of the train is minimised. This objective can be included as a second step (Upadhyay et al., 2017), after the load units to be loaded are known.

The fourth category aims at minimising unproductive movements as a result of using multi-level container stacks at the terminal. Rehandles or reshuffles are unproductive movements required in multilevel stacking areas when the load unit to be picked up is not on top of the stack. This can be reduced by performing non-sequential or backward empty crane movements when loading the train, which is a second type of unproductive operations (Ambrosino and Siri, 2015).

Clearly several of these objectives may be relevant in a specific situation. The type of operator which has the responsibility of constructing and revising the load plan (i.e., the terminal operator or the intermodal operator) influences the considered objectives. Handling cost minimisation is the most common objective in pure train load planning problems. Current models in literature typically use a weighted-sum approach to aggregate the objectives in a single objective function. Anghinolfi et al. (2014) include multiple objectives, but also use a single objective function with a multiplication of the commercial value - representing the load unit urgency - and the weight and length of the load units assigned to a location on the train. To our knowledge, only one

paper (Ambrosino et al., 2016) applies multi-objective optimisation to the train load planning problem. Furthermore, the minimisation of unproductive moves in multi-level stacking areas is only considered in combination with problems which integrate train load planning optimisation with crane and storage planning. No further patterns seem to exist with respect to combinations of objectives. However, optimisation of the weight distribution has not been considered in recent problems.

4.2.2 Constraints

Train component characteristics restrict loading possibilities. As defined by Corry and Kozan (2006a), a load plan provides an assignment of load units to slots on a train, where each slot equals one load unit length. Load units can be divided into three types: trailers, containers and swap bodies (Bruns and Knust, 2012). Each load unit is further characterised by its length, commercial value or urgency, and loaded weight.

Trains have a limited weight and length. They consist of a number of wagons of a specific length and tare wagon weight. Each wagon can carry a limited weight, and has a limited draw gear capacity, which means that the mass that is allowed to trail behind the wagon is limited (Corry and Kozan, 2006b). Each wagon is restricted to a finite number of possible loading patterns or configurations (Corry and Kozan, 2008). They specify how many and which types of load units can be loaded on a wagon. The maximum weights per slot and wagon are limited based on the allowed axle loads. Every change in the configuration implies a set-up cost for changing pin positions (Bruns and Knust, 2012). Bruns and Knust (2012) suggest two options to limit axle loads. The first option models weight distributions discretely. It is based on loading pattern descriptions provided by wagon manufacturers, as is most common in practice. For the second option, axle loads are calculated based on continuous weight distributions instead of using a set of allowed patterns. The authors also mention two ways of categorising load units into types. The first approach considers length-types, where each length-type corresponds to exactly one load unit length. The second approach is based on fixation-types. Load units with the same fixation-type may have various lengths, as long as they fit the same wagon pin configuration. This implies variable overhangs for load units belonging to one fixation type. For example, both a 20 and a 24-foot load unit could be attached to the same pin fixation. Finally, some papers mention the possibility of double stacking, in which more than one load unit can be stacked on a single slot (Bruns and Knust, 2012; Corry and Kozan, 2008). In that case, the train height may be limited for some routes due to the fact that

trains pass low bridges on their route (Corry and Kozan, 2006b). Additional safety constraints relate to the vertical centre-of gravity and stacking rules. A detailed discussion on the loading of double-stack trains is provided by Upadhyay et al. (2017) and Mantovani et al. (2018). However, in Europe this is not allowed due to the presence of low bridges and tunnels and electrical wires above the rails.

Figure 4.2 visualises a number of possible configurations or loading patterns. At the top, the side view of a wagon is shown, indicating the length of the wagon, the length of the loading space and the distance in between bogie attachments in which axles are fixed. In the dashed circles trailer wheels can be attached. In the front of the wagon, a coupling plate is put in place to fix trailers. Next, a top view is shown, presenting the possible pin positions (*). Two boxes ($K1$ and $K2$) show two possible configurations. An arrow ends at a configuration when the pins are used in this configuration. Each configuration defines a maximum number of slots. For example, configuration $K1$ consists of two slots, while $K2$ consists of a single slot. Furthermore, possible lengths per slot are identified (e.g., 20–24 feet), as well as discrete weight restrictions per slot (e.g., a combination of a weight of 0 ton for slot one and 36 ton for slot two).

Depending on the operational environment, additional constraints must be satisfied. Certain types of dangerous goods must be separated by a minimum distance, or at least they cannot be located next to each other on the same wagon. A train can have more than one destination. In that case, wagons must be grouped per destination. Load units carrying refrigerated goods must be attached to wagons with power supply (Corry and Kozan, 2006b). Finally, incompatibilities between certain load units and wagons may exist (Corry and Kozan, 2004).

Corry and Kozan (2004) mention that some factors might conflict with the minimisation of double handling, such as the weight distribution of the train, the wagon axle load, the separation of dangerous goods and the train height limit, as well as aspects related to container handling, such as the travel of equipment and changes in the pin configuration. The authors stress the fact that it may be beneficial to incur carry travel if it results in improvements in double handling or weight distribution (Corry and Kozan, 2006a).

An overview of all train component characteristics and other operational constraints included in current literature is provided in Table 4.2.

Factors influencing the load plan which are related to the train components are now well established. However, pure train load planning models do not account for the urgency with which load units must be shipped, whereas models integrating train load planning with crane and storage planning do not add specific operational constraints.

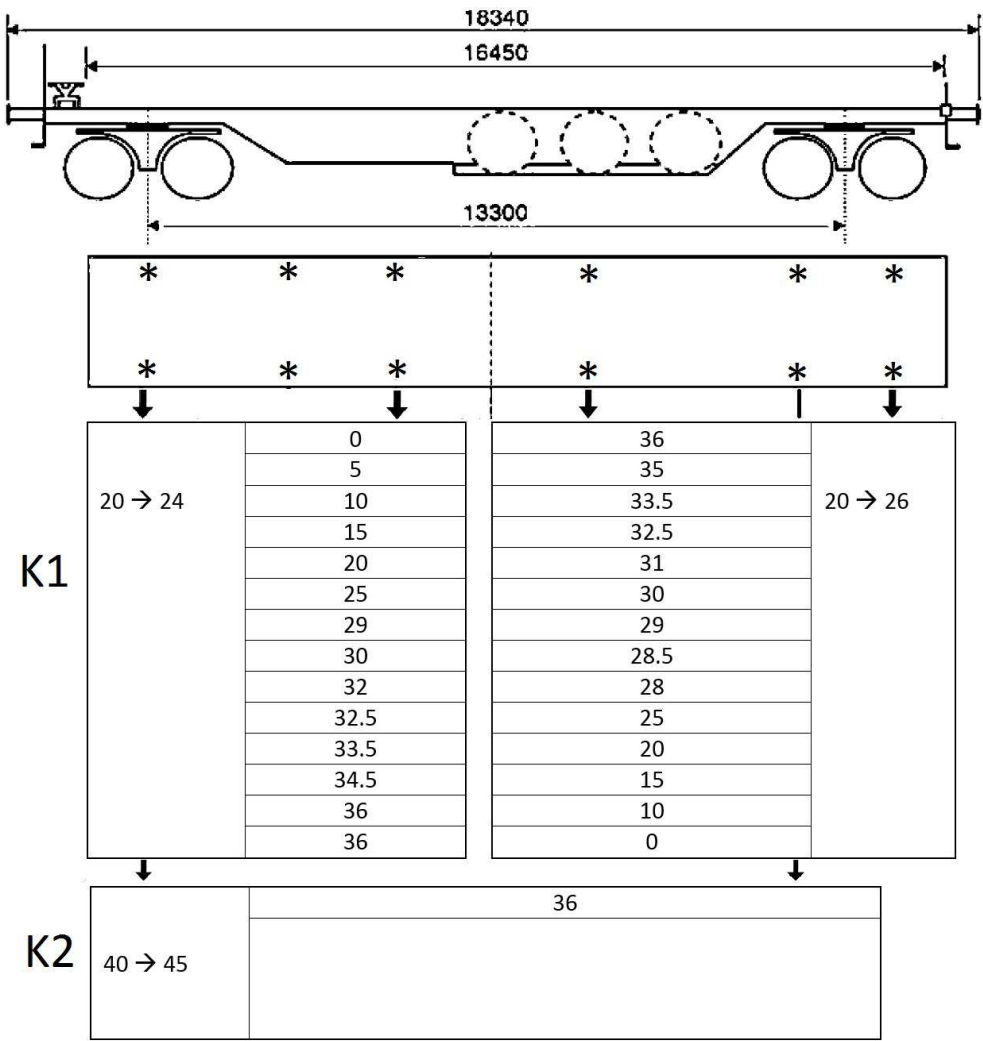


Figure 4.2: Possible configurations (adapted from Bruns and Knust (2012)).

Table 4.2: Constraints for the train load planning problem.

	Train load planning										Integration with crane & storage planning						
	Feo and González-Velarde (1995)	Corry and Kozan (2006a)	Corry and Kozan (2004)	Corry and Kozan (2006b)	Corry and Kozan (2008)	Aggoun et al. (2011)	Bruus and Knust (2012)	Bruus et al. (2014)	Upadhyay et al. (2017)	Mantovani et al. (2018)	Ambrosino et al. (2011)	Ambrosino et al. (2013)	Ambrosino and Siri (2014)	Anghinolfi and Paolucci (2014)	Anghinolfi et al. (2014)	Ambrosino and Siri (2015)	Ambrosino et al. (2016)
<i>Train component characteristics</i>																	
Load unit																	
Type						•	•	•	•								
Length	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
Commercial value/urgency									•		•	•	•	•	•	•	•
Loaded weight		•	•			•	•	•	•	•	•	•	•	•	•	•	•
Wagon																	
Configurations																	
<i>Load unit length types per configuration</i>	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
<i>Pin positions per configuration</i>		•		•			•	•	•	•					•		•
<i>Weight limit per slot</i>						•	•	•			•	•	•	•	•	•	•
Draw gear capacity			•														
Train																	
Weight limit						•	•	•			•	•	•	•	•	•	•
Length limit						•											
Height limit				•					•	•							
Double stacking						•			•	•							
<i>Other operational constraints</i>																	
Dangerous goods separation		(•)	•	•	•					•							
Destination grouping			•														
Incompatibility wagons & load units		(•)				•				•							
Refrigerated containers with power supply			•							•							

(•) = Formulated as a restriction, but not used to solve the model

Finally, the draw gear capacity is only considered in a single article (Corry and Kozan, 2006b).

4.2.3 Planning environment

In practice, train load planning starts with the first booking of a load unit on a train. More load units are assigned to the train until no feasible load plan can be established by adding another load unit, which marks the last booking and the end of the booking process. When the load plan is finalised, it is communicated to the terminal operator and loading starts. In the meantime, load units arrive at the terminal until closing time. The process ends just before train departure, when all load units are loaded onto the train.

We categorise train load planning into three types of planning environments, as presented in Table 4.3.

Table 4.3: Static and dynamic train load planning problems.

Authors	Static	Dynamic	
		Train planning operations:	Terminal operations:
		Uncertainty about information on load units	Uncertainty related to loading operations
Feo and González-Velarde (1995)	•		
Corry and Kozan (2006a)	•		•
Corry and Kozan (2004)			•
Corry and Kozan (2006b)	•		
Corry and Kozan (2008)	•		•
Aggoun et al. (2011)		(•)	
Bruns and Knust (2012)	•		
Bruns et al. (2014)		•	
Upadhyay et al. (2017)	•		
Mantovani et al. (2018)	•		

(•) = Changes are performed by planners and imposed as constraints in the formulation

Static plans are used when terminals receive all load units before the loading process starts, the train is initially empty and all information is known with certainty. A static load plan can also be used as a guide in the booking process each time a load unit is booked to check whether enough capacity is available (Corry and Kozan,

2006b). Problems integrating train load planning with crane and storage planning are all static, as is more common in seaport terminals, with large numbers of available load units, and are not shown in the overview.

Caris et al. (2013) mention that intermodal transport has grown into a dynamic research field. This is even more important in the context of synchronomodal transport, which allows for flexible mode selection based on the operational circumstances and customer requirements, and which is currently receiving increased research attention (SteadieSeifi et al., 2014). Uncertain events can occur in the last hours of the planning process and between the moment the load plan is sent to the terminal operator and train departure. They are related to the train planning environment and make the load planning a dynamic process. An initial load plan is established before loading starts, after which revisions can occur. These revisions may be necessary whenever certain events arise that change the suitability of the current load plan (Corry and Kozan, 2006a). More urgent transport orders can emerge, planned load units might not arrive in time at the terminal and information on the type of load unit may be updated. Moreover, the quality of the input data differs depending on the moment of planning (Bruns et al., 2014). Real weights can differ from the initial data, wagons can be damaged and thus cannot be used temporarily and overhangs can change in comparison with available data, as considered by Bruns et al. (2014). Aggoun et al. (2011) consider the case in which train planners perform changes to the load units to be assigned. The changes are imposed as constraints to the problem, after which a new load plan is generated. In summary, input data adaptations and the occurrence of unexpected events trigger revisions to the load plan, complicating the load planning process.

Another type of dynamic planning occurs when unloading and loading are performed simultaneously, the train is initially not empty and load units arrive while loading (Corry and Kozan, 2004). In that case, it is assumed that all load units are booked and known in advance. It is related to the environment in which terminal operations are performed to execute the load plan. Arrival times of trucks at the terminal are random (within a predefined time window) and thus uncertain. It is impossible to pre-plan train loading and still minimise handling time, because uncertainty exists about the occurrence of double handling. Outbound load units may arrive faster than inbound load units are unloaded. Consequently, a number of outbound load units must be stored in a temporary storage area. Hence, when a truck arrives to collect or deliver a load unit or a load unit is being unloaded from a train (in which case free space becomes available), the load plan is revised.

4.2.4 Related problems

This section provides an overview of literature on related problems. The main focus of these papers is mostly not to establish a train load planning, but the discussed problems are closely associated.

Bostel and Dejax (1998) describe the problem of allocating load units to a location on an arriving train and their location on a departure train to minimise the number of load unit moves as well as the use and sizing of handling equipment in a rapid highly standardised rail-rail transshipment yard. The authors assume that only one type of load unit exists. In a similar context Souffriau et al. (2009) assign containers to slots on trains with different destinations such that the lateral cost of transshipment operations is minimised, without considering detailed load planning or due date restrictions.

Powell and Carvalho (1998) aim to optimally employ the equipment and flow of flatcars, and consequently to manage the size of the fleet. A decision is made on the loaded or empty repositioning of railroad owned equipment by combining known customer requests with demand forecasts, hereby introducing network level information to improve the decisions at a local level to optimise the entire network. However, no decision is made on the assignment of load units to specific wagons or locations on the train. Only the types of wagons satisfying a certain demand and loading pattern for the wagon are determined. The problem is divided into two dynamic assignment models formulated as a logistics queueing network. The first model assigns load units to customer requests, while the other assigns wagons to load units.

Two papers optimise the fuel efficiency of load plans to reduce emissions, for which the load plan serves as an input. It is first introduced by Lai et al. (2008a). The greater the gaps between load units - where the weight associated with a gap between two units decreases as it is farther from the head end -, the less fuel-efficient the train. Hence, two trains with the same slot utilisation might have different aerodynamic efficiency. The model is extended by Lai et al. (2008b) who consider multiple trains, because the aerodynamic efficiency increases with a greater flexibility in placing loads on different trains. However, the benefit of optimising more future trains can be offset by the risk of making wrong decisions due to imperfect information on future trains and incoming load units.

4.3 Conclusions and research opportunities

Automation of the load planning process can support train planners by providing a load plan, and by optimising the capacity utilisation of the train. Load planning models can be used in an environment in which load units are all available before loading, as is for example often the case in seaport terminals. They can offer the decision maker a support tool incorporating real-time, integrated information in such a way that he can make fast decisions (Corry and Kozan, 2006a). The type of operator deciding on the train load plan determines the way in which a train load plan is constructed and revised. The intermodal operator focuses on optimising capacity utilisation of the available trains and managing dynamics during the booking process. The first, main objective of transport planners is maximising the throughput by optimally utilising the given loading space. Improving the capacity utilisation on the level of a single train provides a cheap way of increasing the throughput and reducing the cost per load unit. Within the aim of stimulating intermodal transport, this is an important aspect. Next, the terminal operator minimises handling costs with a given number of load units and dynamically revises the load plan with this cost minimisation as a driver. Cost minimisation of terminal operations is a secondary objective as this objective is usually optimised given a number of predetermined load units to be loaded. Terminal operations also influence the efficiency as they are directly connected with the transshipment cost between two transport modes.

Real-life loading constraints should be included. A simplification of the real situation by omitting loading constraints might possibly lead to an infeasible or undesirable composition of the load plan. Existing literature considers some but not all real-life problem characteristics required in order to assist planners in their decision-making process. For example, currently only wagon axle constraints for wagons with two bogies are considered, while three-bogie wagons are used in practice as well.

Moreover, in a real-world planning environment, train planners look into the future and account for a longer planning horizon. Based on their experience, they look a number of days ahead to establish a first load plan. With information about capacity on future trains and expected or known transport requests, they make the best decision accounting for the future. Detailed load plans are established one day before train departure and the intermodal operator sends them to the terminal operator. Current literature focuses on load planning problems for a single train without looking ahead when making assignment decisions. Only two papers (Anghinolfi and Paolucci, 2014; Lai et al., 2008b) take into account multiple trains. However, when load units are assigned to trains, a longer planning horizon must be incorporated in

the pure train load planning problem. This can be executed by means of prioritising more urgent load units or including a rolling horizon, in order to further assist planners while bookings arrive into the transport system. In this way, it is possible to balance load units to be shipped based on the available capacity according to their urgency. It may be interesting to study multi-period planning problems to provide planners with a longer term view on current capacity plans and address possible issues when they occur.

Although most scientific literature investigates static planning, decision support for planners in a dynamic environment, which is known to be a recent and future trend in line with synchromodal transport, should be available. Dynamic adaptations are triggered by certain events, such as a truck arriving late or urgent orders emerging, and should be communicated to the terminal operator as soon as possible. Current formulations require approaches which are able to account for this dynamic environment train planners daily operate in. More specifically, with the increasing importance of uncertainty about future demand and the occurrence of disturbances in the planning environment, planners must be supported in the decision making during the booking process. Multiple objectives may need to be included to provide a number of load plans.

Moreover, planners often receive a lot of information by means of communication with other operators. Planners of trucking operations can provide information on the locations of available trucks and as such, train plans may need to be revised. Furthermore, at different points in time, a different objective may be more critical, but this is not defined in advance. Therefore, planners could be offered a number of options to plan their train loading which are a priori indifferent to provide appropriate decision support.

To optimise both train load planning and crane and storage planning decisions, information must be available to all parties involved. This is also relevant for the development of dynamic models. Integrating data flows and using real-time information can provide more accurate data to further optimise the train load planning problem. It may support train planning decision-making whenever changes to the load planning occur. For example, when an urgent load unit emerges and needs to be planned on a train, train planners require knowledge on which load unit is already loaded on the train to prevent unloading a loaded load unit. Furthermore, information on the position of load units in stacks may be shared between parties in order to establish well-founded decisions.

A number of studies indicate that the problem is becoming highly complex when all relevant constraints and dynamic influences should be reflected in the train load

planning problem. Heuristics may provide a good solution to the problem after a short amount of computation time, or when small changes occur and an initial plan must be updated. In a dynamic planning environment, in which decisions are made in real-time, this can be an relevant characteristic of train load planning solution methods. Therefore, it may be interesting to further explore the possibility to develop heuristics to solve the train load planning problem.

To conclude, the problem complexity for the train load planning depends on the combination of loading constraints, constraints in the specific operational environment, influencing factors determined by the involved actors (e.g., a terminal operator has other objectives than an intermodal operator) and dynamic influences on the problem. The problem consists of two stages, which might imply that it could be solved in two phases. The initial focus should be on maximising throughput, after which efficient terminal operations can be designed. In a complex intermodal transport system, automated train load planning with adequate decision support can aid intermodal rail decision-makers in their decision-making process. It is a rather young research area, which may be further developed to assist in the decision-making process and optimise train planning.

In Chapters 5 and 6, the focus will be on solution methods with real-life loading constraints. The problem is studied from the viewpoint of an intermodal operator and aims to provide intermodal planning support during the booking process in a multi-objective setting, with the aim of maximising the available capacity utilisation. While a static problem setting is considered, information about the future is indirectly included by considering the urgency with which load units should be transported.

A multi-objective train load planning problem with real-life characteristics: formulation and exact solution approach

5.1 Introduction

The literature review in Chapter 4 identifies the factors influencing the train load planning problem and classifies current research accordingly. The overview indicates that multi-objective approaches for train load planning are a recent area of study. Planners may have multiple conflicting objectives, where no predefined fixed weights exist for the objectives. Therefore, in this chapter, a multi-objective train load planning problem with real-life characteristics is presented, as well as an exact solution approach to solve the problem (Figure 5.1).

The multi-objective approach provides planners with a decision support tool and results in a number of train load plans. Planners can then decide which of the plans suits the current operational environment the most. Ambrosino et al. (2016) focus in a conference paper on multiple objectives regarding crane operations and reshuffles in the storage area at seaports. They maximise the commercial value of loaded units and

This chapter is based on the following paper:
Heggen, H., Braekers, K., Caris, A., 2018. A multi-objective approach for intermodal train load planning. *OR Spectrum*, 40(2), pp. 341-366.

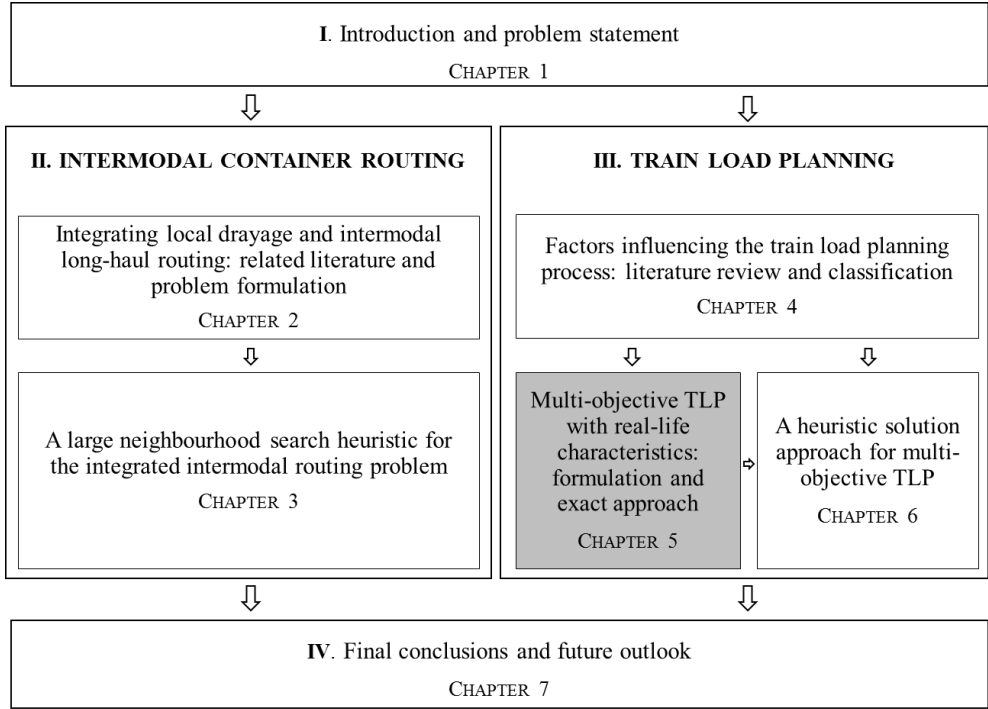


Figure 5.1: Thesis summary - Chapter 5.

minimise the number of reshuffles in the storage area and the number of pin changes. As opposed to Ambrosino et al. (2016), we consider a multi-objective problem from the perspective of an operator who receives transport orders and has to plan these according to the customer requirements and operational circumstances within an intermodal rail-road context. More specifically, our problem is aimed at maximising train length utilisation, scores for assigning load units to a more preferred, flexible destination and the number of urgent load units assigned.

Compared to existing research, additional characteristics of a real-life train planning environment are introduced. The model is able to handle wagons with a third bogie at its centre, while previous models only considered two bogies. Next, it is expanded with practical constraints that train planners are confronted with. The model deals with trains stopping at an intermediate terminal before arriving at the final destination terminal, including a more restrictive path weight limit between the intermediate and final destination terminal. Within this context, load units can have flexible destinations. Therefore, preferences for being unloaded at the intermediate or final terminal are determined by the distance between the unload terminal and the load unit's final destination in order to minimise the amount of kilometres travelled

on the road. Another consideration is the fact that some load units are very urgent and must be assigned to a location on the train to arrive at its destination in time. The inclusion of a measure for urgency aims at accounting for future time periods, hereby incorporating a longer-term planning horizon.

In this chapter, a new train load planning problem is proposed, inspired by a real-life case study as described in Section 5.2. The objective function is based on the viewpoint of a train planning department of a company which manages incoming transport orders and has its own trains. Wagons with three bogies are accounted for. A mathematical formulation of the problem is presented, and practical elements considered important by train planners in a real-life case study are added to the model. Section 5.3 presents an exact adaptive ϵ -constraint solution method to solve the problem and Section 5.4 summarises the characteristics of the problem instances used to analyse the exact method. Computational results for the ϵ -constraint method, obtained using ILOG Cplex 12.6 in C++, are discussed in Section 5.5. The main conclusions are summarised in Section 5.6.

5.2 Problem definition

The model in this chapter focuses on the composition of load plans during the planning process, not on its execution by terminal equipment. Furthermore, it is assumed that the demand for transport of load units is higher than the available space on the train. This makes the problem twofold: a decision should be made on which load units will be loaded, as well as their location on the train. Consequently, the objectives will be based on optimising train utilisation with a given number of wagons. A static model is presented.

5.2.1 Introduction to multi-objective concepts

Train planners have multiple objectives which may be conflicting, hence the importance of having a number of load plans to choose from based on the current real-time operational situation. For example, planners may have more information on the availability of future trains and alternative routes for some load units to be sent to their unload location. The weighted-sum method provides a single solution and assumes predefined weights for the objectives. It may be a difficult task to fix the weights of each objective in advance. Therefore, a multi-objective approach will be proposed.

In a multi-objective maximisation problem, one would prefer to find an ideal point where all objectives have their highest value. However, usually maximising one

objective implies a reduction of the remaining objective values. Therefore, the solution does not consist of a single optimal solution, but is depicted by a Pareto-front, which contains a number of non-dominated solutions. A solution is Pareto-optimal if all other solutions have a lower value for at least one of the objective functions, or have the same value for all of the objective functions (Caramia and Dell’Olmo, 2008). All obtained solutions on the Pareto-front are a priori of equal rank: a solution may be better with respect to one objective, but worse for another. In advance, the solutions are non-comparable. In multi-objective optimisation, the aim is to find the Pareto-optimal solutions. These concepts are illustrated in Figure 5.2 for a bi-objective maximisation problem. The dotted lines are solely for visual representation, and do not represent the solution space.

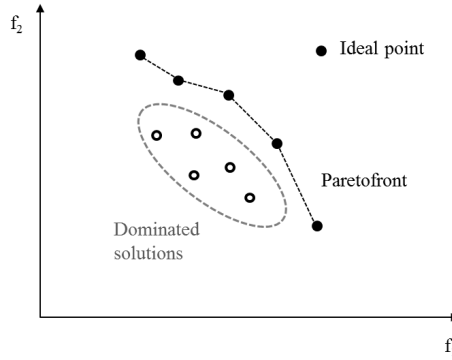


Figure 5.2: Example of a bi-objective maximisation problem.

For the train load planning problem it implies that a number of load plans are provided as output. In this way, train planners can decide a posteriori on the relative weights of each objective implicitly based on information of the current operational situation by choosing one of the load plan solutions.

For a more comprehensive overview of multi-objective optimisation, the reader is referred to Ehrgott (2005).

5.2.2 Multi-objective model formulation

The mathematical formulation builds on the third IP model of Bruns and Knust (2012), which is the only train load planning problem considering axle loads as continuous functions. This provides more flexibility as not all weight configurations must explicitly be defined, and has proven to lead to smaller run times. Our formulation is

based on load unit length-type categorisations instead of fixation types, because only five main length types are available in the real-life case study. Load units of 20, 25, 30, 40 and 45 feet are considered, which are common length types used in practice. A wagon configuration consists of a combination of load units, each with a length type. For example, a 60-feet wagon can be filled with two 30-feet load units, but also with a single 20-feet load unit.

The aim is to find an assignment of load units to slots on an intermodal train, while accounting for a number of weight restrictions. The total train weight, which consists of the weights of all wagons and load units assigned, is limited. Additionally, the weight that each wagon can carry is limited, which is not the case in Bruns and Knust (2012). Finally, the allowed payload on each bogie is restricted, and the bogie payloads should be balanced compared to each other for stability reasons. In our view, the limit on the slot weights as proposed by Bruns and Knust (2012) becomes redundant due to the above restrictions, and thus slot weight limits are not explicitly incorporated in this problem formulation.

Moreover, realistic elements are added to the model based on observations of a real-life train planning department. The model is able to account for wagons with a third bogie at the wagon centre. In that case, the tare wagon weight is distributed over all three axles: 50% over the centre bogie, and 25% over each front and rear bogie. This differs from existing models in which only wagons with two axles are considered. On the one hand, using continuous weight restrictions provides more flexibility. More solutions are feasible compared to a discrete set of feasible weight combinations in a given configuration of load unit types, because more feasible weight combinations are considered (Bruns and Knust, 2012). On the other hand, this requires a unique formula for two- and three-bogie wagons due to different distribution of the weights across the axles, whereas a discrete set of weight combinations can be determined once in advance (based on formulas, which are no longer required once the set is given). Furthermore, the objective is formulated as a multi-objective function, where on the one hand the loading degree in terms of loaded length is maximised and on the other hand the priority scores for urgent load units assigned are maximised to prevent load units not reaching their destination in time. The following notation is used:

Sets, indices and parameters

- I = $\{1, \dots, n\}$ = set of load units with index i
- J = $\{1, \dots, m\}$ = set of wagons with index j
- R = $\{1, \dots, r\}$ = set of wagon types with index r
- κ_r = set of physical configurations for wagons of type r , with index k

S_{jk}	= set of all possible slots of configuration k for wagon j , with index s
l_i	= length of load unit i
p_i	= urgency of load unit i
g_i	= weight of load unit i , tare load unit weight included
w_j	= tare wagon weight of wagon j
G	= train weight limit, tare wagon weights w_j included
W_j	= wagon weight limit for wagon j
τ_j	= wagon type of wagon j , where $\tau_j \in R$
γ_{τ_j}	= maximum feasible payload for the bogies of wagon type τ_j , which is the same for each bogie of a single wagon
d_{τ_j}	= distance between two adjacent bogies for wagon type τ_j
$e_{\tau_j}^{ks}$	= distance between the centre of the load unit and the bogie in front of slot s in configuration k for wagon type τ_j
z_j	= $\begin{cases} 1, & \text{if wagon } j \text{ has two bogies} \\ 0, & \text{if wagon } j \text{ has three bogies} \end{cases}$

Decision variables

a_j	= payload on bogie a for wagon j
b_j	= payload on bogie b for wagon j
c_j	= payload on bogie c for wagon j , only if it has three bogies
y_{jk}	= $\begin{cases} 1, & \text{if configuration } k \text{ is chosen for wagon } j \\ 0, & \text{otherwise} \end{cases}$
x_{ijks}	= $\begin{cases} 1, & \text{if load unit } i \text{ is assigned to slot } s \text{ in configuration } k \text{ of wagon } j, \\ \text{defined only if load unit } i \text{ fits onto slot } s \text{ of wagon } j \text{ in configuration } k \\ 0, & \text{otherwise} \end{cases}$

The problem is formulated as follows:

Problem P_1 :

$$\max f = (f_1, f_2) \quad (5.1)$$

$$f_1 = \sum_{i \in I} \sum_{j \in J} \sum_{k \in \tau_j} \sum_{s \in S_{jk}} l_i \cdot x_{ijks} \quad (5.2)$$

$$f_2 = \sum_{i \in I} \sum_{j \in J} \sum_{k \in \tau_j} \sum_{s \in S_{jk}} p_i \cdot x_{ijks} \quad (5.3)$$

subject to

$$\sum_{j \in J} \sum_{k \in \kappa_{\tau_j}} \sum_{s \in S_{jk}} x_{ijks} \leq 1 \quad \forall i \in I \quad (5.4)$$

$$\sum_{i \in I} x_{ijks} \leq 1 \quad \forall j \in J, k \in \kappa_{\tau_j}, s \in S_{jk} \quad (5.5)$$

$$\sum_{k \in \kappa_{\tau_j}} y_{jk} = 1 \quad \forall j \in J \quad (5.6)$$

$$a_j = \frac{w_j}{2} + \sum_{i \in I} \sum_{k \in \kappa_{\tau_j}} \sum_{s \in S_{jk}} g_i \cdot \frac{d_{\tau_j} - e_{\tau_j}^{ks}}{d_{\tau_j}} \cdot x_{ijks} \quad \forall j \in J, z_j = 1 \quad (5.7)$$

$$b_j = \frac{w_j}{2} + \sum_{i \in I} \sum_{k \in \kappa_{\tau_j}} \sum_{s \in S_{jk}} g_i \cdot \frac{e_{\tau_j}^{ks}}{d_{\tau_j}} \cdot x_{ijks} \quad \forall j \in J, z_j = 1 \quad (5.8)$$

$$a_j = \frac{w_j}{4} + \sum_{i \in I} \sum_{k \in \kappa_{\tau_j}} \sum_{s=1}^2 g_i \cdot \frac{d_{\tau_j} - e_{\tau_j}^{ks}}{d_{\tau_j}} \cdot x_{ijks} \quad \forall j \in J, z_j = 0 \quad (5.9)$$

$$b_j = \frac{w_j}{2} + \sum_{i \in I} \sum_{k \in \kappa_{\tau_j}} \sum_{s=1}^2 g_i \cdot \frac{e_{\tau_j}^{ks}}{d_{\tau_j}} \cdot x_{ijks} \\ + \sum_{i \in I} \sum_{k \in \kappa_{\tau_j}} \sum_{s=3}^4 g_i \cdot \frac{d_{\tau_j} - e_{\tau_j}^{ks}}{d_{\tau_j}} \cdot x_{ijks} \quad \forall j \in J, z_j = 0 \quad (5.10)$$

$$c_j = \frac{w_j}{4} + \sum_{i \in I} \sum_{k \in \kappa_{\tau_j}} \sum_{s=3}^4 g_i \cdot \frac{e_{\tau_j}^{ks}}{d_{\tau_j}} \cdot x_{ijks} \quad \forall j \in J, z_j = 0 \quad (5.11)$$

$$a_j \leq \gamma_{\tau_j} \quad \forall j \in J \quad (5.12)$$

$$b_j \leq \gamma_{\tau_j} \quad \forall j \in J \quad (5.13)$$

$$c_j \leq \gamma_{\tau_j} \quad \forall j \in J \quad (5.14)$$

$$a_j - 3 \cdot b_j \leq 0 \quad \forall j \in J \quad (5.15)$$

$$b_j - 3 \cdot a_j \leq 0 \quad \forall j \in J \quad (5.16)$$

$$b_j - 3 \cdot c_j \leq 0 \quad \forall j \in J \quad (5.17)$$

$$c_j - 3 \cdot b_j \leq 0 \quad \forall j \in J \quad (5.18)$$

$$\sum_{i \in I} \sum_{k \in \tau_j} \sum_{s \in S_{jk}} g_i \cdot x_{ijks} \leq W_j \quad \forall j \in J \quad (5.19)$$

$$\sum_{i \in I} \sum_{j \in J} \sum_{k \in \kappa_{\tau_j}} \sum_{s \in S_{jk}} g_i \cdot x_{ijks} + \sum_{j \in J} w_j \leq G \quad (5.20)$$

$$x_{ijks} \in \{0, 1\} \quad \forall i \in I, j \in J, k \in \kappa_{\tau_j}, s \in S_{jk} \quad (5.21)$$

$$y_{jk} \in \{0, 1\} \quad \forall j \in J, k \in \kappa_{\tau_j} \quad (5.22)$$

The multi-objective function (5.1) maximises the utilisation of the given, available loading length (5.2) as well as the priority scores for urgent load units assigned to locations on the train (5.3). Each load unit can be assigned to at most one slot (5.4), and a slot can only carry one load unit (5.5). Constraint (5.6) guarantees that a single configuration per wagon is chosen. The payload of each bogie is determined by (5.7) and (5.8) for wagons with two bogies, and by (5.9)-(5.11) for wagons with three bogies, where it is assumed that centre of mass of the load unit is at its middle. It is assumed that wagons with three bogies can carry at most four load units, two at each side of the centre bogie. In case a configuration with only two slots is chosen, one at each side of the centre, indices $s = 1$ and $s = 3$ are used for these slots. The bogie payloads are limited in (5.12)-(5.14). Constraints (5.15)-(5.18) ensure that the payloads on each bogie are balanced relative to the adjacent bogies. This balance is set to maximum three times the weight of the adjacent bogie, as done by Bruns and Knust (2012). The allowed wagon weight and train weight are limited by (5.19) and (5.20) respectively. Finally, (5.21) and (5.22) define the domain of the decision variables.

5.2.3 Adding practical considerations

In this section, practical constraints are added to the model, based on the problem context of Move Intermodal, an intermodal transport company with activities throughout Europe. One of its main activities concerns the intermodal rail-road connection between Belgium and the Netherlands on the one hand, and two unload terminals in the north of Italy on the other hand.

Trains leaving Belgium or the Netherlands have two destinations, an intermediate destination and a final destination. The first u wagons of the train are decoupled at the intermediate terminal, while the remaining ones continue to the final terminal. Because of a weight restriction on the railway path between both unload terminals, the total train weight limit is lower for the path between the intermediate rail stop and the final destination. In this case, Eq. (5.20) must be supplemented by an additional restriction (5.23), where G' constitutes the train weight limit for the wagons which will be unloaded at the final destination terminal, after unloading the first u wagons

at the intermediate terminal. Consequently, the following restriction is added to the problem, where $G' < G$.

$$\sum_{i \in I} \sum_{\substack{j \in J \\ j > u}} \sum_{\substack{k \in \tau_j \\ s \in S_{jk}}} g_i \cdot x_{ijks} + \sum_{\substack{j \in J \\ j > u}} w_j \leq G' \quad (5.23)$$

Due to the destination grouping of wagons, load units should be grouped per destination. This can be accomplished by assigning a destination preference score λ_{ij} to each load unit, indicating the preference for assigning load unit i to a wagon j with destination D_j . These preferences can be based on the proximity of the unload terminal to the load unit's final destination, where a larger value of λ_{ij} means that the destination of wagon j is more preferred for load unit i . To assign as much load units as possible to their preferred destination, a maximisation of the destination preference score is added to the objective function, as in (5.24) and (5.25).

$$\max f = (f_1, f_2, f_3) \quad (5.24)$$

$$f_3 = \sum_{i \in I} \sum_{j \in J} \sum_{k \in \tau_j} \sum_{s \in S_{jk}} \lambda_{ij} \cdot x_{ijks} \quad (5.25)$$

Furthermore, to incorporate the urgency with which each load unit must arrive at its destination, load units are divided into three categories based on the real-life setting of the train planning environment. Therefore, priority parameters split load units into three classes: critical load units ($p_i = 2$), urgent load units ($p_i = 1$), and non-urgent load units ($p_i = 0$). Critical load units must be assigned to a location on the train to arrive at its destination in time. They receive the highest priority and must always be assigned to a location on the train. Urgent load units have a margin of one day on the ultimate arrival date, and all other load units are non-urgent. These definitions are used throughout the remainder of this chapter. Assuming that all critical load units can be loaded onto the train, a constraint is added to impose all critical load units to be loaded (5.26). This leaves only one urgency objective, namely assigning as much urgent load units ($p_i = 1$) as possible.

$$\sum_{j \in J} \sum_{k \in \kappa_{\tau_j}} \sum_{s \in S_{jk}} x_{ijks} = 1 \quad \forall i \in I_{p_i=2} \quad (5.26)$$

In the remainder of this chapter, these practical considerations are added to the base model to constitute a single train load planning problem formulation.

5.3 A multi-objective exact solution approach

An exact solution approach is presented for the multi-objective train load planning problem with three objectives. From the perspective of the intermodal operator who manages its own trains, three objectives are considered: maximising train length utilisation, maximising scores for load units assigned to a more preferred destination (which may be based on the proximity of the load unit's final destination to the two unload terminals), and maximising the number of urgent load units ($p_i = 1$) assigned. No fixed or predefined hierarchy between these three objectives exists. For example, loading more metres on the train may be conflicting with the optimal number of urgent load units assigned. Furthermore, it may be beneficial to assign fewer load units to a more preferred destination if this leads to more load units assigned. However, this trade-off is not always defined a priori and depends on the current operational environment. For example, expectations about future bookings and availability of slots on trains departing from terminals in the close vicinity influence the final assignment decision.

The exact ϵ -constraint method can be used to solve problems with multiple objectives. The method consists of maximising one objective (5.27), while iteratively increasing (lower) bounds on the other objectives (5.28) until no feasible solution can be obtained anymore. While ϵ_1 and ϵ_2 represent lower bounds, ϵ'_1 and ϵ'_2 are upper bounds for f_1 and f_2 respectively. The solution space Z (5.29) consists of all feasible solutions. For more information on this method, the reader is referred to Chankong and Haimes (1983). The general three-objective maximisation problem can be formulated as follows:

Problem P_2 :

$$\max f_3(z) \tag{5.27}$$

subject to

$$\epsilon_v \leq f_v(z) < \epsilon'_v \quad \forall v \in \{1, 2\} \tag{5.28}$$

$$z \in Z \tag{5.29}$$

Recent research efforts focus on improving the efficiency of the classical ϵ -constraint method, as it is not trivial to implement for more than two objectives. For example, Mavrotas (2009) and Zhang and Reimann (2014) propose acceleration mechanisms such as early exit at infeasible solutions, additions to guarantee strictly dominating

solutions, methods to limit the range of objective function values and the addition of bouncing steps in the loops. The two-dimensional adaptive ϵ -constraint method of Kovacs et al. (2015) includes a memory (k) for tightening the bounds on one objective, and relies on the knowledge that the number of different values for one of the objectives is usually relatively low. As this holds for the priority scores f_2 , the procedure can be applied to the train load planning problem. For any value of this objective, a bi-objective problem can be solved with modified ϵ -constraints.

In this chapter, an exact solution method based on the ϵ -constraint method is used to solve the train load planning problem with three objectives. The idea is to repeatedly solve a single-objective problem which maximises destination preferences, with varying lower bounds on the other two objectives (length utilisation and priority scores). The upper bounds stay constant and can be calculated easily in advance: ϵ'_1 corresponds to the available train length plus one, ϵ'_2 corresponds to the sum of all priority scores over all available load units plus one. A number of redundant subproblems are avoided by using a hierarchical objective function (5.30) replacing (5.27) in problem P_2 , hereby reducing the number of iterations as the highest possible value for the secondary and tertiary objectives are selected for a given value of the primary objective without missing any non-dominated solution - to become problem P_3 .

$$\max f_3 + \frac{f_1}{UB_1} + \frac{f_2}{UB_2} \quad (5.30)$$

The main objective is to maximise destination preferences f_3 (5.25). Maximising length utilisation f_1 (5.2) and priority scores f_2 (5.3) are the secondary and tertiary objectives respectively. Note that all three objective values are integer when assuming l_i , p_i and λ_{ij} to be integer parameters. In order to avoid iterations with solutions which will be dominated by solutions of later iterations, a hierarchical objective function (5.30) is applied (instead of just maximising f_3) with $\frac{f_2}{UB_2} < \frac{f_1}{UB_1} < 1$, $f_1 < UB_1 = \epsilon'_1$, $f_2 < UB_2$, and UB_2 set such that the above conditions hold, i.e., larger than $f_2 * \frac{UB_1}{f_1}$ (for any combination of possible values for f_1 and f_2) and consequently larger than ϵ'_2 . Further, objectives f_1 and f_2 are bounded by (5.28). This approach guarantees that no non-dominated points are missed in the three-dimensional grid. Finally, the solution space Z (5.29) consists of all solutions which satisfy constraints (5.4)–(5.23), and (5.26).

The two-dimensional adaptive ϵ -constraint algorithm of Kovacs et al. (2015) is used to guide the iterative procedure of finding all solutions on the Pareto-front. In this chapter, the number of urgent load units assigned to a location on the train

(f_2) can only take a limited amount of values, especially when all critical load units are assigned. Consequently, a two-dimensional problem can be solved for each possible value of this objective. The implementation of the two-dimensional adaptive ϵ -constraint approach for the train load planning problem is provided in pseudocode (Algorithm 2) and works as follows. Each single-objective problem maximises the objective function (5.30). In a first phase, the problem is iteratively solved with an increasing lower bound ϵ_1 on the loaded length f_1 (lines 6–10). As the minimum length difference between two load unit types amounts to five feet, the increase is set to five feet per iteration ($\delta_1 = 5$). At this point, f_2 is still unbounded. After each phase-one iteration, the resulting f_1 -value increased with $\delta_1 = 5$ serves as lower bound for the next iteration. In the second phase (lines 3–12), a memory (k) is used for tightening the bounds on f_2 , similar as in Kovacs et al. (2015). The first-phase two-dimensional procedure is repeated with the lower bound ϵ_2 on the number of urgent load units set as the minimum f_2 -value found in the previous first-phase iteration, increased with one load unit ($\delta_2 = 1$). This procedure is repeated until no feasible solution can be obtained anymore.

Algorithm 2 Three-objective ϵ -constraint method for train load planning.

```

1:  $\epsilon_1 = 0$ 
2:  $\epsilon_2 = 0$ 
3: while  $\epsilon_2 < \epsilon'_2$  do
4:    $k = \infty$ 
5:    $\epsilon_1 = 0$ 
6:   while  $\epsilon_1 < \epsilon'_1$  do
7:     Solve problem  $P_3$ 
8:      $\epsilon_1 = f_1 + \delta_1$ 
9:      $k = \min(k, f_2)$ 
10:  end while
11:   $\epsilon_2 = k + \delta_2$ 
12: end while

```

The output of this method will be a set of Pareto-optimal solutions, containing at least one solution for each point on the Pareto-front (Kovacs et al., 2015). From this pool of solutions, train planners select the one which is most appropriate for the current operational circumstances.

One disadvantage of this method is that a large number of iterations might be required. On the other hand, the method guarantees to find all non-dominated solu-

tions, including non-supported ones. This is not the case if alternatively for example a weighted objective, with varying weights, would be used to find non-dominated solutions. Furthermore, setting these weights is also not a trivial task. The main advantage of obtaining this complete set is that it provides insights on what the front looks like, without having to set arbitrary weights.

5.4 Problem instances

In current literature, data is produced based on real-life characteristics of load units, wagons and configurations in Italy (Ambrosino and Siri, 2014, 2015; Ambrosino et al., 2011, 2013, 2016; Anghinolfi and Paolucci, 2014; Anghinolfi et al., 2014), Germany (Bruns and Knust, 2012; Bruns et al., 2014), Australia (Corry and Kozan, 2004, 2006a,b, 2008) and the United States (Feo and González-Velarde, 1995; Upadhyay et al., 2017; Mantovani et al., 2018). As no benchmark instances are publicly available, input data is generated based on historical information about load units transported from Belgium to two locations in Northern Italy by the considered company. As such, the generated instances are representative for the real-life problem. The used instances are available upon request.

To assess the performance of the solution methods in terms of computation times and to examine the resulting generated load plans, a number of experiments are conducted using CPLEX. By changing parameter values of a number of input factors, the influence of these factors on the computation time and the trade-offs are analysed. Three factors are considered: the number of available wagons, the number of critical load units available and the weights of the available load units. All combinations are considered, resulting in 18 (3x2x3) instance classes. An overview of the experimental design is presented in Table 5.1.

Table 5.1: Input data: problem types for train load planning.

Factor	Level		
	1	2	3
Number of wagons	5	10	20
Critical load units (% of available load units)	35%	20%	/
Load unit weights	TRIA(17,20,23)	TRIA(23,26,29)	UNIF(17,29)

The three factors are considered on multiple levels. First, the number of available wagons is assumed to be five, ten or twenty. The problem type with twenty wagons can be considered a realistic wagon set. However, also smaller instances are considered in order to allow the comparison of the multi-directional local search heuristic – which will be presented in Chapter 6 – with the exact solutions, as additional experiments on a preliminary single-objective version of the model indicated that large instances might not always be solved to optimality in short computation times (Heggen et al., 2016). The wagon set is fixed to a representative composition with respect to the relative amount of wagons available for each terminal as well as the combination of wagon types used. The number of load units available is determined in such a way that it is directly related to the available wagons. New load units are randomly added to the pool of available load units until they constitute 1.5 times the wagon space available, while a fixed split of load unit types is maintained based on the analysis of real data. Destination preferences of load units to be assigned to one out of two destinations are determined based on the proximity of their final unload location to each of the two intermodal terminals. A load unit has a strong destination preference if it is very close to one terminal ($\lambda_{ij} = 100$). If they are a little further from one of the terminals, a weak preference for being assigned to a terminal ($\lambda_{ij} = 10$) is assigned. All other load units are assumed not to have a clear preference because the difference in distance between each terminal and the load unit unload location is not substantial. A second factor concerns the number of critical load units available, presented as a percentage of all available load units. It is set to either 35% or 20%, while the percentage of urgent load units is kept constant at 20% in order to clearly show the impact of the number of critical load units. Consequently, differences occur with respect to the amount of non-urgent load units. Finally, three levels of load unit weights are considered in the experimental design: light load units, heavy load units and units with equally distributed weights. Input for light and heavy load units is generated using a triangular distribution with a minimum, a modus and a maximum for the load unit weights ($TRIA(min, modus, max)$). These three parameters are chosen based on realistic values for light and heavy load unit weights. The equally distributed load unit weights are obtained by means of a uniform distribution of the load unit weights ($UNIF(min, max)$) with a minimum weight equal to the minimum weight of the instances with light load units and a maximum weight equal to the maximum weight of the instances with heavy load units. For each combination of factors, 10 instances are randomly sampled, resulting in a total of 180 instances.

5.5 Computational experiments for the ϵ -constraint method

A time limit of one hour per single-objective run is imposed to the exact ϵ -constraint method (line 7 of Algorithm 2) to avoid excessive total run times for the calculation of the pool of load plan solutions for a single instance. Results with respect to the computation times are given in Table 5.2, information on the obtained solutions is displayed in Table 5.3.

The left part of Table 5.2 presents computation times related to the entire ϵ -constraint procedure, i.e., the total time for finding all solutions, either optimal or at maximum run time. On the right-hand side, results are shown for all runs which reached the time limit, i.e., which were not solved to optimality in time and thus result in an optimality gap. The average gap represents the percentage difference between the lower and upper bound obtained for solutions not solved to optimality within the time limit. A full and detailed overview of the obtained results per instance can be found in Appendix A.1.

These results indicate that the train load planning problem cannot always be solved in a short computation time, as shown by the total computation times for the entire procedure in each problem category. Some single-objective runs consume a large amount of computation time, severely influencing average total computation times. Preliminary experiments already showed that axle weight limits and the balancing of payloads make the train load planning problem complex to solve. The size of the problem has an influence on the solution time. Problem instances with five available wagons are all solved in very short computation times, with a maximum total run time of 73.50 seconds to find all solutions on the Pareto-front. Most problems with ten available wagons produce higher average and maximum computation times, which partly may be due to the fact that more single-objective iterations are required. For realistic instances of 20 wagons, the exact Pareto-front consists of an even larger number of non-dominated solutions, and consequently more iterations exist within the procedure, further increasing run times.

The load unit weights also influence computation times for instances with ten available wagons. With heavy load units as input, total computation times increase drastically and not all single-objective iterations are solved to optimality within the given time limit of one hour per iteration. Instances with light load units show the smallest computation times, and instances with load unit weights drawn from a uniform distribution lie in between. An identical pattern can be observed for realistic

Table 5.2: Computational results for the ϵ -constraint method.

Problem instance	Instances	Computation times (s)				Solutions with max runtime		
		Avg	Min	Max	StDev	Instances	Iterations	Avg gap (%)
(5, 35%, light)	1-10	1.22	0.39	5.62	1.57			
(5, 35%, heavy)	11-20	1.25	0.25	6.97	2.06			
(5, 35%, unif)	21-30	11.31	0.31	73.50	22.92			
(5, 20%, light)	31-40	1.56	0.09	3.95	1.23			
(5, 20%, heavy)	41-50	1.15	0.20	3.41	1.01			
(5, 20%, unif)	51-60	2.86	0.27	10.31	3.53			
(10, 35%, light)	61-70	10.98	2.43	32.51	9.95			
(10, 35%, heavy)	71-80	699.93	2.93	3614.83	1458.54	2	2	0.62
(10, 35%, unif)	81-90	8.18	1.33	20.20	5.28			
(10, 20%, light)	91-100	11.69	1.38	57.78	16.93			
(10, 20%, heavy)	101-110	1677.01	4.64	8112.28	3346.11	2	4	1.67
(10, 20%, unif)	111-120	12.29	1.59	57.68	16.53			
(20, 35%, light)	121-130	855.77	7.06	4374.27	1676.84	2	2	0.28
(20, 35%, heavy)	131-140	39917.96	90.50	154278.99	55164.00	7	104	1.01
(20, 35%, unif)	141-150	8614.75	37.33	32524.43	11783.22	5	19	0.73
(20, 20%, light)	151-160	536.24	17.42	4777.64	1491.24	1	1	0.22
(20, 20%, heavy)	161-170	5283.01	118.36	33793.43	10641.46	3	12	0.47
(20, 20%, unif)	171-180	1680.34	28.37	14764.40	4598.26	1	4	0.18

instances with 20 wagons. However, for all problem categories the average optimality gap for solutions which are not solved to optimality does not exceed 2%.

Table 5.3 shows the average number of unique, non-dominated solutions obtained by the multi-objective approach for each problem instance, including both optimal solutions and solutions obtained at maximum run time. It further displays the difference between the minimum and maximum objective value for each objective for one instance, averaged over all instances with identical problem characteristics. The final column contains additional information with respect to the average number of available load units as a basis for comparison for the number of urgent load units assigned.

Table 5.3: Analysis of extreme points for the ϵ -constraint method.

Problem instance	Instances	Avg # solutions per instance	Avg difference between extremes			Avg # urgent LU available
			DP	length (ft)	# urgent LU	
(5, 35%, light)	1-10	5.3	187.0	23.0	2.2	4.0
(5, 35%, heavy)	11-20	4.4	154.0	15.5	2.1	4.1
(5, 35%, unif)	21-30	4.9	197.0	21.5	1.6	3.3
(5, 20%, light)	31-40	5.8	211.0	26.0	1.6	3.7
(5, 20%, heavy)	41-50	3.7	159.0	9.5	1.6	4.0
(5, 20%, unif)	51-60	6.5	223.0	26.5	1.6	3.3
(10, 35%, light)	61-70	11.8	320.0	24.0	3.0	6.5
(10, 35%, heavy)	71-80	8.5	258.0	21.0	3.1	7.0
(10, 35%, unif)	81-90	9.8	320.0	22.0	3.1	5.8
(10, 20%, light)	91-100	7.7	225.0	29.0	1.9	6.2
(10, 20%, heavy)	101-110	10.3	240.0	27.0	2.2	6.0
(10, 20%, unif)	111-120	11.0	223.0	19.5	2.5	6.1
(20, 35%, light)	121-130	18.0	470.0	24.5	5.4	13.3
(20, 35%, heavy)	131-140	29.3	631.0	46.0	6.9	13.2
(20, 35%, unif)	141-150	29.7	664.0	34.0	5.9	12.6
(20, 20%, light)	151-160	12.6	274.0	27.0	3.2	12.0
(20, 20%, heavy)	161-170	16.7	300.0	39.0	4.0	12.9
(20, 20%, unif)	171-180	17.7	325.0	25.5	3.9	11.8

DP = Destination preferences

LU = Load unit

Generally, the multi-objective approach provides a solution pool of four to seven load plans, and eight to twelve load plans for instances with five and ten wagons

respectively. The pool of non-dominated solutions is larger for realistic instances and consists on average of 12 to 30 load plans. The extreme points for each instance are relatively close to each other with respect to the average difference in assigned load unit length, compared to a maximum loading length of 360, 720 and 1440 feet for five, ten and twenty wagons respectively. However, larger differences exist with respect to the number of urgent load units assigned, taking into account that on average 11, 22 and 44 load units are assigned to a wagon set of five, ten and twenty available wagons respectively. Variations reflect around 50% of the average number of urgent load units available, but are smaller for instances which contain only 20% critical load units. Moreover, differences in the objective value for assigning more load units to a more preferred destination are relatively large.

In conclusion, results for the smaller instances show that, although the majority of the problems can be solved in short computation times, a number of single-objective problems require a large amount of computation time, resulting in a high total computation time for the procedure. Therefore, the solution method does not provide the user with a solution (i.e., the Pareto-front) in a reasonable amount of time. For instances of realistic problem size, this effect increases due to the fact that the number of Pareto-optimal solutions is larger for these instances and the single-objective problems become more complex to solve (because more variables and constraints are involved). This results in large total computation times for realistic instances, and is not desirable in practice.

5.6 Conclusions

In this chapter, a new train load planning problem is proposed which reflects realistic characteristics of the train load planning task, such as the urgency with which load units must be sent to their destination and the flexible routing of load units to one out of two terminals. The model further deals with wagons with three bogies. The model further deals with wagons with three bogies. Furthermore, a multi-objective solution approach to the train load planning problem is presented for the first time within the context of intermodal rail-road transport. A large amount of bookings enters the system and must be planned based on the customer requirements. The final train load plan strongly depends on information train planners have on the current operational circumstances, such as expected bookings and the availability of slots for possible alternatives. Therefore, planners have to account for multiple objectives at the same time, while a priori no fixed weights can be attached to each of the objectives.

The proposed model includes destination grouping due to unloading at a single intermediate terminal, and an additional objective to ensure that load units are directed to their preferred destination. Furthermore, load units are classified into three classes based on the priority with which they should be transported. Although these additions are mentioned as context-specific practical considerations, these are often valid in practice. For example, variations in the railway path limits increase for longer railway paths or when connections with different corridors have to be made. If the problem would be expanded with such railway paths or more intermediate stops, even more constraints similar to practical constraints (5.23) have to be added, while the objective of maximising destination preference scores can remain unchanged.

The exact ϵ -constraint method cannot always find the exact Pareto-front in short run times, even for relatively small problems compared to real-life trains. For realistic problems, computation times increase drastically, whereas in a real-life setting, planners expect a load plan in a few minutes. This indicates the need for heuristic methods to solve the problem, in order to guarantee train planners good solutions in a short amount of time, especially in order to deal with instances of realistic size. A heuristic for multi-objective train load planning is the subject of Chapter 6.

A heuristic solution approach for multi-objective train load planning

6.1 Introduction

A multi-objective train load planning problem formulation with realistic loading constraints has been presented in Chapter 5. Experimental results of the proposed exact adaptive ϵ -constraint method show that not all non-dominated solutions can be found within a reasonable amount of time, while planners expect fast decision support. In this chapter, a multi-objective heuristic for train load planning is proposed to solve the problem in a short amount of time (Figure 6.1).

Current literature often presents linear programming problems and solves instances of realistic size using commercial software (Bruns and Knust, 2012; Bruns et al., 2014; Corry and Kozan, 2006a,b, 2008; Feo and González-Velarde, 1995). The solutions are often used as a benchmark. Heuristic solution methods are usually based on a local search, sometimes combined with simulated annealing to overcome the problem of converging to local optima. The neighbourhood commonly consists of switching the location of two load units on a train (Corry and Kozan, 2004, 2006a,b;

This chapter is based on the following papers:

Heggen, H., Braekers, K., Caris, A., 2018. A multi-objective approach for intermodal train load planning. *OR Spectrum*, 40(2), pp. 341-366. (Sections 6.2, 6.5, 6.6);

Heggen, H., Braekers, K., Caris, A., 2017. An efficient heuristic for multi-objective train load planning: a parameter sensitivity analysis. In: Bruzzone, A.G., Piera, M.A., Merkuriev, Y., Longo, F. (Eds.), *Proceedings of HMS: The International Conference on Harbor, Maritime and Multimodal Logistics Modelling and Simulation*. Barcelona, Spain, September 18-20. (Sections 6.3, 6.4)

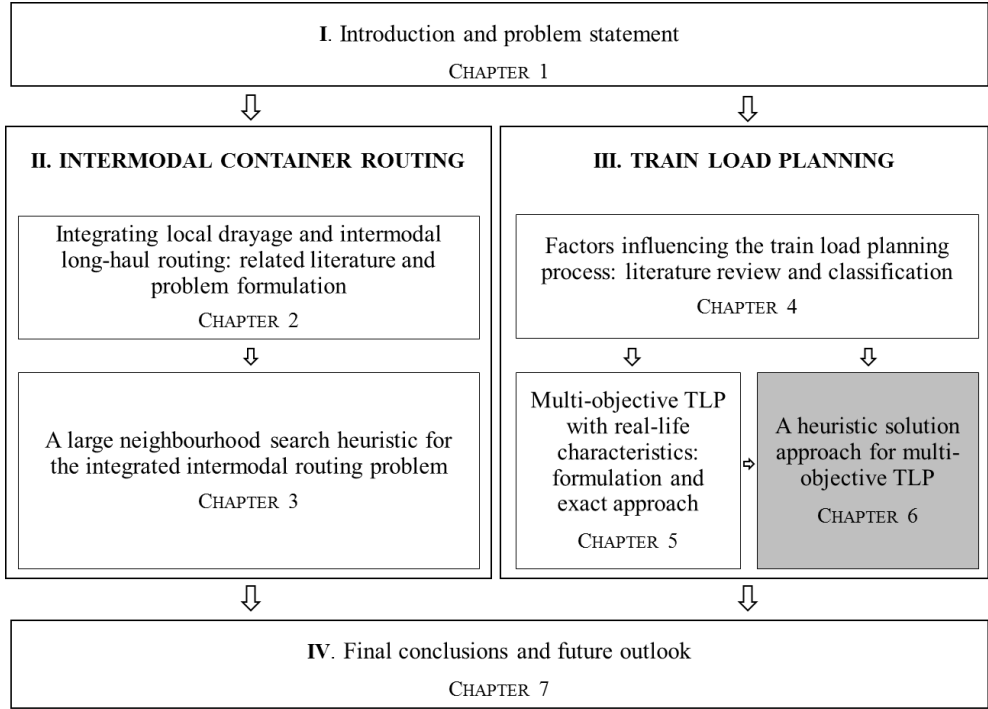


Figure 6.1: Thesis summary - Chapter 6.

Feo and González-Velarde, 1995). Corry and Kozan (2006b, 2008) combine this load unit swap neighbourhood together with a loading pattern neighbourhood, in which loading patterns of two wagons are exchanged while the number of slots is preserved. These heuristic methods are mainly applied to problems in which the load units to be loaded are fixed and known, and only the final assignment to specific locations on the train must be determined. The actual load unit's location on the train is uncertain because not all information about handling operations at the terminal is known in advance and a slot may still be occupied at the moment of arrival of a load unit assigned to that slot. In such circumstances, it becomes a problem with a rolling horizon in which the current load plan serves as initial solution and new events trigger a local search (Corry and Kozan, 2006a, 2004, 2008).

Recently, Ambrosino et al. (2011) integrated the optimisation of train load planning with optimisation of other operational decisions in an intermodal seaport terminal, in particular with the optimisation of crane and storage planning. Linear programs have been solved for this type of problem. Moreover, Anghinolfi and Paolucci (2014) propose a Lagrangian heuristic, which consists of a combination of a Lagrangian relaxation to obtain a first feasible solution with a randomised neighbourhood

search (RANS) matheuristic to improve it, whereas Anghinolfi et al. (2014) propose a GRASP.

In this chapter, a multi-directional local search heuristic is presented, focusing on a number of capacity-related objectives which train planners take into account during their planning process. The algorithm is presented in Section 6.2. Heuristic parameters are tuned in Section 6.3 and a sensitivity analysis is performed in Section 6.4 in order to analyse the impact of variations in values of the heuristic parameters on solution quality and computation time. In Section 6.5, the heuristic performance is assessed by comparing its results with the exact solutions with respect to the solution quality and computation times for small and realistic instances. All algorithms are implemented in C++. Insights in the trade-offs between the three objectives are provided in Section 6.6. Finally, Section 6.7 presents the main conclusions.

6.2 Multi-directional local search heuristic

During the planning process many events can occur which influence the assignment decision, such as the emergence of more urgent load units or planned load units which may not arrive at the terminal in time. Consequently, train planners should continuously monitor the feasibility of the train load plan, while accounting for the characteristics of the available load units. Furthermore, train load planning currently is often performed manually and thus may consume a relatively large amount of time. Because flexible planning is key to a dynamic, synchromodal environment, heuristics can be used if no commercial software is available or if the specific problem characteristics make the real-life problem too complex to solve exactly in a short amount of computation time.

Several types of heuristic algorithms have been proposed to solve multi-objective optimisation problems. Most of these algorithms extend concepts of stochastic local search to a multi-objective setting. The reader is referred to Paquete and Stützle (2018) for a general discussion and review. Especially evolutionary algorithms are popular (e.g. non-dominated sorting genetic algorithm, multi-objective genetic algorithm). New solutions are obtained by combining characteristics of good solutions. Such algorithms assume that combinations of good solutions will also perform well, and that diversity is required in the population. A key issue in these methods is the translation of the general fitness-concept to a multi-objective context (Tricoire, 2012). For a comprehensive overview on multi-objective evolutionary heuristic algorithms, the reader is referred to Coello et al. (2007) and Deb (2014). Alternatively,

local search-based heuristics, such as Pareto local search (Paquete et al., 2004) and multi-directional local search (Tricoire, 2012) have been proposed.

In this chapter, a heuristic based on the multi-directional local search (MDLS) framework of Tricoire (2012) is used to solve the train load planning problem with three objectives, presented in Chapter 5. A main advantage of this method is that a single-objective fitness function can be used to evaluate the fitness of a solution. Furthermore, Tricoire (2012) emphasises that MDLS only keeps track of non-dominated solutions, which differs from most well-known multi-objective heuristics. The method relies on the knowledge that it is sufficient to search in the direction of each of the objectives individually to find new, non-dominated solutions (Braekers et al., 2016a). Consequently, single-objective operators can be implemented in the framework. Other advantages include the flexibility and simplicity of the method. Furthermore, it has proven to lead to good computational results for various problems (e.g., Braekers et al. (2016a), Molenbruch et al. (2017)).

In this section, the proposed multi-objective heuristic to solve the problem (Section 6.2.1) and its components (Section 6.2.2 and 6.2.3) are described, as well as indicators for the heuristic quality assessment (Section 6.2.4).

6.2.1 General structure of the MDLS

The applied multi-directional local search (MDLS) heuristic framework is shown in Algorithm 3. A global pool of non-dominated solutions E is maintained and updated during the search. In an MDLS-iteration, a solution s is selected randomly from the solution pool E . Then, starting from this solution, a distinct local search LS_k for each objective k is performed, and the solution pool is updated.

For the train load planning problem, an initial solution is generated in a first constructive phase. Next, in the MDLS-framework, the local search operators are defined by altering the configuration of two wagons. Preliminary tests with one and three wagons respectively indicated no added value. For one-wagon operators, the applied change proved too small, whereas operators considering three wagons did not result in substantial performance differences. For each objective specifically, wagons are selected in a different way, and a distinct acceptance criterion is used. Each local search ends when a maximum number of consecutive iterations without improvement $n_{it}(LS_k)$ is reached.

Algorithm 3 MDLS-heuristic for the train load planning problem.

```

1: Constructive phase
2: Generate the initial solution;
3: Initialise the set of non-dominated solutions  $E$  with the initial solution;
4:
5: Multi-directional local search
6: for  $i = 1 \rightarrow n_{it(MDLS)}$  do
7:   Select a random solution  $s$  from  $E$ 
8:   for  $k = 1 \rightarrow 3$  do
9:      $LS_k$  on  $s$ , generating temporary solution set  $T$ 
10:    Update  $E$  with solutions in  $T$ 
11:   end for
12: end for
13: Return  $E$ 

```

6.2.2 Constructive phase

An initial solution is constructed by assigning load units to slots on each wagon, one by one, going from the front to the back of the train, using an intelligent candidate list to select load units first based on the highest priority, then highest weight. First, only critical load units are considered. Only after these are feasibly assigned, the remaining load units are considered. For the wagon under consideration, configurations which are more preferred with respect to the available length used are selected first. The available slots of the selected configuration are filled with load units matching the slot dimensions as long as the bogie, wagon and train weight limits are respected. If either not all slots in a configuration can be filled with the remaining available load units or the bogie balance limits are not respected, a following configuration is selected. Otherwise, all slots in the configuration are filled and the assignment procedure continues with the next wagon. This constructive phase results in a single initial solution, which is added to solution pool E .

6.2.3 Local search phase

Next, $n_{it(MDLS)}$ iterations of the MDLS heuristic are performed. One MDLS-iteration consists of three local searches on a single randomly selected solution $s \in E$. Each local search LS_k (Algorithm 4) guides the search primarily towards improving one objective.

Algorithm 4 Local search heuristic (LS_k) for each objective k .

```

1:  $T = \{\}$ 
2:  $n_{it(noImprovement)} = 0$ 
3: repeat
4:    $sol_{new} = sol_{current}$ 
5:    $sol_{new} \leftarrow$  Apply neighbourhood operator
6:    $LS_{Length}$ : change configuration of 2 wagons:
7:     1 with available length, 1 randomly selected
8:    $LS_{Urgency}$ : change configuration of 2 randomly selected wagons
9:    $LS_{DP}$ : change configuration of 2 randomly selected wagons
10:  if normalised, weighted-sum objective function satisfies acceptance criterion
    then
11:     $sol_{current} = sol_{new}$ 
12:  end if
13:  if  $f_k(new) > f_k(best)$  then
14:     $sol_{best} = sol_{current}$ 
15:     $n_{it(noImprovement)} = 0$ 
16:  else
17:     $n_{it(noImprovement)} ++$ 
18:  end if
19:  Update  $T$ 
20: until  $n_{it(noImprovement)} = n_{it(LS_k)}$ 
21: Return  $T$ 

```

6.2.3.1 Local search operators

The neighbourhood is defined by simultaneously altering the configuration of two wagons, i.e., assigning a new configuration to these wagons. The way in which wagons are selected differs depending on the main objective focus of the local search, after which all load units assigned to the two selected wagons are added to the pool of available, currently unassigned load units. Next, configurations for both wagons are selected randomly with a higher probability to be selected if a configuration uses a larger proportion of the available wagon length. The probability of selecting a configuration is determined by the contribution of the length used in that specific configuration compared to the total length of all possible configurations for one wagon type. In this way, the probability of rejecting a solution because it does not satisfy the acceptance criterion, is reduced.

For two given wagons and configurations, unassigned load units are selected (with critical load units first) to be assigned to a slot of these configurations with the same length as the load unit if payload, wagon and train weight limits are respected. If these constraints are not satisfied, new combinations are selected until all slots are tested. If, after testing all combinations, still no feasible assignment is obtained or if not enough unassigned load units are available to fill the remaining empty slots based on their dimensions, two new configurations are selected. Next, if all slots are filled with load units, it is checked whether all critical load units are assigned and the bogie balancing is respected. If not, the entire procedure restarts by selecting new wagons. Only if all critical load units can be assigned, dimensions of the selected load units match the slot dimensions and all constraints related to train, wagon and bogie weight limits as well as the bogie balancing are satisfied, the selected configurations and load unit assignments are final for both wagons.

For the local search in the direction of improving the length utilisation, two wagons are selected as follows: the first wagon is chosen randomly from all wagons for which not all loading length is utilised; the second wagon is selected purely random from all available wagons. Finally, the local search procedure is terminated if a maximum number of sequential iterations without improvement $n_{it(LS_k)}$, or a solution with the maximum possible length utilisation is reached. The search in the direction of improving destination preferences and priority scores is guided by a local search which consists of a similar configuration change operator. However, both neighbourhoods are defined by selecting a new configuration for two randomly selected wagons.

6.2.3.2 Evaluation of the solutions

While other MDLS-approaches use pure single-objective local search procedures, the proposed operators are guided by a normalised, weighted-sum objective function which takes into account all three objectives and showed good performance after preliminary results. This function assigns a weight w_k to the primary objective k , while the remaining objectives r_1 and r_2 each receive a weight of $w_r = \frac{1-w_k}{2}$ (with $w_r \ll w_k$) in order to avoid a large negative change in these remaining objectives. The weight of the remaining objectives should be set small relative to w_k such that non-dominated solutions can be obtained. New solutions with normalised objective function values nf_{new} (6.1) are compared with the current best solution with normalised objective values nf_{best} (6.2).

$$nf_{new} = w_k \cdot nf'_k + w_r \cdot (nf'_{r1} + nf'_{r2}) \quad (6.1)$$

$$nf_{best} = w_k \cdot (nf_k + A_{LS_k}) + w_r \cdot (nf_{r1} + nf_{r2}) \quad (6.2)$$

The possibility of accepting solutions with slightly worse, equal or better weighted objective function values is defined by A_{LS_k} , indicating the acceptance threshold as the absolute change allowed in each objective k for accepting new solutions. By accepting worse solutions (i.e., a negative absolute change), a larger neighbourhood may be explored in comparison with accepting only improving solutions in order to obtain a more diverse solution pool. It could be required to aim at entirely new configurations in order to find new solutions which satisfy axle weight constraints and payload limits and, at the same time, result in better solutions. On the other hand, for destination preference scores, it can be expected that small changes might be satisfactory to obtain improvements. Concrete criteria for each local search LS_k with regard to the evaluation of the solutions (defined by A_{LS_k}) are examined in the tuning phase (Section 6.3). The local search with a main focus on destination preferences evaluates solutions with $nf_{new} > nf_{best}$, whereas solutions within local searches for length utilisation and priority scores are evaluated using $nf_{new} \geq nf_{best}$.

Further, a temporary set of non-dominated solutions T is updated with new solutions within one local search. If a solution is non-dominated by the solutions in the temporary set, it is added to this set, while dominated solutions are removed. Working with this temporary solution set avoids updating solution set E too often when new solutions are found within one local search, especially because one local search primarily focuses on one objective only. A single local search may generate multiple non-dominated solutions, and solutions found at the start of a local search are often

dominated by solutions found in further iterations. Finally, the local search ends with updating the global archive of non-dominated solutions E with the temporary set of solutions T obtained in LS_k .

6.2.4 Quality indicators

In order to assess the performance of the MDLS, two complementary quality indicators, the hypervolume indicator and the multiplicative epsilon indicator, are used to quantify the quality of an approximation set A obtained by the heuristic relative to a reference set R . The reference set can be either the exact Pareto-front or, in case the entire exact Pareto-front is unknown, a set which consists of all non-dominated solutions found by the heuristic algorithm and the exact method over all available solutions for all runs (Knowles et al., 2006).

The hypervolume indicator I_H measures the hypervolume covered by a set of solutions relative to a reference point. The multiplicative epsilon indicator I_ϵ determines the factor by which each point in an approximation set should be multiplied such that the reference set is weakly dominated by the approximation set (Knowles et al., 2006). Zitzler et al. (2003) perform the calculation for a minimisation problem. It can be adapted to a maximisation problem as follows for two solutions sets A and R , with n objectives:

$$I_\epsilon(A, R) = \min_{z^2 \in R} \max_{z^1 \in A} \min_{1 \leq i \leq n} \frac{z_i^1}{z_i^2}$$

After calculating the ratio of each objective for all solutions in A and R , the minimum ratio obtained out of all objectives n is kept in order to compare each solution in the approximation set with the reference set. When taking the maximum value for all solutions in A for each solution in R , and next the minimum value out of all solutions in R , the epsilon indicator is obtained.

Both indicators are visualised in Figure 6.2 for a bi-objective maximisation problem. The left-hand side shows the hypervolumes covered by a reference set and an approximation set. On the right-hand side, the crosses indicate the weakly dominated set obtained by the epsilon indicator.

The hypervolume indicator I_H is defined as the hypervolume of the reference set which is covered by the approximation set generated by the MDLS ($\frac{I_{Ha}}{I_{Hr}}$). The indicator value is obtained using the algorithm provided by Beume and Rudolph (2006). Moreover, both indicators are calculated after normalisation of the objective values to ensure that each objective contributes approximately equally. The minimum reference point is determined based on the characteristics of the critical load units:

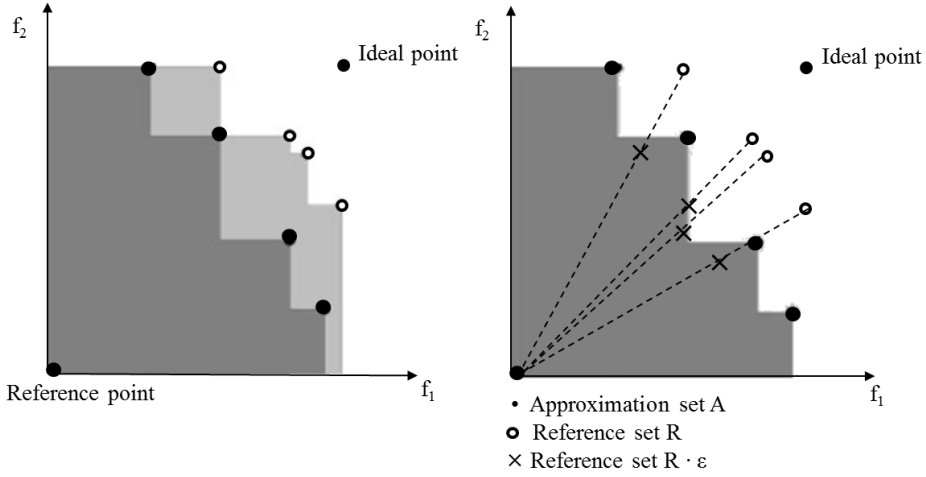


Figure 6.2: Quality indicators (adapted from Parragh et al. (2009)).

the urgency score and length of the critical load units of an instance minus one and a minimum destination preference score of zero.

The closer both indicators are to one, the better the quality of the approximation set. In order to compare the approximation sets obtained by two variants of a heuristic design, the two quality indicators can be used together, as each indicator measures slightly different information. Furthermore, if the indicators show opposite preference, the sets can be considered incomparable.

6.3 Parameter tuning

Irace (López-Ibáñez et al., 2016), a promising iterated racing procedure for tuning algorithm parameters, is used in order to find a parameter setting which leads to solutions of good quality, such that the heuristic can be used as a decision support tool by practitioners. First, the characteristics of the tuning instances (Section 6.3.1) and considered MDLS-parameters (Section 6.3.2) are presented. Next, the iterated racing procedure for the multi-objective train load planning problem is described and the parameters are tuned (Section 6.3.3).

6.3.1 Instance classes

A heterogeneous set of instance classes is used as input, generated based on historical data of an intermodal company, using the same classes as the problem instances in Section 5.4. Sets of load units with weights (in tonnes) being either light with $TRIA(17, 20, 23)$, heavy with $TRIA(23, 26, 29)$ or uniformly distributed with $UNIF(17, 29)$ are considered. Furthermore, the number of critical load units is varied and can be 35% or 20% of the total amount of available load units. These characteristics are representative for real-life problems. The six instance classes are applied to a wagon set of 5, 10 and 20 wagons, resulting in 18 classes, where 20-wagon instances are realistically sized. For the parameter tuning phase, two instances per class are used.

6.3.2 Parameters of the MDLS

Two important parameters for the MDLS are the number of times a new solution is selected from the pool of non-dominated solutions ($n_{it(MDLS)}$), and the number of consecutive iterations without improvement after which the local search phase in the direction of each objective ends ($n_{it(LS_k)}$). Clearly, a trade-off between the values of these parameters can be expected if a limited computation time is available. In this section, no limit on the computation time is considered, but Section 6.4 examines the relationship between solution quality and computation time for different amounts of MDLS-iterations and LS-iterations.

Within the scope of a single local search LS_k , the weight w_k attached to the main objective k of a local search should be tuned carefully. Moreover, within the normalised, weighted objective function used for accepting new solutions, possible values for A_{LS_k} are evaluated. Independent of this, only non-dominated, accepted solutions are added to the solution pool. Finally, it is tested whether the local search in the direction of improving destination preferences performs better when changing the configuration of two wagons with different destinations (i.e., $DP_{wgn\text{destin}} = 1$).

Table 6.1 presents an overview of the heuristic parameters under consideration, as well as their considered range or set of values.

The acceptance criterion in the local search for the length utilisation ($A_{LS_{Length}}$) considers only a set of specified ordinal values (o), while for the other parameters values between the specified minimum and maximum bounds can be selected by the racing procedure. The weight attached to the principal objective w_k is a continuous value with a precision of two decimals (r), all other parameter values are integers (i). The ranges of all parameters are defined intuitively after preliminary testing.

Table 6.1: MDLS algorithm parameters for tuning.

Parameter	Description	Range	Type
$n_{it(MDLS)}$	Number of times a random solution is selected	(10, 1000)	i
$n_{it(LS_{Length})}$	Number of consecutive non-improving iterations for LS_{Length}	(100, 2000)	i
$n_{it(LS_{Urgency})}$	Number of consecutive non-improving iterations for $LS_{Urgency}$	(100, 2000)	i
$n_{it(LS_{DP})}$	Number of consecutive non-improving iterations for LS_{DP}	(100, 2000)	i
w_k	Weight attached to the main objective k in LS_k	(0.8, 1)	r
$A_{LS_{Length}}$	Accepted length difference (in feet) in LS_{Length}	{-45;-20;0;20}	o
$A_{LS_{Urgency}}$	Accepted difference in priority scores in $LS_{Urgency}$	(-2, 2)	i
$A_{LS_{DP}}$	Accepted difference in destination preference scores in LS_{DP}	(-2, 2)	i
$DP_{wgn\text{destin}}$	Select two wagons with different destinations in LS_{DP}	(0, 1)	i

For the allowed absolute deviation from the current best solution in each local search, A_{LS_k} , a negative value indicates the acceptance of worse solutions, whereas a positive value relates to accepting improvements only. Depicted values of $A_{LS_{Length}}$ vary from a length difference (in feet) of 45 feet worse (with $A_{LS_{Length}} = -45$) – which relates to a proportional difference of approximately 12.5%, 6.25% or 3.125% of the available train length for instances with 5, 10 and 20 wagons respectively – until a length increase of 20 feet (i.e., $A_{LS_{Length}} = 20$). One additional unit in the considered value for $A_{LS_{Urgency}}$ and $A_{LS_{DP}}$ corresponds respectively to a single urgent load unit or a load unit with a weak destination preference score additionally assigned to a location on the train.

6.3.3 Iterated racing procedure (*irace*)

The iterated racing procedure (Algorithm 5) is able to automatically configure algorithms, providing a set of parameter values which performs well for a particular problem (López-Ibáñez et al., 2016).

Algorithm 5 Iterated Racing Procedure (López-Ibáñez et al. (2016)).

Require: $I = \{I_1, I_2, \dots\} \sim \mathcal{I}$,
parameter space: X ,
cost measure: $C(\theta, i) \in \mathbb{R}$,
tuning budget: B

- 1: $\Theta_1 \sim \text{SampleUniform}(X)$
- 2: $\Theta^{elite} := \text{Race}(\Theta_1, B_1)$
- 3: $j := 1$
- 4: **while** $B_{used} \leq B$ **do**
- 5: $j := j + 1$
- 6: $\Theta^{new} \sim \text{Sample}(X, \Theta^{elite})$
- 7: $\Theta_j := \Theta^{new} \cup \Theta^{elite}$
- 8: $\Theta^{elite} := \text{Race}(\Theta_j, B_j)$
- 9: **end while**
- 10: **Output:** Θ^{elite}

The racing procedure starts with T_{first} instances on which a number of uniformly sampled candidate parameter configurations are tested. After these T_{first} tested instances, the candidate configurations which perform worse than at least one other configuration – calculated by a statistical Friedman test – are discarded (lines 1–2). The best configurations (i.e., the configurations with the best objective values) are

selected as an elite set, and new configurations are added for the following race based on well-performing parent elite configurations found so far (lines 5–8). In the next iterations or races, each time T_{each} instances are evaluated before discarding any configuration. Furthermore, the standard deviation is reduced for each parameter as the number of iterations increases in order to search closer around better performing values. The procedure is terminated if a predefined computational budget B is reached. This budget corresponds to a maximum number of experiments, where one experiment consists of one parameter configuration tested on a single instance.

As multiple objectives must be considered, the cost function is represented by the quality indicator value, which should be maximised. López-Ibáñez et al. (2016) tested *irace* for their problem with multiple objectives using the hypervolume and the epsilon indicator and could not find significant differences. Therefore, only the hypervolume indicator is used as measure of the solution quality at this stage. The calculation of this quality indicator requires a reference set, which can be the exact Pareto-front. If not all Pareto-optimal solutions are known, the reference set consists of all non-dominated solutions found by a number of MDLS-runs, combined with the non-dominated solutions found so far in the exact procedure. Therefore, all candidate parameter configurations in a single iteration are tested on one instance before evaluating the cost function, i.e., calculating the quality indicator. Normalisation bounds and the reference point can be calculated in advance, independent of the approximation sets found by the heuristic.

The total set of tuning instances consists of two blocks of 18 instances, with a representative set of characteristics. The total amount of 36 instances, containing two instances from each instance class, is first tested before discarding any candidate configuration. In this way, two instances of every class are evaluated before a first elimination occurs, to cope with a possible outlier instance. Next, after every block of 18 instances (such that each instance class is represented), the configurations under consideration are again evaluated. Sampling of instances does not occur randomly, but in the order of the instance classes within one block in order to avoid elite configurations being biased towards only a subset of the instance classes.

Table 6.2 provides the adapted *irace* parameters used. All other parameters are at their default values.

The best configurations presented by *irace* are summarised in Table 6.3. These configurations are ordered according to their mean performance, but do not show a statistically significant difference with respect to the solution quality. The average hypervolume indicator value of the best configuration across all considered instances amounts to 0.9996.

Table 6.2: *Irace* parameters.

<i>Irace</i> parameter	Value
Tuning budget B	5000
Cost measure C	Hypervolume
T_{first}	36
T_{each}	Random samples
Random samples	Off

Table 6.3: Best configurations obtained by *irace*.

Parameter		C_1	C_2	C_3	C_4
$n_{it(MDLS)}$		801	949	759	775
$n_{it(LS_{Length})}$		987	377	434	271
$n_{it(LS_{Urgency})}$		1826	1498	1745	1384
$n_{it(LS_{DP})}$		1716	1541	1388	1631
w_k		0.91	0.92	0.90	0.90
$A_{LS_{Length}}$	$(nf_{new} \geq nf_{best})$	0	0	0	0
$A_{LS_{Urgency}}$	$(nf_{new} \geq nf_{best})$	0	0	0	0
$A_{LS_{DP}}$	$(nf_{new} > nf_{best})$	0	0	0	0
$DP_{wgndestin}$		0	0	0	0

Relatively large differences exist with respect to the range of the number of iterations $n_{it(MDLS)}$ and $n_{it(LS_k)}$. This can be explained by the fact that no limit exists on the computation time, and thus, more iterations lead to a better heuristic performance. However, the difference in solution quality is not necessarily substantial for additional iterations.

The acceptance thresholds in the evaluation of the normalised, weighted objective functions are identical for all best configurations C_1 to C_4 . The parameter values for $A_{LS_{Length}}$ and $A_{LS_{Urgency}}$ within their respective local searches indicate that solutions are accepted if they are at least as good as the current best solution. The obtained parameter value for $A_{LS_{DP}}$ indicates that solutions are accepted only if they are better than the current best solution within a single local search.

This difference can be explained by the fact that the neighbourhood for the length utilisation and urgency scores should be larger in order to find solutions which differ more from previous solutions. Entirely new configurations or new combinations of load units may need to be established in order to find new, allowed combinations with respect to the payload limits which increase the length utilisation or priority scores. On the other hand, in order to improve destination preference scores it can be sufficient to switch the load unit assignment between two wagons (instead of interchanging assigned and non-assigned load units or recombining assigned load units), and keep this solution as current best solution to evaluate new solutions against.

However, as also indicated by parameter $DP_{wgn_{destin}} = 0$, at some moments the local search may benefit from a larger neighbourhood, selecting wagons purely random in order to further improve the current best solution, as opposed to selecting wagons of two different destinations.

One important limitation of *irace* is that the automatic algorithm configuration does not take into account computation times when selecting parameter configurations. Therefore, these results should be further tested and a sensitivity analysis is performed next on the parameters to analyse differences in computation times and solution quality for different parameter settings.

6.4 Parameter sensitivity analysis

As *irace* is used to find a suitable set of parameter values, no insights are obtained in the behaviour of the heuristic performance under varying values of the parameters. *Ir-ace* executes the tuning phase and presents the user with parameter values. However, no insights are presented on which combinations of parameters work well. Therefore, a

sensitivity analysis is performed. The best parameter configuration C_1 resulting from the *irace* tuning procedure is analysed to examine the influence of changes in these parameter values on solution quality and computation times. For this purpose, two new test instances per class are used, which are not considered in the tuning phase, resulting in 36 instances. Varying parameter values are tested on these instances with respect to differences in solution quality, expressed as a proportional deviation from the hypervolume of the reference set (I_{Hr}), as well as differences in computation time. The heuristic is implemented in C++. The experiments are conducted on an Intel Xeon E5-2680v3 processor at 2.5 GHz of the Flemish Supercomputer Center (VSC).

First, interactions between the number of times a new random solution is selected in the MDLS $n_{it(MDLS)}$ and the number of consecutive iterations without improvement $n_{it(LS_k)}$ after which each local search ends are considered (Section 6.4.1). Variations in parameter values w_k , A_{LS_k} and the possibility of selecting two wagons with different destinations ($DP_{wgn\text{destin}}$) in LS_{DP} are tested (Section 6.4.2). A discussion to determine the final parameter configuration is summarised in Section 6.4.3.

6.4.1 Interaction between $n_{it(MDLS)}$ and $n_{it(LS_k)}$

It can be expected that a higher number of iterations, $n_{it(MDLS)}$ and $n_{it(LS_k)}$, corresponds to a higher solution quality, but at the cost of larger computation times. Therefore, a point should be determined as from which additional gains in solution quality become small relative to the increase in computation time. The heuristic performance is evaluated for different parameter values. Values up to 1000 MDLS-iterations ($n_{it(MDLS)}$) are considered with steps of 200 iterations. As the largest gains may be obtained during the first MDLS-iterations, results for 10, 50 and 100 iterations are also evaluated. To examine its interaction with the number of consecutive iterations without improvement $n_{it(LS_k)}$ after which each local search is ended, for each local search LS_k separately multiples of 250 consecutive non-improving iterations are considered with a maximum of 2000 iterations.

Figure 6.3 shows the average proportional deviation from the hypervolume of the reference set when either varying the number of LS-iterations for destination preferences (DP), length utilisation (Length) or urgency scores (Urgency). The remaining parameters are set at the values of the best-performing configuration. Variations in the number of non-improving LS-iterations after which the local search with respect to the length utilisation ends do not significantly influence solution quality. This corresponds to the relatively small parameter values for $n_{it(LS_{Length})}$ in the best *irace* configurations. For the destination preference scores and urgency scores, lar-

ger differences can be observed for low numbers of MDLS-iterations. Clearly, major improvements with respect to the solution quality are reached during the first MDLS-iterations. These results are consistent with the best configurations found by *irace*, as $n_{it}(LS_{DP})$ and $n_{it}(LS_{Urgency})$ are always larger than 1250.

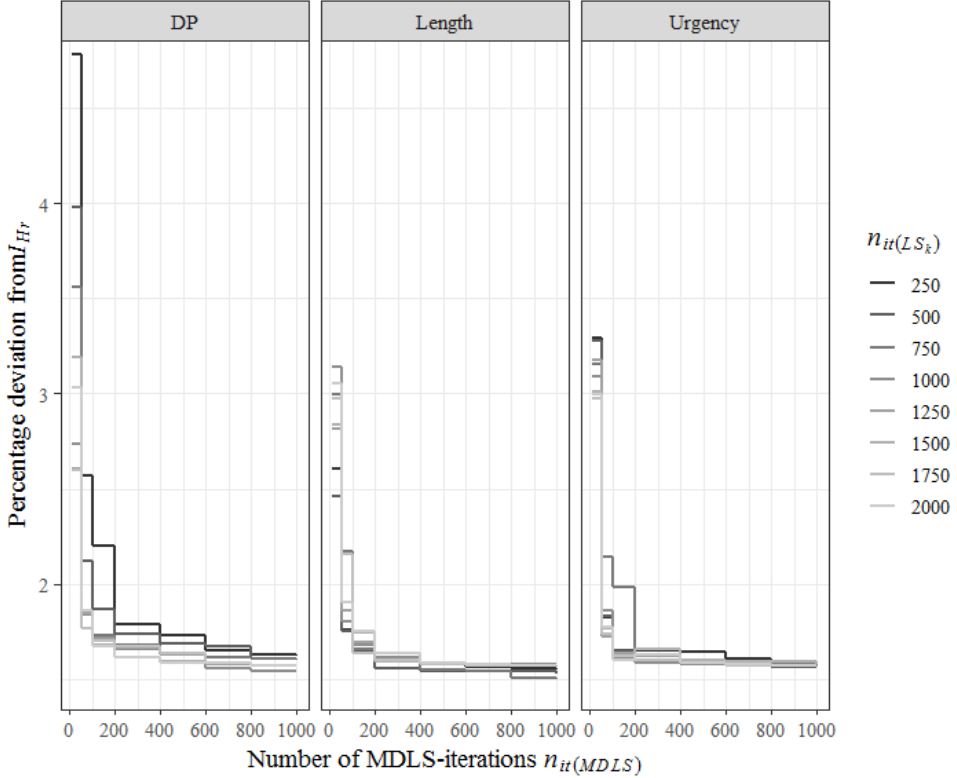


Figure 6.3: Solution quality based on $n_{it}(MDLS)$ and $n_{it}(LS_k)$.

Additionally, Figure 6.4 displays the average solution quality depending on the number of MDLS-iterations over all experiments, grouped per instance size. These results show that a clear difference exists with respect to the average solution quality: the heuristic performance is highest for instances with 10 and 20 wagons, while the mean performance is lower for 5 wagons. This may be due to the fact that one instance could be an outlier. The cause can be further investigated when analysing the impact of the remaining heuristic parameters on the solution quality in Section 6.4.2. Nevertheless, the mean deviation is still lower than 5%.

Generally, for low numbers of MDLS-iterations, differences in solution quality are

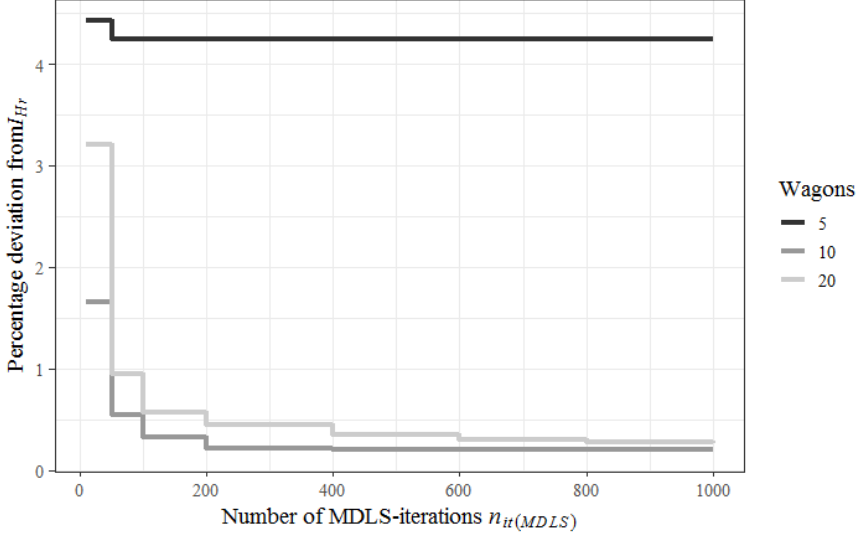


Figure 6.4: Average solution quality per instance size.

larger between a high and low number of non-improving local search iterations. As from 400 MDLS-iterations, the improvement in solution quality becomes relatively small, independent of the number of non-improving local search iterations, although performance improvements are still obtained after 400 iterations for the realistic instances consisting of 20 wagons. Although the best configurations obtained in the tuning phase indicate that at least 775 iterations should be performed, no limit on the computation time or trade-off between the solution quality and computational budget per instance is incorporated.

Figure 6.5 shows average computation times in seconds for all instances of each possible combination of MDLS- and LS-iterations. The number of MDLS-iterations mainly influences computation time, as it involves additional iterations for all three local searches at the same time. The difference in the slope of the destination preference graph compared to the length and urgency graphs indicates that each LS_{Length} and $LS_{Urgency}$ reached the number of consecutive iterations without improvement earlier, ending the local search. For LS_{DP} , this implies that more improvements are found at a later stage of the respective local search (i.e., after a larger number of consecutive iterations without improvement), obtaining a new best solution without ending the local search.

Computation time primarily increases with the number of MDLS-iterations, as

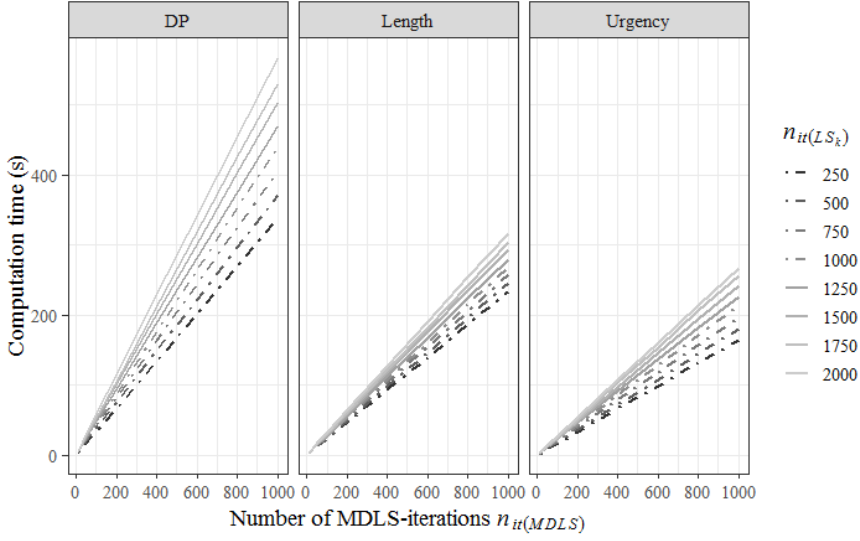


Figure 6.5: Average computation times based on $n_{it(MDLS)}$ and $n_{it(LS_k)}$.

one iteration consists of three distinct local search improvement phases. At a small number of MDLS-iterations, a higher number of LS-iterations is preferred, while the solution quality remains relatively stable beyond 400 MDLS-iterations, regardless of the number of LS-iterations. Beyond that point, either a small or a large number of LS-iterations produces a steady heuristic performance. The smaller the number of MDLS-iterations, the larger the impact of the number of consecutive iterations without improvement after which each local search is ended becomes. If computation time is a decisive factor, a smaller number of MDLS-iterations combined with a relatively high number of LS-iterations should be preferred.

For the best configurations C_1 to C_4 the number of iterations for LS_{Length} varies between 250 and 1000, and for $LS_{Urgency}$ and LS_{DP} between 1250 and 2000. For these intervals of LS-iterations per objective, the solution quality is high and stable as from 200 MDLS-iterations (Figure 6.3) with a substantially shorter computation time compared to 400 MDLS-iterations (Figure 6.5). Based on these observations, for the remainder of this sensitivity analysis we use $n_{it(MDLS)} = 200$, while $n_{it(LS_k)}$ remain at the best values found by *irace*. This configuration leads to high-quality solutions in relatively short computation times.

6.4.2 Sensitivity of w_k , A_{LS_k} and $DP_{wgndestin}$

For all other parameters under consideration, Table B.1 (Appendix B) shows an overview of the average solution quality obtained for each instance size as well as the overall average solution quality for each parameter value. For most of the parameters, the sensitivity analysis shows results identical to the *irace* parameter configurations. However, for $A_{LS_{Length}}$ the parameter value resulting in the overall average best result (indicated with an asterisk) does not correspond to the parameter values selected by *irace* (indicated in bold). This small deviation may be explained by the decision to work with 200 MDLS-iterations instead of a parameter value out of one of the best configurations as well as by the fact that different instances are used for the sensitivity analysis. In the remainder of this section, individual results for each parameter are discussed in detail.

With respect to the **weight** attached to the main objective of a single local search, w_k , values between 0.6 and 1 are tested with an interval of 0.05. The parameter value of 0.91 obtained in the best configuration C_1 is also added. The solution quality for each parameter value and each instance size (5, 10 and 20 wagons) is displayed in Figure 6.6. Results show that a weighted objective function can be beneficial in comparison with a function which only attaches weight to the main objective of the local search. Based on the results of the test instances used, for small instances with 5 wagons, the weight could be set to either 0.91 or 0.95, both leading to a high solution quality. For other values of the parameter, one instance severely impacts the average performance. For the larger and realistic instances with 10 and 20 wagons, differences in solution quality are smaller. Generally, a weight of 0.91 provides the highest average solution quality and the most favourable minimum guaranteed solution quality.

Similarly, the influence of the criterion for **accepting solutions** in each local search LS_k for each objective k is evaluated. Figure 6.7 shows that the best parameter values for $A_{LS_{Length}}$ are consistent with the *irace* results for the considered test instances of 10 and 20 wagons. However, a difference in solution quality exists between a small instance size of 5 wagons and larger instance sizes. The performance for instances with 10 and 20 wagons does not substantially differ for varying parameter values. Accepting worse solutions clearly results in an improved performance for 5-wagon instances, as is obvious for these smaller instances. Therefore, the acceptance of worse solutions (with a length of 20 or 45 feet less than the current best solution) could provide a higher overall solution quality for these test instances.

With respect to $A_{LS_{Urgency}}$, the overall average solution quality (Table B.1, Appendix B) is the highest at $A_{LS_{Urgency}} = 0$, which is compatible with the *irace* results.

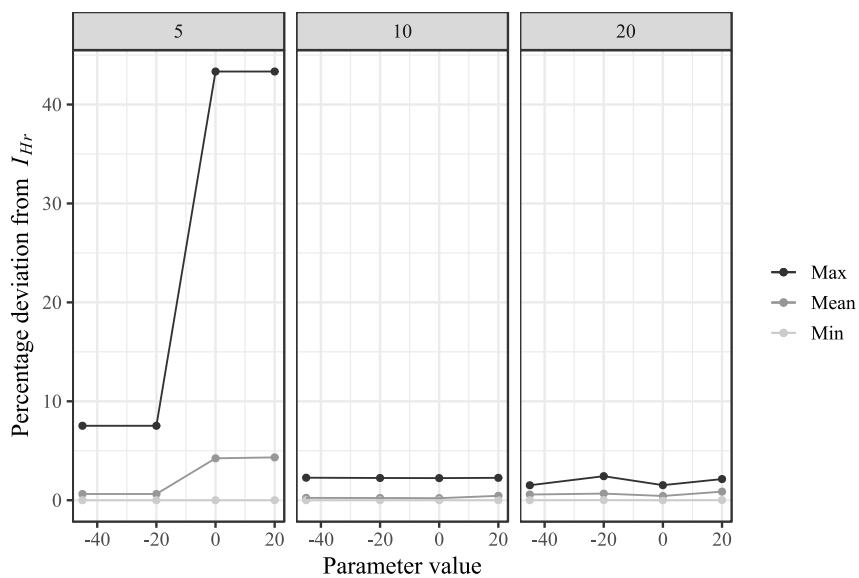
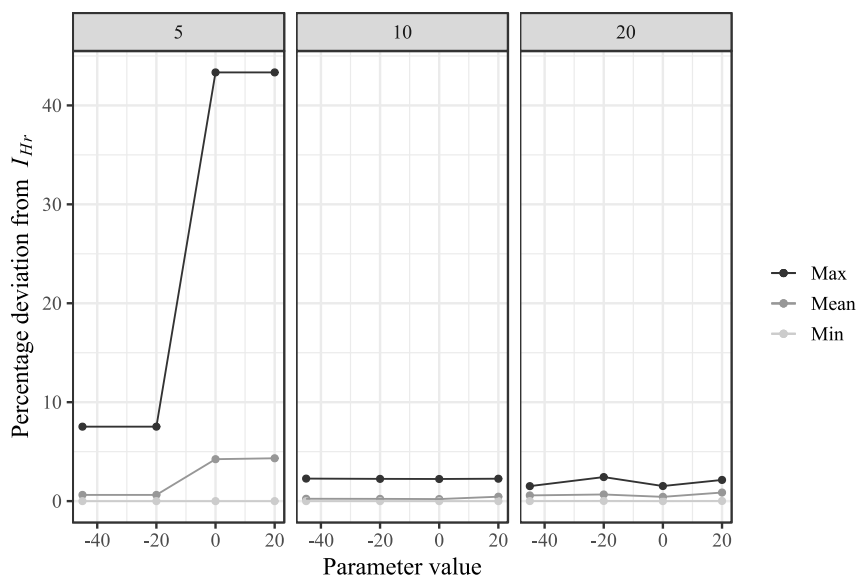
Figure 6.6: Sensitivity of w_k on the solution quality.Figure 6.7: Sensitivity of $A_{LS_{Length}}$ on the solution quality.

Figure 6.8 depicts a difference in solution quality for small instances in comparison with the instances of 10 and 20 wagons. For the latter instance categories, the heuristic performance is the highest at $A_{LS_{Urgency}} = 0$, and minimum and maximum solution quality are extremely close, whereas small instances benefit from the acceptance of worse solutions.

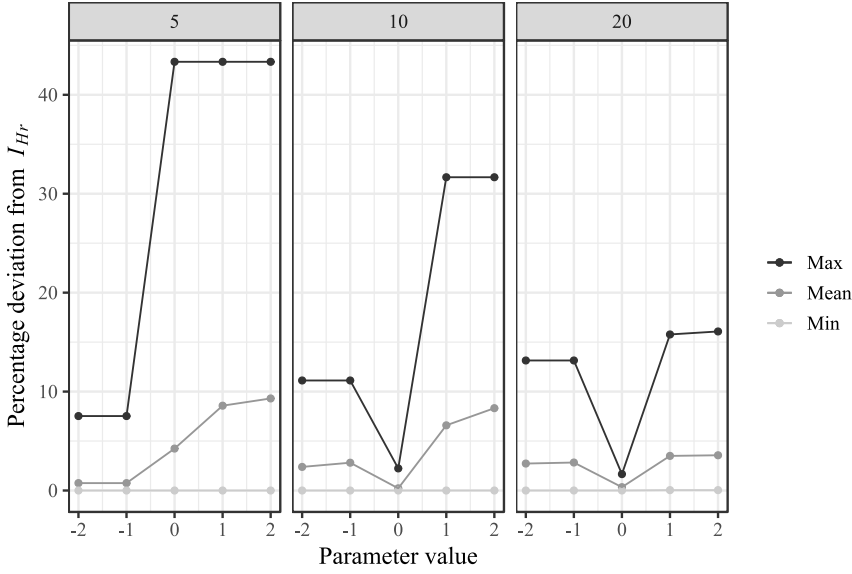


Figure 6.8: Sensitivity of $A_{LS_{Urgency}}$ on the solution quality.

Figure 6.9 displays a pattern similar to $A_{LS_{Urgency}}$ for possible values of $A_{LS_{DP}}$ with respect to the solution quality. Although instances with 5 wagons perform worse if $A_{LS_{Urgency}} = 0$, the overall performance is highest.

Finally, a parameter $DP_{wgndestin}$ is added and evaluated in order to test whether in the local search focusing on destination preference scores, selecting two wagons with different **destinations** ($DP_{wgndestin} = 1$) leads to a higher solution quality in comparison with two random wagons ($DP_{wgndestin} = 0$). As shown in Figure 6.10, no substantial difference exists with respect to the solution quality. Therefore, selecting two wagons with different destinations does not add value to the heuristic.

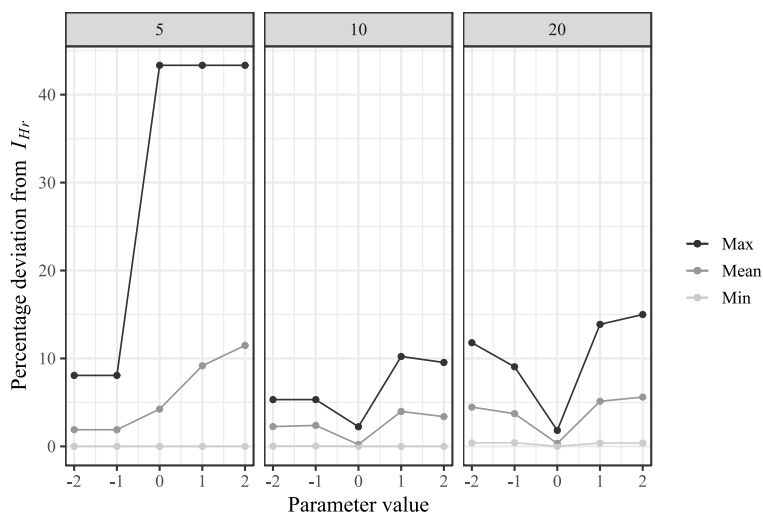


Figure 6.9: Sensitivity of $A_{LS_{DP}}$ on the solution quality.

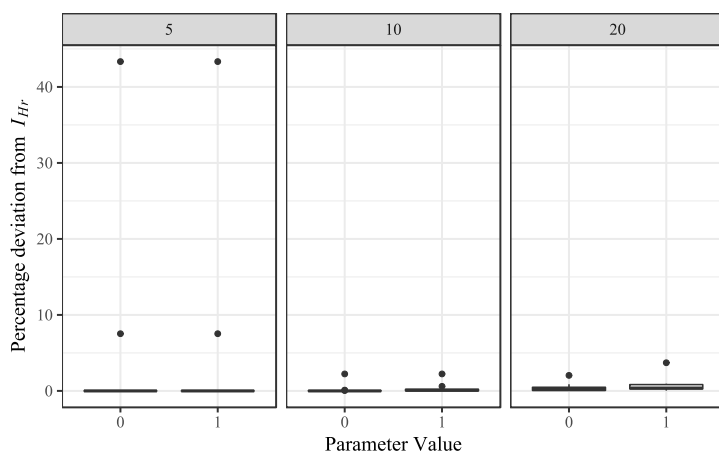


Figure 6.10: Sensitivity of $DP_{wgndestin}$ on the solution quality.

6.4.3 Practical implications

Table 6.4 summarises the results of *irace* and the sensitivity analysis. Generally, the parameter configuration C_1 found by *irace* using 36 tuning instances provides high-quality results for the test instances as well, even for a smaller number of MDLS-iterations. The 36 new test instances used in the sensitivity analysis show a heuristic performance consistent with the results obtained by *irace*. However, one possible outlier test instance in the instance category of 5 wagons might have influenced the results of the sensitivity analysis. For that specific category a different configuration may be more suitable. Changing $A_{LS_{Length}}$ to accepting solutions with a length utilisation of 20 feet worse ($A_{LS_{Length}} = -20$) than the current best solution so far leads to the exploration of a larger search area, escaping from local optima, and contributes to an enhanced performance for this specific instance size.

Table 6.4: Recommended parameter values based on *irace* and sensitivity analysis.

Parameter	<i>irace</i>	Sensitivity Analysis	Recommendation
$n_{it(MDLS)}$	801	200	≤ 200
$n_{it(LS_{Length})}$	987	987	250-1000
$n_{it(LS_{Urgency})}$	1826	1826	1250-2000
$n_{it(LS_{DP})}$	1716	1716	1250-2000
w_k	0.91	0.91	0.91
$A_{LS_{Length}}$	0	0	-20 or 0
$A_{LS_{Urgency}}$	0	0	0
$A_{LS_{DP}}$	0	0	0
$DP_{wgndestin}$	0	0	0

For practical applications good solutions should be obtained in short computation times. Based on the discussed results, train load plans obtained by the MDLS heuristic with a maximum of 200 MDLS-iterations will be valuable for practitioners, as the solution quality is high and stable as from this point. For a higher number of MDLS-iterations, improvements in solution quality are extremely small, while the computational burden increases. Therefore, we now investigate the relationship between the solution quality and computation time for a smaller number of MDLS-

iterations with a maximum of 200 iterations (i.e., 30, 100 and 200 iterations) and include the influence of parameter values $A_{LS_{Length}} = -20$ and $A_{LS_{Length}} = 0$ on the heuristic performance in order to obtain a parameter configuration to be used by practitioners. The number of iterations after which each local search is finished, is set at the minimum value for which the solution quality becomes stable, i.e., at 200 MDLS-iterations

Figures 6.11 and 6.12 respectively display the average deviation from the hypervolume of the reference set HV_R and computation times (in seconds) for each category of instances (5, 10 and 20 wagons; high, low or uniform weights of the load units; 35% and 20% of the available load units which are critical). All instances used in *irace* as well as in the sensitivity analysis are included, resulting in a total of four instances per category.

Figure 6.11 confirms that for small instances, temporarily allowing worse solutions ($A_{LS_{Length}} = -20$, indicated by the solid lines) clearly has a positive impact on the solution quality. Obtained solutions are either optimal or really close to optimality. By accepting worse solutions, we are able to escape from a local optimum, hereby increasing the solution quality. However, for realistic instances, the opposite holds. The solution quality is the highest if only solutions which are at least as good as the current best solution ($A_{LS_{Length}} = 0$, indicated by the dotted lines) are accepted for following iterations. Moreover, the number of MDLS-iterations does not influence the solution quality for small instances with 5 wagons, and 10-wagon instances show only minor differences with a maximum deviation of less than 2% from the optimal solution. Differences in parameter values for $n_{it(MDLS)}$ become important for 20-wagon realistic instances. The solution quality is definitely inferior for $n_{it(MDLS)}=30$, while results for 100 and 200 iterations are similar.

Figure 6.12 compares computation times for the exact solution method with different heuristic parameter values for the number of MDLS-iterations $n_{it(MDLS)}$ and for the acceptance criterion within the local search focused on the length utilisation $A_{LS_{Length}}$. No substantial difference in computation time is demonstrated by varying the acceptance criterion $A_{LS_{Length}}$. Each value of $n_{it(MDLS)}$ (indicated by the line colours), displays a similar pattern, while a slightly larger computation time is observed for instances with low load unit weights for most instance categories. The added value of the heuristic is demonstrated for realistic instances with 20 wagons. Heuristic computation times are below 100 seconds, while the exact solution method requires much larger computation times.

Without loss of generality, when focusing on solving realistic instances, the best results are obtained if $A_{LS_{Length}} = 0$ for all load unit weights and distributions of

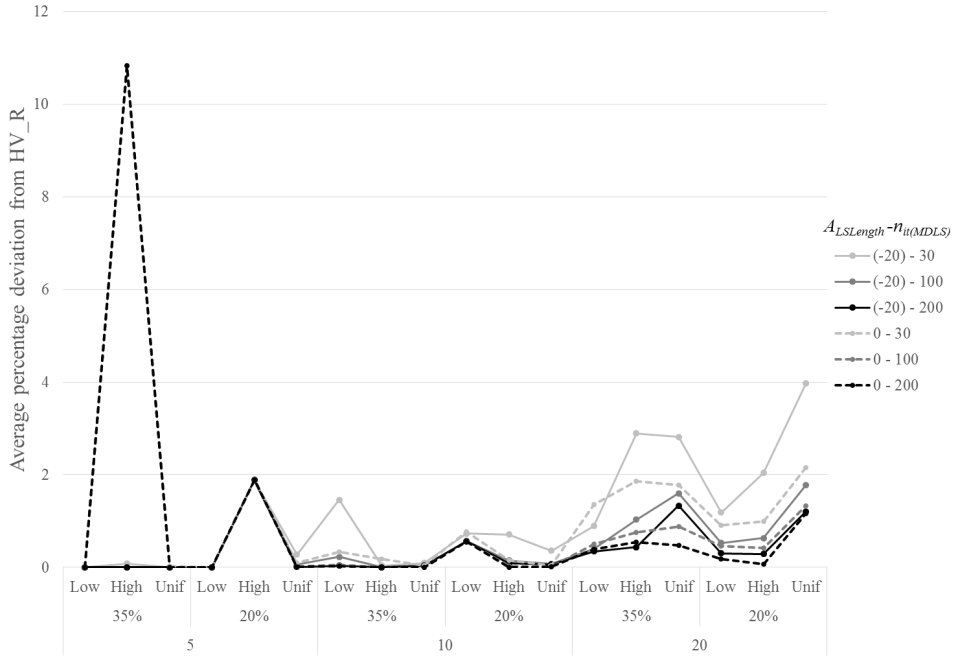


Figure 6.11: Average solution quality per category of instances for different combinations of $A_{LSLength}$ and $n_{it}(MDLS)$.

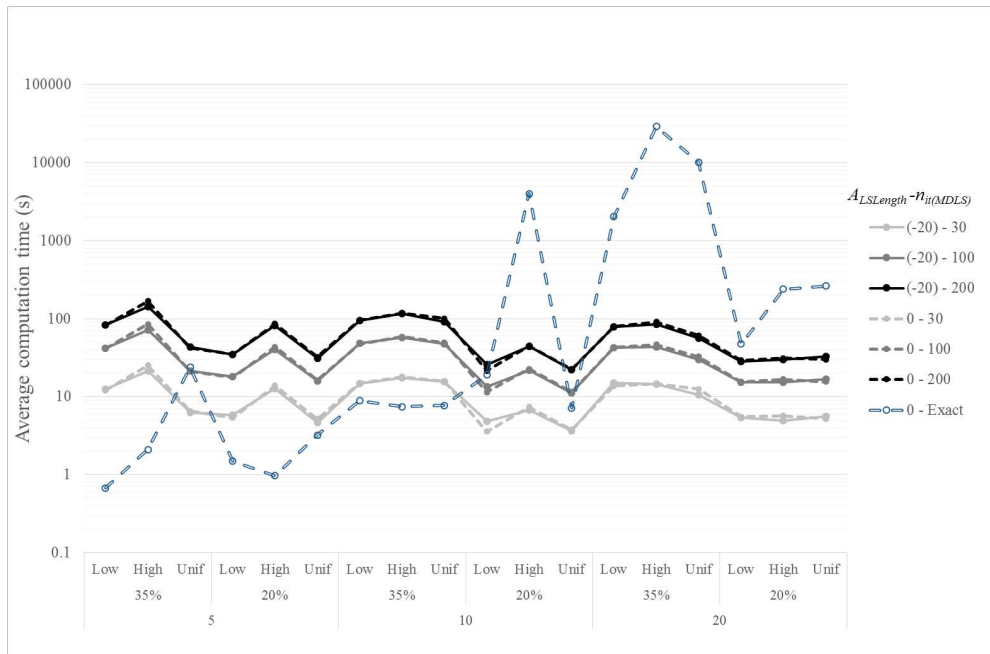


Figure 6.12: Average computation times per category of instances for different combinations of $A_{LSLength}$ and $n_{it}(MDLS)$.

the urgency of available load units, accepting only solutions at least as good as the current best solution. As the solution quality is slightly higher for a larger number of times a new random solution is selected, $n_{it(MDLS)}$ is set to 200 iterations. This parameter setting provides the best results for realistic instances, while the solution quality is also high for all other instance sizes.

6.4.4 Conclusions

In this section, we obtained a configuration which generally performs efficiently for a heterogeneous set of instance classes. With this parameter configuration, the multi-directional local search heuristic is able to find solutions of high quality within a reasonable amount of computation time. One limitation of this research is that only the hypervolume is used as a performance indicator. The obtained results may be validated with the results of the epsilon indicator, although it is expected that similar results would be obtained. Further, the considered instances are heterogeneous with respect to their characteristics and maybe different parameter configurations would be selected if each category would be considered separately. More specifically, a distinct parameter configuration could be obtained for each instance size. Further research may focus on finding specific configurations for each category of instances, depending on the intended use of the heuristic.

Moreover, while the interaction between the number of MDLS-iterations and the number of LS-iterations is investigated, no interaction effects are studied with respect to the parameters for w_k , A_{LS_k} , $DP_{wgndestin}$. Due to the relatively high number of iterations, the effect of the acceptance criterion included in the weighted objective function may be underestimated. Therefore, one opportunity lies in investigating the interaction between the various acceptance thresholds A_{LS_k} and the number of LS- and MDLS-iterations.

6.5 Performance of the multi-directional local search heuristic

In order to analyse the performance of the proposed multi-directional search heuristic, its results are compared with results of the exact ϵ -constraint method. Both the hypervolume and the epsilon indicator are used to evaluate the quality of the approximation sets found by the proposed heuristic. The heuristic results in this section are obtained using the heuristic parameter setting summarised in Table 6.5.

Table 6.5: Best parameter setting for the multi-directional local search heuristic for the train load planning problem.

Parameter	Definition	Value
$n_{it(MDLS)}$	Number of times a new random solution is selected	200
$n_{it(LS_{Length})}$	Number of non-improving local search iterations for f_1	250
$n_{it(LS_{Urgency})}$	Number of non-improving local search iterations for f_2	1250
$n_{it(LS_{DP})}$	Number of non-improving local search iterations for f_3	1250
w_k	Weight attached to the prevailing objective k in LS_k	0.91
w_r	Weight attached to the remaining objectives in LS_k	0.045
$A_{LS_{Length}}$	Accepted length difference (in feet) in LS_{Length}	0
$A_{LS_{Urgency}}$	Accepted priority difference (in feet) in $LS_{Urgency}$	0
$A_{LS_{DP}}$	Accepted difference in destination preference score in LS_{DP}	0
$DP_{wgn_{destin}}$	Selection of two wagons with different destinations in LS_{DP}	No

In Figure 6.13 and Table 6.6, exact computational results are compared with the average performance of the MDLS-heuristic for the instances used in Chapter 5, ten instances per class, with ten runs per instance in order to analyse the solution quality of the approximation sets obtained by the heuristic approach. In this table, the hypervolume indicator I_H is stated as the hypervolume of the reference set, i.e., the exact Pareto-front, which is covered by the approximation set generated by the MDLS ($I_H = I_{H_a}/I_{H_r}$). Each category consists of 10 instances with 10 runs of the heuristic algorithm per instance, and 100 times is the maximum number of times the entire Pareto-front can be found. The number of times this front is found, is indicated by $\#PF_{found}$.

Generally, the average epsilon and hypervolume indicator values are close to one for all considered instances, which indicates that on average, solutions of good quality are found by the MDLS. With respect to the small instances with five wagons, the heuristic is even able to find the exact Pareto-front in a large number of runs in each problem category. Moreover, both quality indicators show a similar pattern. However, the solution quality slightly decreases if the problem size increases, as shown by both quality indicators.

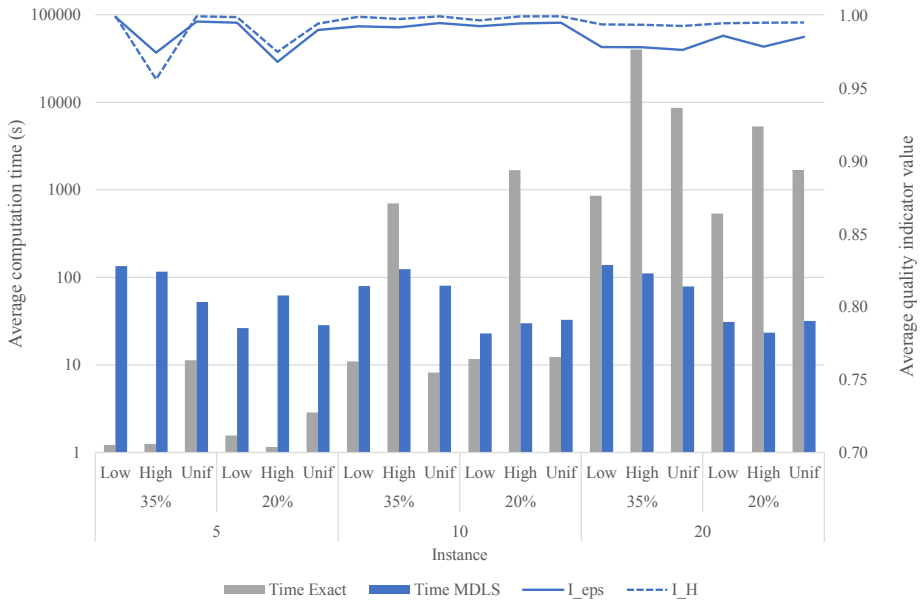


Figure 6.13: Average solution quality and time per category of instances for the best parameter setting.

Table 6.6: Comparison of average results for the exact ϵ -constraint method and the MDLS-heuristic.

Problem instance	Instance	Exact		MDLS				
		Time	Sol	Time	Sol	I_ϵ	I_H	#PF found
(5, 35%,light)	1-10	1.22	5.3	134.23	5.3	0.999432	0.999983	98
(5, 35%,heavy)	11-20	1.25	4.4	116.24	4.2	0.975000	0.956664	90
(5, 35%,uniform)	21-30	11.31	4.9	52.43	4.8	0.996319	0.999923	82
(5, 20%,light)	31-40	1.56	5.8	26.33	5.7	0.995575	0.999311	90
(5, 20%,heavy)	41-50	1.15	3.7	62.23	3.6	0.968579	0.975409	80
(5, 20%,uniform)	51-60	2.86	6.5	28.47	6.3	0.990463	0.994861	70
(10, 35%,light)	61-70	10.98	11.8	79.53	11.9	0.993013	0.999610	63
(10, 35%,heavy)	71-80	699.93	8.5	124.09	8.9	0.992400	0.997966	75
(10, 35%,uniform)	81-90	8.18	9.8	80.31	9.4	0.995184	0.999926	53
(10, 20%,light)	91-100	11.69	7.7	22.94	8.0	0.993118	0.997094	73
(10, 20%,heavy)	101-110	1677.01	10.3	29.91	9.8	0.994947	0.999856	57
(10, 20%,uniform)	111-120	12.29	11.0	32.71	10.5	0.995447	0.999816	62
(20, 35%,light)	121-130	855.77	18.0	138.86	15.4	0.978773	0.994388	18
(20, 35%,heavy)	131-140	39917.96	29.3	111.27	25.6	0.978619	0.994002	1
(20, 35%,uniform)	141-150	8614.75	29.7	78.66	25.7	0.976734	0.993225	0
(20, 20%,light)	151-160	536.24	12.6	31.04	11.7	0.986570	0.995180	8
(20, 20%,heavy)	161-170	5283.01	16.7	23.35	14.0	0.978968	0.995577	1
(10, 20%,uniform)	171-180	1680.34	17.7	31.77	14.9	0.985799	0.995610	7

With respect to the computation times (in seconds), computation times of the heuristic for instance categories with a small number of critical load units are clearly shorter than for instances with a large number of critical load units. For small instances, heuristic run times are large compared to the exact ϵ -constraint method. However, for realistic instances, the advantages of the MDLS-heuristic are obvious. Furthermore, categories of instances which show extreme computation times by the exact ϵ -constraint method are solved in a much shorter computation time by the heuristic solution method, which shows the added value of a heuristic solution method for the multi-objective train load planning problem.

High-quality approximation sets for realistic instances with 20 available wagons can be found by the MDLS within less than three minutes on average, which is consistently smaller than the run time for finding the exact Pareto-front. These approximation sets cover on average at least 99% of the hypervolume of the reference sets, as shown by I_H . In conclusion, the results indicate that the MDLS-heuristic can be used in a realistic environment. It guarantees solutions of good quality in short times, which can be critical for train planners in a real-time operational environment.

6.6 Trade-off analysis

To show the added value of the multi-objective approach, results are visualised for two realistic instances. Figure 6.14 and 6.15 respectively show objectives f_1 and f_3 , loaded length and scores for destination preferences, on the x- and y-axes. Objective f_2 , the number of urgent load units assigned, is shown by adding sized and colour-scaled dots. First, solutions are shown for an instance with 35% critical load units and load unit weights drawn from a uniform distribution. The Pareto-front for this instance consists of 33 different solutions, visualised in Figure 6.14.

Clearly, trade-offs can be observed for this instance. The higher the number of urgent load units assigned, the lower the destination preference scores. Consequently, the graph shows a clear decrease in destination preference scores from points with small dots of a light colour to larger and darker dots. Further, when 11 or 12 urgent load units are assigned, planners have a large flexibility, with a larger range of possible load plans, going from plans with a high length utilisation but low destination preference scores to plans with higher destination preference scores but less of the available length used. With either a high number of urgent load units assigned or high destination preference scores, the possibilities to compose a load plan become smaller. On the other hand, if length utilisation is important at a specific moment,

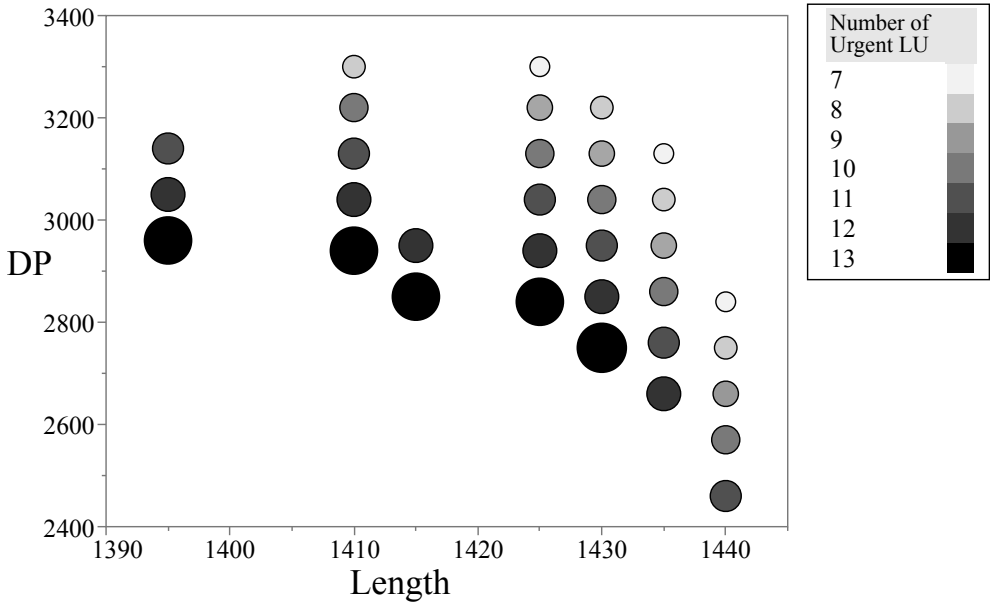


Figure 6.14: Trade-offs for a realistic instance with 35% critical load units and a uniform weight distribution.

planners still have flexibility selecting a plan with either higher destination preference scores or a high number of urgent load units.

The obtained solutions for a second instance are shown in Figure 6.15. It has the same characteristics, except for the fact that only 20% of the available load units is critical. This instance is further characterised by 20 different Pareto-optimal solutions, and trade-offs are very clear.

For this instance, differences between the number of urgent load units assigned are more explicit. The dotted markers for 5 and 6 urgent load units are relatively close to each other, which indicates that in this case, up to 6 load units can be assigned without compromising too much with respect to the other objectives, destination preference scores and length utilisation. However, the more urgent load units are assigned, the larger the decrease in the other objectives becomes. With a high number of urgent load units, planners should decide whether, at a specific moment, either the length utilisation or assigning load units to a more preferred unload terminal is more important. With a small number of urgent load units, planners can at the same time optimise destination preference scores and length utilisation.

One possibility for future research may be the selection of load plans from the pool

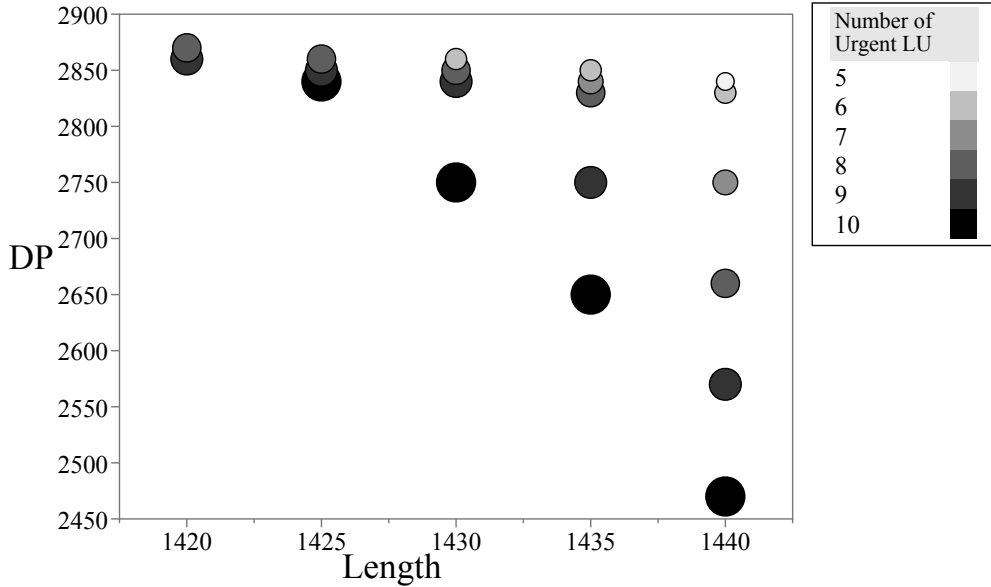


Figure 6.15: Trade-offs for a realistic instance with 20% critical load units and a uniform weight distribution.

of available alternative solutions generated by the multi-objective problem, based on knowledge of characteristics of load plans which are often chosen in certain operational circumstances. For a literature overview on multi-criteria classification and sorting methods to support decision makers in selecting the most suitable alternative, we refer to Zopounidis and Doumpos (2002).

In conclusion, results show that the multi-objective approach provides train load planners with a number of train load plans, such that they can decide a posteriori on the importance of each objective by selecting the most suitable load plan at that moment. Planners can narrow down the options based on information they receive about the prevailing transport conditions. They can optimise the scores for destination preferences or the number of urgent load units assigned to a location on the train without implying large changes with respect to the loaded length on the train. However, in some circumstances, these small differences in length utilisation may be essential.

6.7 Conclusions and future research

This chapter shows the added value of a multi-directional local search heuristic, incorporating real-life characteristics of train load planning, as with the exact ϵ -constraint method presented in Chapter 5, some single-objective problems involve long computation times. Train planners can benefit from the advantages of the proposed MDLS-heuristic, which provides them with a number of non-dominated solutions, i.e., possible train load plans. The MDLS is able to solve real-world instances and provides solutions of good quality in short computation times, which is valuable in real-life. It is able to provide fast assistance to train planners during the assignment process in a realistic, operational environment. The parameter tuning phase determined a parameter setting which performs well on average for all instance classes and thus for multiple possible inputs. For practical use, the detailed sensitivity analysis provides a good overview to select a good parameter setting for any specific problem setting.

The proposed multi-directional local search heuristic does not include a measure for the diversity of the solution pool. One such indicator concerns the crowding distance of a solution (Lian et al., 2016). It measures the density of solutions surrounding one solution. Including measures of the distance between any solution and all other obtained solutions allows to focus on less-crowded areas of the non-dominated solution set, which in turn may improve the diversity of the solution pool. It could be interesting to analyse the inclusion of distance measures on the heuristic performance.

Future research could focus on methods to narrow down the selection of load plans to provide decision support under the specific operational circumstances of a given moment. Furthermore, it would be interesting to extend the proposed heuristic to a dynamic setting, in which new information continuously becomes available, when the load plan must be updated accordingly.

A final interesting opportunity lies within the use of column generation to solve the train load planning problem. Both the axle weight limits and payload balancing are restrictive to the problem. The problem can be reformulated into a set partitioning problem, in which a list of feasible configurations of load units is created for each wagon type. These configurations should satisfy all loading constraints. For each wagon of a given type, exactly one configuration has to be selected and all load units should be assigned.

Part IV

Conclusions

Final conclusions and future outlook

This doctoral thesis aims at providing decision support for intermodal transport planners in order to optimally employ their service network and maximise transport throughput with minimised total transport costs. Two intermodal planning problems are considered, intermodal container routing and train load planning, in order to stimulate a modal shift from unimodal road transport to intermodal rail transport.

Chapters 2 and 3 (Part II) focus on integrating decisions in intermodal transport. In Chapter 2, related literature on the separate planning problems is presented and a new integrated intermodal routing problem is proposed, which combines drayage and long-haul routing decisions. In Chapter 3, a heuristic algorithm is developed to solve the integrated intermodal routing problem, quantify the advantages of the integrated approach, and provide managerial insights on how tactical decisions could influence the total cost of the transport system. Chapters 4 to 6 (Part III) investigate the train load planning problem with real-life characteristics. The train load planning problem is considered from a multi-objective perspective, enabling planners to maximise the throughput of transport requests with a given capacity during the train planning phase within the context of intermodal rail transport.

In this final chapter, the major conclusions of both main research areas studied are presented (Figure 7.1). Final conclusions are summarised in Section 7.1. Future research directions are presented in Section 7.2.

7.1 Final conclusions

With the rising focus towards sustainable transport, intermodal transport is a promising alternative for road transport and provides a lot of opportunities. However, it

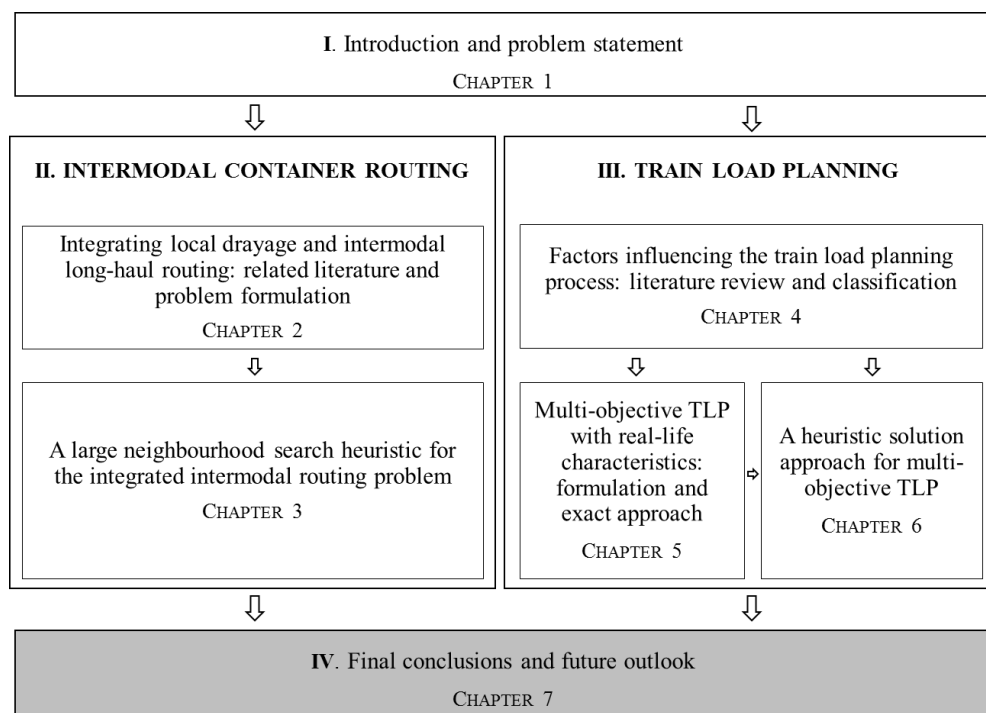


Figure 7.1: Thesis summary - Chapter 7.

also presents a number of challenges for intermodal operators and planners who are in charge of executing transport planning in an increasingly dynamic environment. As multiple transport modes and decision makers are involved, integrated decision-making tools are a prerequisite for adequate decision support. Due to low financial margins for transport companies, cost minimisation is a major objective. Although road transport is only used for drayage and presents a small portion of the intermodal trajectory, drayage costs constitute a substantial part of total transport costs. Furthermore, network capacity is expensive and should be managed carefully to maximise total transport throughput, which could further reduce total costs.

This dissertation proposes fast planning algorithms which include real-life characteristics and allow for planning flexibility in the decision process. These concepts are inherent to the vision of synchromodal transport in order to encourage a modal shift away from unimodal road transport. Two decision support tools for intermodal operators are proposed by means of heuristic algorithms for integrated intermodal routing and train load planning in order to execute intermodal transport operations both effectively and efficiently. They are applied to real-life cases in order to demonstrate their advantages for practitioners. Both problems aim at maximising network capacity utilisation and minimising total transport costs, only the time horizon differs. Support on intermodal routing decisions is provided for a weekly planning for which computation times can be higher. Decision support for train load planning is focused on daily operations and as such, requires a shorter computation time.

In the first part of this thesis, the focus is on providing integrated decision support for intermodal operators. Due to the complex planning environment in which intermodal service providers operate, decision within and between decision levels are interrelated and the need for an integrated approach rises. Although research on integrated decision support models in intermodal freight transport is limited, research efforts are growing. Therefore, a literature review on service network design, intermodal routing and drayage problems indicating the tackled problem characteristics in current research is presented. The review demonstrates that problem integration with respect to transport modes, decision levels and problem characteristics is required, but research efforts are still limited. The proposed, new integrated approach on intermodal routing includes more planning flexibility by allowing transport requests to be routed through another service than the long-haul service with the closest pre- and end-haul distance if this reduces total drayage costs. Moreover, the integrated viewpoint allows to use more information when making intermodal routing decisions compared to a sequential approach, resulting in better-informed decision processes. Results show that, although no total cost reduction is obtained for demand character-

ised by random customer locations, an increased capacity utilisation of the network in the current planning week can be achieved. If customer demand locations are clustered around intermodal terminals, both a reduction in drayage costs and an increased network capacity utilisation are obtained. This integrated approach is employed to analyse the impact of tactical decisions on total transport costs and possibly adapt a given service network in order to provide better service levels to customers. It is the first time that the impact of vehicle routing costs at the operational level are included when making decisions on the long-haul service network design at the tactical level. Within this aim, a case study demonstrates the way in which companies can use this solution method in practice and adapt their tactical decisions to the characteristics of their service network and demand data.

With the shift towards synchronised and real-time planning, the task of transport planners becomes more challenging. Therefore, the second part of this thesis focuses on the integration of real-life considerations into the train load planning process, i.e., at a detailed planning level. Due to the varying importance of multiple objectives, each situation requires different weights of each objective. Therefore, no unique best solution exists. This research is the first to present a multi-objective approach which provides a set of solutions which are a priori of equal value. Computational results of a multi-objective approach show that different load plans are obtained with respect to the length utilisation of the train, the urgency of the load units assigned and destination preferences related to drayage operations. The current state of the transport network and incoming transport orders determines which of the plans is the most suitable for a given train departure. With the proposed method, planners can select the most appropriate load plan for each individual departing train.

To summarise, this dissertation provides decision support in an intermodal planning environment which deals with real-life characteristics and allows for flexibility, either by taking an integrated perspective or by providing planning assistance in the form of multiple plans to choose from. The proposed heuristic algorithms can be incorporated in transport management systems. The real-life case studies presented for both planning problems show the advantages and applicability in practice.

7.2 Future research opportunities

The integrated intermodal routing problem proposed in Part II lays the groundwork for further studies on innovative concepts to stimulate a modal shift away from unimodal road transport. First, the real-life case studies provide interesting results using

realistic data. In future research, forecasting techniques could be used to estimate future demand. By using more accurate, forecasted input data, the obtained intermodal routing decisions and service network insights could be further optimised. Second, the heuristic algorithm for the integrated approach can be adapted to a dynamic approach for real-time decision support. Instead of considering a planning horizon of an entire week, the planning horizon could be set to a single day, or even one hour in order to utilise this approach to its full potential. When new load units enter the system, the heuristic algorithm can assign them simultaneously to their most cost-efficient long-haul service and determine their inclusion in truck routes. In this way, the transport plan could deal with continuous updates and optimise the long-haul capacity utilisation in a cost-efficient way. Besides providing decision support for planners, it might also enhance the service levels offered. Finally, it would be interesting to analyse and quantify the effect of such integrated decision-making on empty container repositioning within each of the service areas.

The research on train load planning performed in Part III of this dissertation allows planners to flexibly select a load plan which is most relevant for a specific operational situation. Depending on external information, such as, for example, delays in train departures, planners can decide themselves which objective receives the most priority at a specific moment in time and select the most appropriate load plan due to the proposed multi-objective solution approaches. Future research could focus on reducing the number of load plans presented to planners to reduce the time required for decision-making. Moreover, the proposed solution method indirectly accounts for future orders by assigning a priority value based on the urgency with which load units should be transported. In the future, this might be extended to a look-ahead approach, in which information about future costs is included in present decisions. Such approach uses a multi-period viewpoint, but only a decision is made for the load plan in the current period, i.e., only the next train departure is considered. In this way, new information can be included in each period. As already emphasised, most decisions in intermodal transport are interrelated. In this thesis, train load planning is considered from the viewpoint of a logistics service provider. In a subsequent phase, operations executed by a terminal operator at the terminal are optimised, once the load units to be assigned are known. Future research might incorporate the impact of load planning on terminal operations by means of a second-phase optimisation of handling operations at the terminal. Finally, column generation might especially be an interesting solution method for train load planning in future research, as the problem can be reformulated as a set partitioning problem in which a single feasible configuration must be selected for each wagon.

Appendix A

Detailed results for the exact ϵ -constraint method

Table A.1 shows a detailed overview of the computational results of the exact adaptive ϵ -constraint method discussed in Chapter 5. The left part of the table indicates the total number of iterations for each instance, and how many of these iterations led to infeasible and unique solutions with respect to the objective values respectively. It further shows total computation times for the entire procedure and the minimum and maximum computation times for a single iteration. Finally, the right-hand side of the table shows the number of iterations not solved to optimality within the maximum run time of one hour and the average optimality gap for these solutions.

Table A.1: Detailed results for the ϵ -constraint approach.

Inst	It.	Infeas.	Unique sol.	Procedure			Non-optimal sol	
				Total time (s)	Min Time (s)	Max Time (s)	It.	Avg gap (%)
1	4	1	3	0.77	0.15	0.36		
2	11	0	5	0.77	0.06	0.08		
3	8	0	5	0.76	0.08	0.16		
4	5	2	3	0.39	0.12	0.14		
5	12	3	5	1.2	0.12	0.16		
6	8	0	6	1.12	0.12	0.16		
7	6	1	3	0.73	0.14	0.15		
8	5	1	3	0.4	0.08	0.12		
9	32	4	16	5.62	0.09	0.38		
10	5	1	4	0.46	0.11	0.12		

Inst	It.	Infeas.	Unique sol.	Procedure			Non-optimal sol	
				Total time (s)	Min Time (s)	Max Time (s)	It.	Avg gap (%)
11	15	5	9	6.97	0.14	2.55		
12	9	3	5	0.54	0.08	0.11		
13	7	1	4	0.56	0.07	0.11		
14	8	4	3	0.34	0.08	0.09		
15	2	0	2	0.25	0.12	0.12		
16	5	1	4	0.37	0.09	0.1		
17	8	3	4	0.49	0.09	0.11		
18	6	1	4	0.63	0.09	0.19		
19	10	3	5	1.92	0.18	0.31		
20	7	3	4	0.44	0.09	0.13		
21	13	0	6	73.5	0.11	59.21		
22	5	0	4	0.6	0.1	0.14		
23	7	0	5	20.6	0.11	15.87		
24	8	3	4	0.64	0.09	0.22		
25	4	0	3	0.51	0.09	0.19		
26	8	1	5	1.4	0.13	0.25		
27	6	3	2	0.31	0.1	0.11		
28	5	0	4	0.62	0.11	0.14		
29	8	0	7	13.17	0.12	10.44		
30	12	0	9	1.72	0.08	0.26		
31	1	0	1	0.09	0.09	0.09		
32	14	0	8	2.43	0.11	0.28		
33	9	2	6	1.66	0.12	0.59		
34	16	5	8	1.79	0.1	0.26		
35	2	0	2	0.39	0.16	0.23		
36	11	0	8	1.96	0.09	0.38		
37	22	0	12	3.95	0.09	0.29		
38	3	0	2	0.28	0.08	0.11		
39	13	0	7	2.48	0.16	0.25		
40	6	1	4	0.61	0.09	0.15		
41	6	1	4	0.58	0.08	0.22		
42	4	0	3	0.48	0.07	0.2		
43	9	3	6	0.84	0.13	0.18		
44	7	0	6	1.98	0.11	0.46		
45	3	0	2	0.34	0.11	0.12		
46	8	3	5	3.41	0.11	1.91		
47	2	0	2	0.2	0.08	0.12		
48	3	1	2	1.95	0.57	1.38		

Inst	It.	Infeas.	Unique sol.	Procedure			Non-optimal sol	
				Total time (s)	Min Time (s)	Max Time (s)	It.	Avg gap (%)
49	4	1	2	1.25	0.15	0.82		
50	7	1	5	0.52	0.08	0.1		
51	6	0	5	1.02	0.07	0.23		
52	13	0	12	10.31	0.3	2.77		
53	4	0	3	0.27	0.07	0.07		
54	14	4	6	1.34	0.08	0.23		
55	15	0	10	6.55	0.1	1.67		
56	12	0	10	6.18	0.25	0.98		
57	7	2	4	0.4	0.07	0.09		
58	14	0	8	1.66	0.1	0.14		
59	6	2	3	0.51	0.11	0.14		
60	4	0	4	0.37	0.07	0.13		
61	23	5	16	8.99	0.37	0.72		
62	16	0	13	10.18	0.37	0.95		
63	23	2	18	12.36	0.31	1.28		
64	6	0	6	3.94	0.44	0.84		
65	8	0	5	3.79	0.29	0.79		
66	28	4	16	32.51	0.5	6.96		
67	4	0	3	2.43	0.33	1.1		
68	12	3	7	4.61	0.35	0.71		
69	38	4	23	24.59	0.35	1.88		
70	14	1	11	6.34	0.32	0.93		
71	18	2	11	7.45	0.37	0.57		
72	9	1	8	7.53	0.44	3.02		
73	17	6	10	5.28	0.44	0.56		
74	24	7	16	9.42	0.37	1.68		
75	24	2	15	3313.14	0.65	2088.9	1	*
76	8	1	5	3.26	0.37	0.54		
77	3	0	3	5.12	1.29	2.42		
78	6	0	5	2.93	0.33	0.59		
79	7	1	4	3614.83	0.97	3600.03	1	0.6152
80	13	3	7	30.32	0.98	6.57		
81	4	0	4	1.33	0.23	0.43		
82	18	0	15	10.49	0.36	0.9		
83	18	0	14	8.28	0.28	0.58		
84	12	2	10	10.87	0.71	1.46		
85	17	2	13	20.2	0.44	5.74		
86	10	0	9	5.17	0.31	1.04		

Inst	It.	Infeas.	Unique sol.	Procedure			Non-optimal sol	
				Total time (s)	Min Time (s)	Max Time (s)	It.	Avg gap (%)
87	13	0	9	10.01	0.3	1.61		
88	16	1	11	6.53	0.26	0.82		
89	10	3	6	3.16	0.29	0.71		
90	12	3	7	5.71	0.36	1.18		
91	20	4	10	57.78	0.41	25.11		
92	4	0	3	1.43	0.3	0.44		
93	13	0	10	7.42	0.3	1.22		
94	7	0	7	9.3	0.33	5.23		
95	11	0	7	8	0.43	1.6		
96	29	5	15	18.47	0.55	1.77		
97	3	0	3	1.38	0.24	0.65		
98	10	0	9	4.27	0.35	0.62		
99	9	0	6	4.39	0.34	0.6		
100	12	1	7	4.46	0.26	0.51		
101	21	0	9	8112.28	0.64	3600.02	2	0.7312
102	7	1	5	4.64	0.48	1.16		
103	41	1	28	7916.45	0.97	3600.02	2	2.6024
104	11	3	7	4.84	0.28	1.17		
105	10	2	8	11.09	0.49	3.09		
106	28	6	16	33.34	0.43	11.24		
107	10	2	5	640.21	0.87	613.19		
108	6	0	6	6.44	0.69	1.82		
109	10	0	9	33.95	0.79	6.11		
110	16	0	8	6.87	0.32	0.71		
111	14	0	12	8.37	0.45	0.99		
112	15	0	12	8.96	0.25	0.92		
113	7	0	6	5.05	0.4	1.04		
114	12	1	8	6.21	0.28	1.12		
115	4	0	4	5.08	0.65	1.89		
116	4	0	4	1.59	0.28	0.53		
117	4	0	4	2.01	0.41	0.74		
118	17	0	15	15.54	0.42	1.64		
119	41	2	28	57.68	0.45	6.83		
120	21	0	17	12.44	0.25	2.07		
121	4	0		7.06	1.29	2.82		
122	22	0	4	3667.16	1.72	3600.07	1	0.2066
123	30	0	18	77.5	1.68	4.46		
124	22	0	26	4374.27	1.58	3600.18	1	0.3485

Inst	It.	Infeas.	Unique sol.	Procedure			Non-optimal sol	
				Total time (s)	Min Time (s)	Max Time (s)	It.	Avg gap (%)
125	24	2	16	96.23	2.93	8.83		
126	16	1	18	33.94	1.39	4.56		
127	15	0	14	33.17	1.39	3.67		
128	25	0	15	67.17	1.41	11.05		
129	14	1	21	51.37	2.06	9.22		
130	55	2	10	149.83	1.53	3.93		
131	35	1	36	3851.56	2.11	3600.1	1	0.0018
132	20	0	12	90.5	1.87	14.22		
133	72	4	32	108832.8	3.11	3600.53	29	1.0951
134	10	0	6	3672.58	3.68	3600.11	1	0.1332
135	18	2	9	11051.72	3.54	3600.16	3	0.659
136	54	8	23	38610.62	2.3	3600.17	10	2.0811
137	58	4	23	77673.42	3.08	3600.14	20	0.628
138	18	0	15	1017.85	1.33	903.86		
139	29	0	25	99.58	1.94	10.57		
140	100	5	41	154279	5.24	3600.75	40	0.9539
141	11	0	10	37.33	1.64	5.56		
142	24	0	20	2506.89	4.06	2176.04		
143	26	1	14	26402.83	3.8	3600.17	7	1.4687
144	26	0	19	11940.58	9.54	3600.13	1	1.5724
145	45	1	36	8609.13	2.57	3600.14	2	0.6567
146	13	0	6	3707.09	2.87	3600.09	1	0.1788
147	57	3	43	211.14	2.02	27.37		
148	39	2	33	117.26	1.91	10.12		
149	131	3	87	32524.43	1.83	3600.15	8	0.0697
150	24	1	20	90.76	2.1	7.45		
151	11	0	9	67.37	3.41	14.31		
152	32	0	22	71.04	1.34	4.55		
153	11	0	7	26.16	1.33	3.43		
154	10	0	8	25.63	1.57	3.46		
155	6	0	5	17.42	1.5	6.44		
156	17	0	11	48.89	1.56	8.1		
157	45	5	30	204.38	1.63	17.45		
158	20	0	16	41.4	1.43	2.79		
159	9	0	5	4777.64	2.68	3600.12	1	0.2162
160	21	0	12	82.49	1.4	12.09		
161	34	0	15	205.3	2.15	46.78		
162	15	0	12	149.99	4.86	34.26		

Inst	It.	Infeas.	Unique sol.	Procedure			Non-optimal sol	
				Total time (s)	Min Time (s)	Max Time (s)	It.	Avg gap (%)
163	35	0	23	240.54	2.35	43.13		
164	31	6	20	361.97	6.19	63.6		
165	21	0	14	147.87	2.44	18.84		
166	31	0	22	11482.88	3.6	3600.17	2	2.3488
167	59	8	17	33793.43	3.23	3600.23	9	0.0993
168	47	0	21	4595.46	5.22	3600.19	1	0.0002
169	11	0	6	118.36	4.16	26.21		
170	10	0	9	1734.28	3.02	1698.95		
171	42	0	36	202.29	2.11	24.68		
172	11	0	7	234.23	3.4	67.65		
173	21	0	19	358.68	2.21	191.08		
174	39	0	29	254.52	1.91	70.03		
175	54	0	36	218.31	1.96	7.82		
176	17	0	8	14764.4	4.15	3600.13	4	0.1849
177	4	0	4	246.57	5.71	225.37		
178	29	0	20	352.91	1.82	107.58		
179	20	0	10	143.09	1.11	81.12		
180	6	0	6	28.37	4.09	5.43		

* Unknown solution status

Appendix B

Average proportional deviation from I_{Hr} for solutions obtained using the MDLS-heuristic

Table B.1 indicates the performance of newly generated instances using the parameter configuration proposed by *irace* in Chapter 6, expressed as a percentage deviation of the hypervolume of the reference set. Results with the parameter configuration proposed by *irace* are indicated in bold. An asterisk indicates the parameter configuration resulting in the best overall solution quality based on the sensitivity analysis on different instances.

Table B.1: Average proportional deviation from I_{Hr} (in %).

Parameter	Value	Wagons			Overall
		5	10	20	
w_k	0.60	4.3306	0.6157	1.0049	1.9837
	0.65	4.3306	0.2217	0.9655	1.8393
	0.70	4.2386	0.2011	0.7245	1.7214
	0.75	4.2476	0.2095	0.5645	1.6739
	0.80	4.2386	0.2615	0.4667	1.6556
	0.85	4.2386	0.3367	0.4415	1.6723
	0.90	4.2386	0.1963	0.3564	1.5971
	0.91	0.6272	0.2155	0.3957	*0.4128
	0.95	0.6272	0.2771	2.0622	0.9888
	1	4.2386	0.2001	0.3852	1.608
$A_{LS_{Length}}$	-45	0.6272	0.2374	0.5776	*0.4807
	-20	0.6272	0.2269	0.671	0.5084
	0	4.2386	0.2073	0.4302	1.6254
	20	4.3306	0.4452	0.8677	1.8812
$A_{LS_{Urgency}}$	-2	0.7481	2.3843	2.725	1.9525
	-1	0.7481	2.8099	2.8284	2.1288
	0	4.2386	0.2215	0.3379	*1.5993
	1	8.5737	6.5884	3.4988	6.2203
	2	9.3047	8.3223	3.5622	7.0631
$A_{LS_{DP}}$	-2	1.8935	2.2498	4.4634	2.8689
	-1	1.8935	2.3854	3.7206	2.6665
	0	4.2386	0.2105	0.3344	*1.5945
	1	9.1638	3.9724	5.128	6.0881
	2	11.4811	3.3848	5.6054	6.8238
$DP_{wgn_{destin}}$	0	4.2386	0.1991	0.4045	*1.6141
	1	4.2386	0.31	0.6993	1.7493

Detailed tuning results for the large neighbourhood search heuristic for the integrated intermodal routing problem

For the sequential and integrated approach respectively, Figures C.1 and C.2 show detailed results for different parameter settings, visualised for each instance class individually. The solution quality is indicated on the primary axis on the left, visualised using lines, and the computation time in seconds is displayed on the secondary axis on the right, and shown by the bars in the graph.

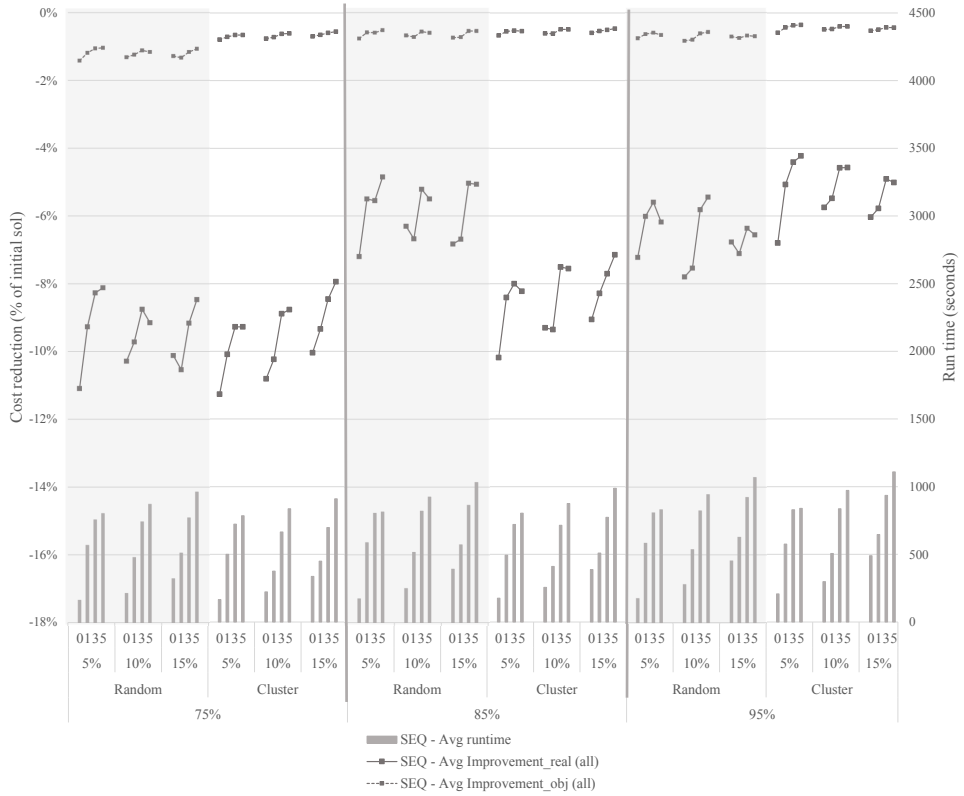


Figure C.1: LNS parameter tuning: experimental results for the sequential approach, for each combination of parameters and all instance classes.

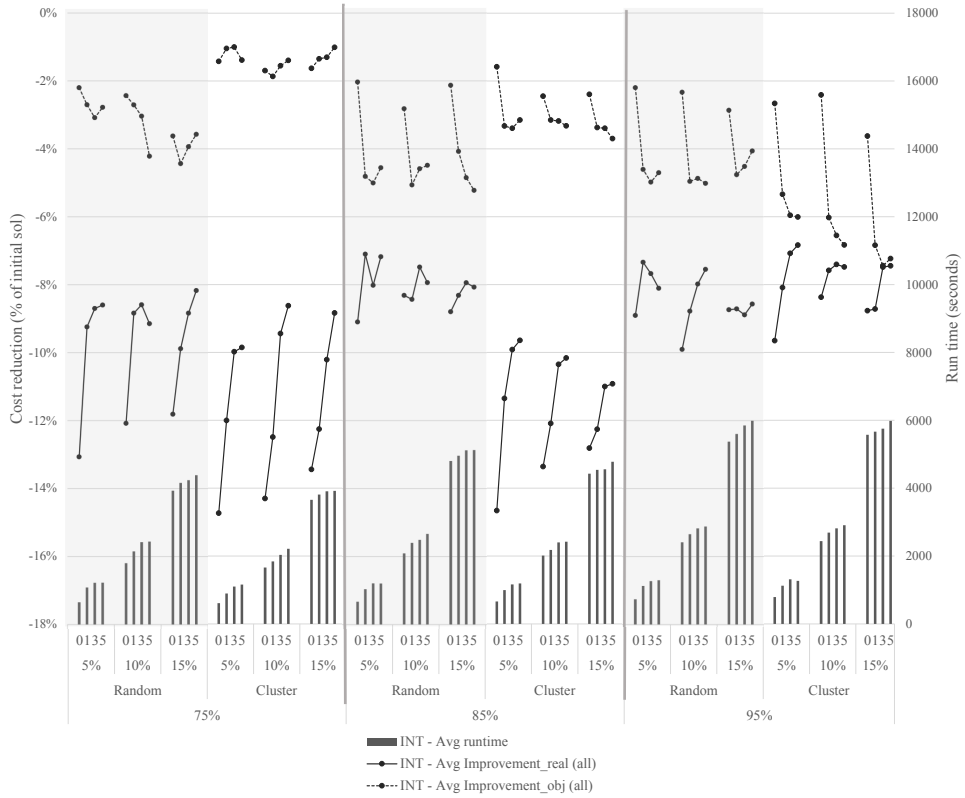


Figure C.2: LNS parameter tuning: experimental results for the integrated approach, for each combination of parameters and all instance classes.

Bibliography

- Aggoun, A., Rhiat, A., Grassien, J.P., 2011. Online Assignments of Containers to Trains Using Constraint Programming, in: Sombattheera, C., Agarwal, A., Udgata, S.K., Lavangnananda, K. (Eds.), Multi-disciplinary Trends in Artificial Intelligence. Springer, Heidelberg. volume 7080 of *LNAI*, pp. 395–405. doi:10.1007/978-3-642-25725-4_35.
- Alliance for Logistics Innovation through Collaboration in Europe, 2014. Corridors, hubs and synchromodality research and innovation roadmap.
- Ambra, T., Caris, A., Macharis, C., 2019. Towards freight transport system unification: reviewing and combining the advancements in the physical internet and synchromodal transport research. *International Journal of Production Research* 57, 1606–1623. doi:10.1080/00207543.2018.1494392.
- Ambrosino, D., Bernocchi, L., Siri, S., 2016. Multi-objective optimization for the train load planning problem with two cranes in seaport terminals, in: *IFAC-PapersOnLine*, pp. 383–388. doi:10.1016/j.ifacol.2016.07.064.
- Ambrosino, D., Bramardi, A., Pucciano, M., Sacone, S., Siri, S., 2011. Modeling and solving the train load planning problem in seaport container terminals, in: *Proceedings of the 2011 IEEE conference on Automation science and engineering*, IEEE. pp. 208–213. doi:10.1109/CASE.2011.6042439.
- Ambrosino, D., Caballini, C., Siri, S., 2013. A mathematical model to evaluate different train loading and stacking policies in a container terminal. *Maritime Economics & Logistics* 15, 292–308. doi:10.1057/me1.2013.7.
- Ambrosino, D., Siri, S., 2014. Models for train load planning problems in a container terminal, in: de Sousa, J.F., Rossi, R. (Eds.), *Computer-based Modelling and Optimization in Transportation*. Springer International Publishing, Cham. volume 262, pp. 15–25.

- Ambrosino, D., Siri, S., 2015. Comparison of solution approaches for the train load planning problem in seaport terminals. *Transportation Research Part E: Logistics and Transportation Review* 79, 65–82. doi:10.1016/j.tre.2015.04.003.
- Andersen, J., Christiansen, M., 2009. Designing new European rail freight services. *Journal of the Operational Research Society* 60, 348–360. doi:10.1057/palgrave.jors.2602559.
- Andersen, J., Crainic, T.G., Christiansen, M., 2009a. Service network design with asset management: Formulations and comparative analyses. *Transportation Research Part C: Emerging Technologies* 17, 197–207. doi:10.1016/j.trc.2008.10.005.
- Andersen, J., Crainic, T.G., Christiansen, M., 2009b. Service network design with management and coordination of multiple fleets. *European Journal of Operational Research* 193, 377–389. doi:10.1016/j.ejor.2007.10.057.
- Anghinolfi, D., Caballini, C., Sacone, S., 2014. Optimizing train loading operations in innovative and automated container terminals, in: *Proceedings of the 19th international federation of automatic control World Congress, 2014*, pp. 9581–9586. doi:10.3182/20140824-6-ZA-1003.02553.
- Anghinolfi, D., Paolucci, M., 2014. A general purpose lagrangian heuristic applied to the train loading problem. *Procedia - Social and Behavioral Sciences* 108, 37–46. doi:10.1016/j.sbspro.2013.12.818.
- Assadipour, G., Ke, G.Y., Verma, M., 2015. Planning and managing intermodal transportation of hazardous materials with capacity selection and congestion. *Transportation Research Part E: Logistics and Transportation Review* 76, 45–57. doi:10.1016/j.tre.2015.02.003.
- Bai, R., Kendall, G., Li, J., 2010. An efficient guided local search approach for service network design problem with asset balancing, *IEEE*. pp. 110–115. doi:10.1109/ICLSIM.2010.5461456.
- Barnhart, C., Ratliff, D.H., 1993. Modeling intermodal routing 14, 205–223.
- Baykasoğlu, A., Subulan, K., 2016. A multi-objective sustainable load planning model for intermodal transportation networks with a real-life application. *Transportation Research Part E: Logistics and Transportation Review* 95, 207–247. doi:10.1016/j.tre.2016.09.011.

- Behdani, B., Fan, Y., Wiegmans, B., Zuidwijk, R., 2016. Multimodal schedule design for synchromodal freight transport systems. *European Journal of Transport and Infrastructure Research* 16, 424–444.
- Beume, N., Rudolph, G., 2006. Faster S-metric calculation by considering dominated hypervolume as Klee's measure problem, in: Kovalerchuk, B. (Ed.), *Proceedings of the Second IASTED Conference on Computational Intelligence*, ACTA Press, Anaheim. pp. 231–236.
- Boardman, B., Malstrom, E., Butler, D., Cole, M., 1997. Computer assisted routing of intermodal shipments. *Computers & Industrial Engineering* 33, 311–314. doi:10.1016/S0360-8352(97)00100-9.
- Bock, S., 2010. Real-time control of freight forwarder transportation networks by integrating multimodal transport chains. *European Journal of Operational Research* 200, 733–746. doi:10.1016/j.ejor.2009.01.046.
- Bontekoning, Y.M., Macharis, C., Trip, J.J., 2004. Is a new applied transportation research field emerging? –A review of intermodal railtruck freight transport literature. *Transportation Research Part A: Policy and Practice* 38, 1–34. doi:10.1016/j.tra.2003.06.001.
- Bostel, N., Dejax, P., 1998. Models and Algorithms for Container Allocation Problems on Trains in a Rapid Transshipment Shunting Yard. *Transportation Science* 32, 370–379. doi:10.1287/trsc.32.4.370.
- Boysen, H.E., 2012. General model of railway transportation capacity, in: *13th International Conference on Design and Operation in Railway Engineering (Comprail)*, Wessex Institute of Technology, New Forest. pp. 335–347. doi:10.2495/CR120281.
- Boysen, H.E., 2014. Øresund and Fehmarnbelt high-capacity rail corridor standards updated. *Journal of Rail Transport Planning & Management* 4, 44–58. doi:10.1016/j.jrtpm.2014.09.001.
- Boysen, N., Flidner, M., Jaehn, F., Pesch, E., 2013. A survey on container processing in railway yards. *Transportation Science* 47, 312–329. doi:10.1287/trsc.1120.0415.
- Boysen, N., Flidner, M., Kellner, M., 2010. Determining fixed crane areas in rail–rail transshipment yards. *Transportation Research Part E: Logistics and Transportation Review* 46, 1005–1016. doi:10.1016/j.tre.2010.05.004.

- Braekers, K., Caris, A., Janssens, G.K., 2013. Integrated planning of loaded and empty container movements. *OR Spectrum* 35, 457–478. doi:10.1007/s00291-012-0284-5.
- Braekers, K., Hartl, R.F., Parragh, S.N., Tricoire, F., 2016a. A bi-objective home care scheduling problem: Analyzing the trade-off between costs and client inconvenience. *European Journal of Operational Research* 248, 428–443. doi:10.1016/j.ejor.2015.07.028.
- Braekers, K., Ramaekers, K., Van Nieuwenhuyse, I., 2016b. The vehicle routing problem: State of the art classification and review. *Computers & Industrial Engineering* 99, 300–313. doi:10.1016/j.cie.2015.12.007.
- Bruns, F., Goerigk, M., Knust, S., Schöbel, A., 2014. Robust load planning of trains in intermodal transportation. *OR Spectrum* 36, 631–668. doi:10.1007/s00291-013-0341-8.
- Bruns, F., Knust, S., 2012. Optimized load planning of trains in intermodal transportation. *OR Spectrum* 34, 511–533. doi:10.1007/s00291-010-0232-1.
- Caramia, M., Dell’Olmo, P., 2008. *Multi-objective Management in Freight Logistics*. Springer London, London.
- Caramia, M., Guerriero, F., 2009. A heuristic approach to long-haul freight transportation with multiple objective functions. *Omega* 37, 600–614. doi:10.1016/j.omega.2008.02.001.
- Caris, A., Janssens, G., 2009. A local search heuristic for the pre- and end-haulage of intermodal container terminals. *Computers & Operations Research* 36, 2763–2772. doi:10.1016/j.cor.2008.12.007.
- Caris, A., Janssens, G.K., 2010. A deterministic annealing algorithm for the pre-and end-haulage of intermodal container terminals. *International Journal of Computer Aided Engineering and Technology* 2.
- Caris, A., Macharis, C., Janssens, G.K., 2008. Planning problems in intermodal freight transport: Accomplishments and prospects. *Transportation Planning and Technology* 31, 277–302. doi:10.1080/03081060802086397.
- Caris, A., Macharis, C., Janssens, G.K., 2013. Decision support in intermodal transport: A new research agenda. *Computers in Industry* 64, 105–112. doi:10.1016/j.compind.2012.12.001.

- Chang, T.S., 2008. Best routes selection in international intermodal networks. *Computers & Operations Research* 35, 2877–2891. doi:10.1016/j.cor.2006.12.025.
- Chankong, V., Haimes, Y., 1983. *Multiobjective Decision Making Theory and Methodology*. Elsevier Science, New York.
- Cho, J.H., Kim, H.S., Choi, H.R., 2012. An intermodal transport network planning algorithm using dynamic programming—A case study: from Busan to Rotterdam in intermodal freight routing. *Applied Intelligence* 36, 529–541. doi:10.1007/s10489-010-0223-6.
- Chouman, M., Crainic, T.G., 2015. Cutting-Plane Matheuristic for Service Network Design with Design-Balanced Requirements. *Transportation Science* 49, 99–113. doi:10.1287/trsc.2013.0493.
- Coello, C.A.C., Lamont, G.B., Veldhuizen, D.A.V., 2007. MOP Evolutionary Algorithm Approaches, in: *Evolutionary Algorithms for Solving Multi-Objective Problems: Second Edition*. Springer US, Boston, MA. Genetic and Evolutionary Computation Series, pp. 61–130. doi:10.1007/978-0-387-36797-2_2.
- Cordeau, J.F., Gendreau, M., Laporte, G., 1997. A tabu search heuristic for periodic and multi-depot vehicle routing problems. *Networks* 30, 105–119. doi:10.1002/(SICI)1097-0037(199709)30:2<105::AID-NET5>3.0.CO;2-G.
- Corry, P., Kozan, E., 2004. Dynamic container train planning, in: Kozan, E. (Ed.), *Proceedings of the Fifth Asia Pacific Industrial Engineering and Management Systems Conference 2004*, Queensland University of Technology, Brisbane, Australia. pp. 30.4.1–30.4.20.
- Corry, P., Kozan, E., 2006a. An assignment model for dynamic load planning of intermodal trains. *Computers & Operations Research* 33, 1–17. doi:10.1016/j.cor.2004.05.013.
- Corry, P., Kozan, E., 2006b. A decision support system for intermodal train planning, in: Kozan, E. (Ed.), *Proceedings of The Second International Intelligent Logistics Systems Conference*, pp. 13.1–13.17.
- Corry, P., Kozan, E., 2008. Optimised loading patterns for intermodal trains. *OR Spectrum* 30, 721–750. doi:10.1007/s00291-007-0112-5.

- Crainic, T.G., 2000. Service network design in freight transportation. *European Journal of Operational Research* 122, 272–288. doi:10.1016/S0377-2217(99)00233-7.
- Crainic, T.G., Kim, K.H., 2007. Chapter 8 Intermodal Transportation, in: *Handbooks in Operations Research and Management Science*. Elsevier. volume 14, pp. 467–537. doi:10.1016/S0927-0507(06)14008-6.
- Deb, K., 2014. Multi-objective Optimization, in: Burke, E.K. (Ed.), *Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques*. Springer US, Boston, MA, pp. 403–449. doi:10.1007/978-1-4614-6940-7_15.
- Demir, E., Burgholzer, W., Hrušovský, M., Arıkan, E., Jammernegg, W., Woensel, T.V., 2016. A green intermodal service network design problem with travel time uncertainty. *Transportation Research Part B: Methodological* 93, 789–807. doi:10.1016/j.trb.2015.09.007.
- Di Febbraro, A., Sacco, N., Saeednia, M., 2016. An agent-based framework for co-operative planning of intermodal freight transport chains. *Transportation Research Part C: Emerging Technologies* 64, 72–85. doi:10.1016/j.trc.2015.12.014.
- Dotoli, M., Epicoco, N., 2016. A technique for the optimal management of containers' drayage at intermodal terminals, *IEEE*. pp. 000566–000571. doi:10.1109/SMC.2016.7844300.
- Dotoli, M., Epicoco, N., Falagario, M., Angelico, B., Vinciullo, A., 2015. A two-step optimization model for the pre- and end-haulage of containers at intermodal freight terminals, *IEEE*. pp. 3472–3477. doi:10.1109/ECC.2015.7331071.
- Dragomir, A.G., Nicola, D., Soriano, A., Gansterer, M., 2018. Multidepot pickup and delivery problems in multiple regions: a typology and integrated model. *International Transactions in Operational Research* 25, 569–597. doi:10.1111/itor.12473.
- Ehrgott, M., 2005. *Multicriteria Optimization*. Springer-Verlag, Berlin/Heidelberg.
- Erera, A.L., Morales, J.C., Savelsbergh, M., 2005. Global intermodal tank container management for the chemical industry. *Transportation Research Part E: Logistics and Transportation Review* 41, 551–566. doi:10.1016/j.tre.2005.06.004.
- Escudero, A., Muñuzuri, J., Guadix, J., Arango, C., 2013. Dynamic approach to solve the daily drayage problem with transit time uncertainty. *Computers in Industry* 64, 165–175. doi:10.1016/j.compind.2012.11.006.

- European Commission, 2009. White paper: A sustainable future for transport –towards an integrated, technology-led and user friendly system.
- European Commission, 2011a. Memo/11/197: Transport 2050: The major challenges, the key measures.
- European Commission, 2011b. White paper: Roadmap to a single European transport area –Towards a competitive and resource efficient transport system.
- European Commission, 2016. Staff working document: The implementation of the 2011 White Paper on Transport "Roadmap to a Single European Transport Area – towards a competitive and resource-efficient transport system" five years after its publication: achievements and challenges.
- Eurostat, 2016. Modal split of freight transport. URL: <https://ec.europa.eu/eurostat/data/database>.
- Feo, T.A., González-Velarde, J.L., 1995. The intermodal trailer assignment problem. *Transportation Science* 29, 330–341. doi:10.1287/trsc.29.4.330.
- Francis, P., Zhang, G., Smilowitz, K., 2007. Improved modeling and solution methods for the multi-resource routing problem. *European Journal of Operational Research* 180, 1045–1059. doi:10.1016/j.ejor.2006.03.054.
- Grasman, S.E., 2006. Dynamic approach to strategic and operational multimodal routing decisions. *International Journal of Logistics Systems and Management* 2, 96–106. doi:10.1504/IJLSM.2006.008220.
- Gromicho, J.A., Oudshoorn, E., Post, G., 2011. Generating price-effective intermodal routes: Generating intermodal routes. *Statistica Neerlandica* 65, 432–445. doi:10.1111/j.1467-9574.2011.00493.x.
- Heggen, H., Braekers, K., Caris, A., 2016. Optimizing Train Load Planning: Review and Decision Support for Train Planners. *Computational Logistics* , 193–208doi:10.1007/978-3-319-44896-1_13.
- Imai, A., Nishimura, E., Current, J., 2007. A Lagrangian relaxation-based heuristic for the vehicle routing with full container load. *European Journal of Operational Research* 176, 87–105. doi:10.1016/j.ejor.2005.06.044.
- ITF, 2015. ITF Transport Outlook 2015. ITF Transport Outlook, OECD Publishing. URL: <http://www.oecd-ilibrary.org/>

- transport/itf-transport-outlook-2015_9789282107782-en, doi:10.1787/9789282107782-en.
- Jiang, X., Bai, R., Atkin, J., Kendall, G., 2017. A scheme for determining vehicle routes based on Arc-based service network design. *INFOR: Information Systems and Operational Research* 55, 16–37. doi:10.1080/03155986.2016.1262580.
- Knowles, J., Thiele, L., Zitzler, E., 2006. A Tutorial on the Performance Assessment of Stochastic Multiobjective Optimizers. TIK-Report No 214. Computer Engineering and Networks Laboratory (TIK), ETH Zurich, Switzerland. Revised version.
- Kovacs, A.A., Parragh, S.N., Hartl, R.F., 2015. The multi-objective generalized consistent vehicle routing problem. *European Journal of Operational Research* 247, 441–458. doi:10.1016/j.ejor.2015.06.030.
- Lai, Y.C., Barkan, C.P., Önal, H., 2008a. Optimizing the aerodynamic efficiency of intermodal freight trains. *Transportation Research Part E: Logistics and Transportation Review* 44, 820–834. doi:10.1016/j.tre.2007.05.011.
- Lai, Y.C., Ouyang, Y., Barkan, C.P.L., 2008b. A rolling horizon model to optimize aerodynamic efficiency of intermodal freight trains with uncertainty. *Transportation Science* 42, 466–477. doi:10.1287/trsc.1080.0245.
- Lenstra, J.K., Kan, A.H.G.R., 1981. Complexity of vehicle routing and scheduling problems. *Networks* 11, 221–227. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/net.3230110211>, doi:10.1002/net.3230110211.
- Li, L., Negenborn, R.R., De Schutter, B., 2015. Intermodal freight transport planning – A receding horizon control approach. *Transportation Research Part C: Emerging Technologies* 60, 77–95. doi:10.1016/j.trc.2015.08.002.
- Li, X., Wei, K., Aneja, Y., Tian, P., 2017. Design-balanced capacitated multicommodity network design with heterogeneous assets. *Omega* 67, 145–159. doi:10.1016/j.omega.2016.05.001.
- Lian, K., Milburn, A.B., Rardin, R.L., 2016. An improved multi-directional local search algorithm for the multi-objective consistent vehicle routing problem. *IIE Transactions* 48, 975–992. doi:10.1080/0740817X.2016.1167288.
- López-Ibáñez, M., Dubois-Lacoste, J., Pérez Cáceres, L., Birattari, M., Stützle, T., 2016. The irace package: Iterated racing for automatic algorithm configuration. *Operations Research Perspectives* 3, 43–58. doi:10.1016/j.orp.2016.09.002.

- Macharis, C., Bontekoning, Y., 2004. Opportunities for OR in intermodal freight transport research: A review. *European Journal of Operational Research* 153, 400–416. doi:10.1016/S0377-2217(03)00161-9.
- Macharis, C., Van Hoeck, E., Pekin, E., van Lier, T., 2010. A decision analysis framework for intermodal transport: Comparing fuel price increases and the internalisation of external costs. *Transportation Research Part A: Policy and Practice* 44, 550–561. doi:10.1016/j.tra.2010.04.006.
- Mantovani, S., Morganti, G., Umang, N., Crainic, T.G., Frejinger, E., Larsen, E., 2018. The load planning problem for double-stack intermodal trains. *European Journal of Operational Research* 267, 107–119. doi:10.1016/j.ejor.2017.11.016.
- Mavrotas, G., 2009. Effective implementation of the ϵ -constraint method in Multi-Objective Mathematical Programming problems. *Applied Mathematics and Computation* 213, 455–465. doi:10.1016/j.amc.2009.03.037.
- Medina, J., Hewitt, M., Lehuédé, F., Péton, O., 2018. Integrating long-haul and local transportation planning: the Service Network Design and Routing Problem. *EURO Journal on Transportation and Logistics* doi:10.1007/s13676-017-0114-7.
- Moccia, L., Cordeau, J.F., Laporte, G., Ropke, S., Valentini, M.P., 2011. Modeling and solving a multimodal transportation problem with flexible-time and scheduled services. *Networks* 57, 53–68. doi:10.1002/net.20383.
- Molenbruch, Y., Braekers, K., Caris, A., Vanden Berghe, G., 2017. Multi-directional local search for a bi-objective dial-a-ride problem in patient transportation. *Computers & Operations Research* 77, 58–71. doi:10.1016/j.cor.2016.07.020.
- Nossack, J., Pesch, E., 2013. A truck scheduling problem arising in intermodal container transportation. *European Journal of Operational Research* 230, 666–680. doi:10.1016/j.ejor.2013.04.042.
- Paquete, L., Chiarandini, M., Stützle, T., 2004. Pareto Local Optimum Sets in the Biobjective Traveling Salesman Problem: An Experimental Study, in: Gandibleux, X., Sevaux, M., Sörensen, K., T'kindt, V. (Eds.), *Metaheuristics for Multiobjective Optimisation*, Springer Berlin Heidelberg. pp. 177–199.
- Paquete, L., Stützle, T., 2018. Stochastic local search algorithms for multiobjective combinatorial optimization: A review, in: Gonzalez, T.F. (Ed.), *Handbook of Approximation Algorithms and Metaheuristics: Methodologies and Traditional Applications*. 2 ed.. Chapman and Hall/CRC Press. volume 1.

- Parragh, S.N., Doerner, K.F., Hartl, R.F., Gandibleux, X., 2009. A heuristic two-phase solution approach for the multi-objective dial-a-ride problem. *Networks* 54, 227–242. doi:10.1002/net.20335.
- Pedersen, M.B., Crainic, T.G., Madsen, O.B.G., 2009. Models and Tabu Search Metaheuristics for Service Network Design with Asset-Balance Requirements. *Transportation Science* 43, 158–177.
- Pérez Rivera, A.E., Mes, M.R.K., 2017. Scheduling Drayage Operations in Synchronodal Transport, in: Bektaş, T., Coniglio, S., Martinez-Sykora, A., Voß, S. (Eds.), *Computational Logistics*. Springer International Publishing, Cham. volume 10572, pp. 404–419. doi:10.1007/978-3-319-68496-3_27.
- Pinto, J.T.d.M., Mistage, O., Bilotta, P., Helmers, E., 2018. Road-rail intermodal freight transport as a strategy for climate change mitigation. *Environmental Development* 25, 100–110. doi:10.1016/j.envdev.2017.07.005.
- Posada, M., Andersson, H., Häll, C.H., 2017. The integrated dial-a-ride problem with timetabled fixed route service. *Public Transport* 9, 217–241. doi:10.1007/s12469-016-0128-9.
- Powell, W.B., Carvalho, T.A., 1998. Real-Time Optimization of Containers and Flatcars for Intermodal Operations. *Transportation Science* 32, 110–126. doi:10.1287/trsc.32.2.110.
- Qu, Y., Bektaş, T., Bennell, J., 2016. Sustainability SI: Multimode Multicommodity Network Design Model for Intermodal Freight Transportation with Transfer and Emission Costs. *Networks and Spatial Economics* 16, 303–329. doi:10.1007/s11067-014-9227-9.
- Reinhardt, L.B., Pisinger, D., Spoorendonk, S., Sigurd, M.M., 2016. Optimization of the drayage problem using exact methods. *INFOR: Information Systems and Operational Research* 54, 33–51. doi:10.1080/03155986.2016.1149919.
- Reis, V., Fabian Meier, J., Pace, G., Palacin, R., 2013. Rail and multi-modal transport. *Research in Transportation Economics* 41, 17–30. doi:10.1016/j.retrec.2012.10.005.
- Resat, H.G., Turkay, M., 2015. Design and operation of intermodal transportation network in the Marmara region of Turkey. *Transportation Research Part E: Logistics and Transportation Review* 83, 16–33. doi:10.1016/j.tre.2015.08.006.

- van Riessen, B., Negenborn, R.R., Dekker, R., 2016. Real-time container transport planning with decision trees based on offline obtained optimal solutions. *Decision Support Systems* 89, 1–16. doi:10.1016/j.dss.2016.06.004.
- van Riessen, B., Negenborn, R.R., Dekker, R., Lodewijks, G., 2015. Service network design for an intermodal container network with flexible transit times and the possibility of using subcontracted transport. *International Journal of Shipping and Transport Logistics* 7, 457–477. doi:10.1504/IJSTL.2015.069683.
- Ropke, S., Pisinger, D., 2006. An Adaptive Large Neighborhood Search Heuristic for the Pickup and Delivery Problem with Time Windows. *Transportation Science* 40, 455–472. doi:10.1287/trsc.1050.0135.
- Rudi, A., Fröhling, M., Zimmer, K., Schultmann, F., 2016. Freight transportation planning considering carbon emissions and in-transit holding costs: a capacitated multi-commodity network flow model. *EURO Journal on Transportation and Logistics* 5, 123–160. doi:10.1007/s13676-014-0062-4.
- Savelsbergh, M.W.P., 1992. The Vehicle Routing Problem with Time Windows: Minimizing Route Duration. *ORSA Journal on Computing* 4, 146–154. doi:10.1287/ijoc.4.2.146.
- Shaw, P., 1998. Using Constraint Programming and Local Search Methods to Solve Vehicle Routing Problems, in: Goos, G., Hartmanis, J., van Leeuwen, J., Maher, M., Puget, J.F. (Eds.), *Principles and Practice of Constraint Programming — CP98*. Springer Berlin Heidelberg, Berlin, Heidelberg. volume 1520, pp. 417–431. doi:10.1007/3-540-49481-2_30.
- Shiri, S., Huynh, N., 2016. Optimization of drayage operations with time-window constraints. *International Journal of Production Economics* 176, 7–20. doi:10.1016/j.ijpe.2016.03.005.
- Soriano, A., Gansterer, M., Hartl, R.F., 2018. The two-region multi-depot pickup and delivery problem. *OR Spectrum* doi:10.1007/s00291-018-0534-2.
- Souffriau, W., Vansteenwegen, P., Berghe, G.V., Oudheusden, D.V., 2009. Variable Neighbourhood Descent for Planning Crane Operations in a Train Terminal, in: Sörensen, K., Sevaux, M., Habenicht, W., Geiger, M.J. (Eds.), *Metaheuristics in the Service Industry*. Springer Berlin Heidelberg, Berlin, Heidelberg. volume 624, pp. 83–98. doi:10.1007/978-3-642-00939-6_6.

- Stahlbock, R., Voß, S., 2008. Operations research at container terminals: a literature update. *OR Spectrum* 30, 1–52. doi:10.1007/s00291-007-0100-9.
- StadieSeifi, M., Dellaert, N., Nuijten, W., Van Woensel, T., 2017. A metaheuristic for the multimodal network flow problem with product quality preservation and empty repositioning. *Transportation Research Part B: Methodological* 106, 321–344. doi:10.1016/j.trb.2017.07.007.
- StadieSeifi, M., Dellaert, N., Nuijten, W., Van Woensel, T., Raoufi, R., 2014. Multimodal freight transportation planning: A literature review. *European Journal of Operational Research* 233, 1–15. doi:10.1016/j.ejor.2013.06.055.
- Steenken, D., Voß, S., Stahlbock, R., 2004. Container terminal operation and operations research - a classification and literature review. *OR Spectrum* 26, 3–49. doi:10.1007/s00291-003-0157-z.
- Sterzik, S., Kopfer, H., 2013. A Tabu Search Heuristic for the Inland Container Transportation Problem. *Computers & Operations Research* 40, 953–962. doi:10.1016/j.cor.2012.11.015.
- Toth, P., Vigo, D. (Eds.), 2014. Vehicle routing: problems, methods, and applications. Number 18 in MOS-SIAM series on optimization. 2. ed ed., SIAM [u.a.], Philadelphia, Pa. OCLC: 935554111.
- Tricoire, F., 2012. Multi-directional local search. *Computers & Operations Research* 39, 3089–3101. doi:10.1016/j.cor.2012.03.010.
- Upadhyay, A., Gu, W., Bolia, N., 2017. Optimal loading of double-stack container trains. *Transportation Research Part E: Logistics and Transportation Review* 107, 1–22. doi:10.1016/j.tre.2017.08.010.
- Verma, M., Verter, V., Zufferey, N., 2012. A bi-objective model for planning and managing rail-truck intermodal transportation of hazardous materials. *Transportation Research Part E: Logistics and Transportation Review* 48, 132–149. doi:10.1016/j.tre.2011.06.001.
- Vidal, T., Crainic, T.G., Gendreau, M., Prins, C., 2015. Timing problems and algorithms: Time decisions for sequences of activities. *Networks* 65, 102–128. doi:10.1002/net.21587.

- Vu, D.M., Crainic, T.G., Toulouse, M., 2013. A three-phase matheuristic for capacitated multi-commodity fixed-cost network design with design-balance constraints. *Journal of Heuristics* 19, 757–795. doi:10.1007/s10732-013-9225-y.
- Wang, X., Regan, A.C., 2002. Local truckload pickup and delivery with hard time window constraints. *Transportation Research Part B: Methodological* 36, 97–112. doi:10.1016/S0965-8564(00)00037-9.
- Wolfiger, D., Tricoire, F., Doerner, K.F., 2018. A matheuristic for a multimodal long haul routing problem. *EURO Journal on Transportation and Logistics* doi:10.1007/s13676-018-0131-1.
- Woodburn, A., 2015. An empirical study of the variability in the composition of British freight trains. *Journal of Rail Transport Planning & Management* 5, 294–308. doi:10.1016/j.jrtpm.2015.12.001.
- Yang, X., Low, J.M., Tang, L.C., 2011. Analysis of intermodal freight from China to Indian Ocean: A goal programming approach. *Journal of Transport Geography* 19, 515–527. doi:10.1016/j.jtrangeo.2010.05.007.
- Zhang, M., Pel, A., 2016. Synchromodal hinterland freight transport: Model study for the port of Rotterdam. *Journal of Transport Geography* 52, 1–10. doi:10.1016/j.jtrangeo.2016.02.007.
- Zhang, R., Yun, W.Y., Kopfer, H., 2010. Heuristic-based truck scheduling for inland container transportation. *OR Spectrum* 32, 787–808. doi:10.1007/s00291-010-0193-4.
- Zhang, R., Yun, W.Y., Moon, I., 2009. A reactive tabu search algorithm for the multi-depot container truck transportation problem. *Transportation Research Part E: Logistics and Transportation Review* 45, 904–914. doi:10.1016/j.tre.2009.04.012.
- Zhang, W., Reimann, M., 2014. A simple augmented ϵ -constraint method for multi-objective mathematical integer programming problems. *European Journal of Operational Research* 234, 15–24. doi:10.1016/j.ejor.2013.09.001.
- Zhu, E., Crainic, T.G., Gendreau, M., 2014. Scheduled Service Network Design for Freight Rail Transportation. *Operations Research* 62, 383–400. doi:10.1287/opre.2013.1254.

- Ziliaskopoulos, A., Wardell, W., 2000. An intermodal optimum path algorithm for multimodal networks with dynamic arc travel times and switching delays. *European Journal of Operational Research* 125, 486–502. doi:10.1016/S0377-2217(99)00388-4.
- Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C., da Fonseca, V., 2003. Performance assessment of multiobjective optimizers: an analysis and review. *IEEE Transactions on Evolutionary Computation* 7, 117–132. doi:10.1109/TEVC.2003.810758.
- Zopounidis, C., Doumpos, M., 2002. Multicriteria classification and sorting methods: A literature review. *European Journal of Operational Research* 138, 229–246. doi:10.1016/S0377-2217(01)00243-0.

Nederlandstalige samenvatting

Intermodaal transport betreft het gebruik van meerdere transportmodi, waarbij het hoofdtransport per trein of schip uitgevoerd wordt. Het biedt een interessant alternatief voor unimodaal wegvervoer, dat een grote, negatieve impact heeft op het milieu. Intermodaal transport kan de impact op het milieu reduceren doordat wegvervoer enkel tijdens het voor- en natransport wordt gebruikt. Daarnaast worden schaalvoordelen bereikt omdat grotere volumes gebundeld worden tijdens transport over langere afstanden. Intermodaal transport draagt zo bij aan de ontwikkeling van een duurzamer transportsysteem en wordt daarom sterk gestimuleerd door de Europese Commissie.

Echter, meerdere transportmodi en operatoren zijn betrokken in dit proces, wat diverse uitdagingen met zich meebrengt voor transportplanners in een complexe en dynamische planningsomgeving. Hierbij is een naadloze integratie van intermodale processen en beslissingen cruciaal. De mate waarin geïntegreerde beslissingsondersteuning geboden kan worden alsook transportkosten bepalen samen de aantrekkelijkheid en perceptie van intermodaal transport. Daarom focust dit onderzoek op de ontwikkeling van innovatieve transportsystemen die problemen met realistische kenmerken en planningsbeperkingen kunnen oplossen, om zo intermodaal transport verder te stimuleren.

In deze thesis worden twee planningsinstrumenten voorgesteld ter ondersteuning van de beslissingen van intermodale transportplanners. Het betreft intermodale routing enerzijds en de laadplanning van treinen anderzijds. Beide instrumenten bestaan uit snelle planningsalgoritmes die kunnen omgaan met realistische probleemeigenschappen. De doelstelling is de minimalisatie van de totale transportkosten en een maximale benutting van de beschikbare transportcapaciteit.

Een literatuurstudie over intermodale lange-afstandsroutering en rittenplanningsproblemen voor het voor- en natransport toont aan dat deze twee beslissingen traditioneel sequentieel genomen worden. In deze thesis worden beide beslissingen gecombineerd tot een geïntegreerd optimalisatieprobleem voor intermodale routing. De kosten en beperkingen van de rittenplanning van vrachtwagens worden betrokken in de beslissing over de lange-afstandsroutering doorheen het intermodale netwerk.

De geïntegreerde benadering wordt vervolgens gebruikt om het intermodale netwerk te analyseren. Op basis van dergelijke analyses kan het netwerk aangepast worden dankzij inzichten in het intermodale netwerk op het tactische beslissingsniveau onder bepaalde verwachtingen omtrent marktomstandigheden. Deze inzichten zijn bijvoorbeeld nuttig bij onderhandelingen van nieuwe transportcontracten.

Na de toewijzing van laadeenheden aan intermodale treinroutes doorheen het intermodale transportnetwerk, dient voor elke vertrekkende trein een laadplan met gedetailleerde ladingsbeperkingen opgesteld te worden waarin laadeenheden worden toegewezen aan specifieke locaties op een trein. Door het operationele, dynamische karakter van deze beslissingen kunnen op korte termijn nog kleine wijzigingen vereist zijn. Zowel een exacte methode als een heuristisch algoritme worden voorgesteld om dergelijk probleem met realistische ladingsbeperkingen en meerdere doelstellingen op te lossen. Het resultaat is een voorstel van enkele plannen, waarbij het belang van elke doelstelling verschilt per plan. Hierdoor kunnen planners voor een specifiek treinvertrek het plan selecteren dat het meest aansluit bij de kenmerken van de planningsomgeving. Het levert aanzienlijke tijdswinst op en biedt zekerheid dat aan alle ladingsbeperkingen is voldaan.

Valorisation

In this addendum, a discussion on the practical relevance of this academic thesis is presented. As emphasised in the introduction, an important challenge to the transport sector and, by extension, the supply chain is the sustainability of the entire system. The environmental impact of our current transport system is high, with producers and consumers often separated geographically by larger distances. In order to achieve a more sustainable transport system, the full cost of transport should be accounted for. Despite the fact that vehicles on the road are becoming greener due to reduced emissions, external costs of congestion and, related to that, costs of pollution increase. A modal shift of transport flows from unimodal road to intermodal transport may relieve congested highways and reduce the environmental impact.

In an ideal transport network, multiple complimentary transport modes and connections should be available. In that way, the most appropriate route can be selected for each transport request at a specific moment in time. This requires flexible and real-time planning, which poses a number of challenges for planners. In this context, adequate decision support and fast planning algorithms with real-life characteristics are needed to support a synchromodal vision. Innovative support systems should be available to help planners decide in a more complex environment. The number of research papers focusing on synchromodal transport and its prerequisites is rising. However, a gap between academic research and practice can be observed, because real-life characteristics of intermodal transport companies are ignored. Therefore, the focus of this thesis is on decision support for human planners at the operational decision level during the planning process of intermodal rail transport.

In order to stimulate intermodal transport, on the one hand, additional costs should be reduced, while on the other hand fast planning algorithms with real-life problem characteristics should be available to accommodate decisions in a complex intermodal planning environment. This thesis offers two intermodal planning tools which can be used by intermodal operators in order to minimise total transport costs and maximise service capacity utilisation, which in turn results in decreasing costs of the transport system.

The *first contribution* of this thesis is the introduction of an integrated intermodal routing problem. After a transport order is received, planners must assign each load unit to an intermodal long-haul service to maximise the overall network capacity utilisation, and local drayage routes must be established to transport load units between load and unload locations and the transshipment terminals for long-haul transport. Usually, these decisions are made in a sequential way. As the level of integration of different transport modes and decision levels influences the attractiveness of intermodal transport, an integrated intermodal routing problem is proposed. By taking local truck routing and long-haul transport planning decisions simultaneously, the aim is to reduce total transport costs. In a *second contribution*, the intermodal routing model is applied to a real-life intermodal network to support the analysis of the impact of tactical service network design decisions made by an intermodal operator. Insights will be presented on how to best utilise a given service network, in order to reduce total transport costs and synchronise available services to the expected demand of orders. The approach aims at reducing the number of road kilometres and increases bundling opportunities by maximising the long-haul capacity utilisation. In this way, it contributes to the modal shift towards intermodal transport and a more sustainable transport system.

At a more operational level, train load planning is concerned with the assignment of load units to specific locations on intermodal trains, accounting for real-life loading restrictions in order to maximise the on-train capacity utilisation. As including these real-life aspects are key to deciding which load units will be assigned to which location, the *third contribution* is the introduction of an exact and heuristic algorithm with real-life train load problem characteristics. It can be used as a decision support tool to provide human planners with a number of load plans from which they can choose the best one for a specific moment in time. Early-stage planning assistance during the booking process may improve the overall capacity utilisation of the available rail services.

Decisions on these two planning problems, intermodal routing and train load planning, influence the throughput of the intermodal transport system and aim at minimising overall transport costs while accounting for customer service requirements. Fast decision support by means of planning algorithms are provided for improved, more efficient planning. Results contribute to a better understanding of the way in which intermodal operators can maximise their service network's transport capacity based on the expected demand of transport orders, and lead to an increased transport capacity utilisation and a minimisation of total transport costs.

Biography and publications

Hilde Heggen was born on the 12th of January 1991 in Maastricht, The Netherlands. She graduated in 2014 as a Business Engineer with major Operations Management and Logistics at Hasselt University, Belgium. After working as a transportation specialist in the private sector, she started her doctoral studies in April 2015 at Hasselt University, within the scope of a double-degree programme with Maastricht University. Working papers on the conducted research were presented at multiple national and international conferences. The results of her doctoral research are presented in this dissertation.

Journal publications

Heggen, H., Braekers, K., Caris, A., 2018. A multi-objective approach for intermodal train load planning. *OR Spectrum*, 40(2), pp. 341-366.

Heggen, H., Molenbruch, Y., Caris, A., Braekers, K., 2019. Intermodal Container Routing: Integrating Long-Haul Routing and Local Drayage Decisions. *Sustainability*, 11, 1634.

Publications in conference proceedings

Heggen, H., Braekers, K., Caris, A., 2014. Analyzing the practical use of load planning models in intermodal rail transport. In: Bruzzone, Agostino G., Del Rio Vilas, David, Longo, Francesco, Merkuriev, Yury, Piera, Miquel Angel (Ed.), *Proceedings of HMS: The International Conference on Harbor, Maritime and Multimodal Logistics Modelling and Simulation*. Bordeaux, France, September 10-12, pp. 11-21.

Heggen, H., Braekers, K., Caris, A., 2015. Optimizing train planning in intermodal rail transportation: Explorative case study and future research directions. In:

Weijers, S., Kraan, M. (Eds.), *Bijdragen Vervoerslogistieke werkdagen 2015*. Breda, the Netherlands, November 26-27, pp. 69-82.

Heggen, H., Braekers, K., Caris, A., 2016. Optimizing Train Load Planning: Review and Decision Support for Train Planners. In: Paias, A., Ruthmair, M., Voß, S. (Eds.), *Lecture Notes in Computer Science 9855, Computational Logistics: 7th International Conference, ICCL 2016*. Lisbon, Portugal, September 7-9, pp. 193-208.

Heggen, H., Ambra, T., Braekers, K., Caris, A., 2017. Rail-road planning systems in the context of synchromodal transportation. In: Cools, Mario; Limbourg, Sabine (Ed.). *Proceedings of the BIVEC-GIBET Transport Research Days 2017: Towards an Autonomous and Interconnected Transport Future*. Liege, Belgium, May 17-18, pp. 316-325.

Heggen, H., Braekers, K., Caris, A., 2017. An efficient heuristic for multi-objective train load planning: a parameter sensitivity analysis. In: Bruzzone, A.G., Piera, M.A., Merkuriev, Y., Longo, F. (Eds.), *Proceedings of HMS: The International Conference on Harbor, Maritime and Multimodal Logistics Modelling and Simulation*. Barcelona, Spain, September 18-20.

Heggen, H., Molenbruch, K., Braekers, K., Caris, A., 2019. Integrating drayage and intermodal routing decisions: a real-life case study. *Proceedings of the BIVEC-GIBET Transport Research Days 2019: Towards a cleaner and smarter transport and mobility*. Ghent, Belgium, May 23-24.

Conference presentations based on abstract

Heggen, H., Braekers, K., Caris, A., 2016. Train load planning optimization : state-of-the-art and opportunities for future research. In: *30th Annual Conference of the Belgian Operations Research Society*. Louvain-la-Neuve, Belgium, January 28-29.

Heggen, H., Braekers, K., Caris, A., 2017. Multi-objective decisions for train load planning. In: *31st Annual Conference of the Belgian Operations Research Society*. Brussels, Belgium, February 2-3.

Heggen, H., Caris, A., Braekers, K., 2018. The simultaneous vehicle routing and service network design problem in intermodal rail transportation. In: *32nd Annual*

Conference of the Belgian Operations Research Society. Liège, Belgium, February 1-2.

Heggen, H., Caris, A., Braekers, K., 2018. Integrating local drayage and network flow planning in intermodal transport. In: 28th European Conference on Operational Research. Valencia, Spain, July 8-11.

Braekers, K., Heggen, H., Crauwels, N., 2018. Intermodal terminal selection and its effect on pre- and end-haulage costs. In: 28th European Conference on Operational Research. Valencia, Spain, July 8-11.

Heggen, H., Caris, A., Braekers, K., 2018. On the advantages of including local drayage costs into intermodal routing decisions. In: The annual international conference of the German Operations Research Society (GOR), OR2018. Brussels, Belgium, September 12-14.

Heggen, H., Caris, A., Braekers, K., 2019. A large neighborhood search heuristic for the integrated intermodal routing problem. In: 33rd Annual Conference of the Belgian Operations Research Society. Hasselt, Belgium, February 7-8.

