Deliverable D 3.1
Analysis of the gap between daily timetable and operational traffic

<table>
<thead>
<tr>
<th>Project acronym:</th>
<th>FR8RAIL II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starting date:</td>
<td>01/08/2018</td>
</tr>
<tr>
<td>Duration (in months):</td>
<td>33</td>
</tr>
<tr>
<td>Call (part) identifier:</td>
<td>H2020-S2R-CFM-IP5-01-2018</td>
</tr>
<tr>
<td>Grant agreement no:</td>
<td>826206</td>
</tr>
<tr>
<td>Due date of deliverable:</td>
<td>30/11/2019</td>
</tr>
<tr>
<td>Actual submission date:</td>
<td>16/12/2019</td>
</tr>
<tr>
<td>Responsible/Author:</td>
<td>Magnus Wahlborg, TrV</td>
</tr>
<tr>
<td>Dissemination level:</td>
<td>PU</td>
</tr>
<tr>
<td>Status:</td>
<td>Draft</td>
</tr>
</tbody>
</table>

Reviewed: (yes/no)
## Document history

<table>
<thead>
<tr>
<th>Revision</th>
<th>Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>20/05/2019</td>
<td>First issue. /AP</td>
</tr>
<tr>
<td>0.2</td>
<td>13/09/2019</td>
<td>Second issue. /AP</td>
</tr>
<tr>
<td>0.21</td>
<td>16/09/2019</td>
<td>Editorial changes based on meeting. /AP</td>
</tr>
<tr>
<td>0.3</td>
<td>02/10/2019</td>
<td>Third issue. /AP</td>
</tr>
<tr>
<td>0.4</td>
<td>15/10/2019</td>
<td>Fourth issue. /AP</td>
</tr>
<tr>
<td>1.0</td>
<td>31/10/2019</td>
<td>1st complete version for Shift2Rail reviewing. /AP</td>
</tr>
<tr>
<td>1.01</td>
<td>13/12/2019</td>
<td>Reviewed by TMT/MW</td>
</tr>
<tr>
<td>1.02</td>
<td>05/02/2020</td>
<td>Minor quality improvements./SG</td>
</tr>
</tbody>
</table>

## Report contributors

<table>
<thead>
<tr>
<th>Name</th>
<th>Beneficiary</th>
<th>Details of contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anders Peterson</td>
<td>LiU</td>
<td>Editor</td>
</tr>
<tr>
<td>Magnus Wahlborg</td>
<td>TrV</td>
<td>Project leader</td>
</tr>
<tr>
<td>Carl Henrik Häll</td>
<td>LiU</td>
<td>Ch. 5</td>
</tr>
<tr>
<td>Christiane Schmidt</td>
<td>LiU</td>
<td>Ch. 4, 5, 11</td>
</tr>
<tr>
<td>Anders Peterson</td>
<td>LiU</td>
<td>Ch. 3, 4.1, 5, 11, 12</td>
</tr>
<tr>
<td>Behzad Kordnejad</td>
<td>KTH</td>
<td>Ch. 3, 5, 11</td>
</tr>
<tr>
<td>Jennifer Warg</td>
<td>KTH</td>
<td>Ch. 5</td>
</tr>
<tr>
<td>Ingrid Johansson</td>
<td>KTH</td>
<td>Ch. 5</td>
</tr>
<tr>
<td>Martin Joborn</td>
<td>RISE</td>
<td>Ch. 6, 11</td>
</tr>
<tr>
<td>Sara Gestrelius</td>
<td>RISE</td>
<td>Ch. 6</td>
</tr>
<tr>
<td>Johanna Törnquist Krasemann</td>
<td>BTH</td>
<td>Ch. 3, 8.1, 9.1, 10, 11</td>
</tr>
<tr>
<td>Sai Prashanth Josyula</td>
<td>BTH</td>
<td>Ch. 10</td>
</tr>
<tr>
<td>Carl-William Palmqvist</td>
<td>LU</td>
<td>Ch. 6</td>
</tr>
<tr>
<td>Tomas Lidén</td>
<td>VTI</td>
<td>Ch. 4.4, 7, 11, 12</td>
</tr>
<tr>
<td>Magnus Wahlborg</td>
<td>TrV</td>
<td>Ch. 3, 11</td>
</tr>
<tr>
<td>Martin Aronsson</td>
<td>RISE</td>
<td>Review of Ch. 4.4 and 7</td>
</tr>
</tbody>
</table>
# Table of Contents

3.1. Related Shift2Rail Projects .......................................................... 6
3.2. Task Descriptions ........................................................................ 8
3.3. Scope and Objectives .................................................................. 9
3.4. Outline of Deliverable D3.1 ............................................................ 10
4.1. Track Capacity Allocation ............................................................ 12
4.2. Long-Term Planning Process ....................................................... 14
4.3. Ad-Hoc Planning Process ............................................................ 16
4.4. Engineering Works and Maintenance ........................................... 18
5.1. Introduction to Timetable Modelling ............................................ 21
5.2. Timetable Simulation ................................................................. 22
5.3. Simulation for Evaluation of Timetable Alternatives ..................... 23
5.4. Timetable Changes Close to Operation: Early and Delayed Departures ............................................................................. 25
5.5. Need for Methodological Developments ...................................... 27
6.1. Introduction and Background ...................................................... 30
6.2. Method ....................................................................................... 33
6.3. Results ....................................................................................... 33
6.5. Discussion and Conclusions ....................................................... 41
7.1. Train Path Adaptations due to Engineering Works and Maintenance ......................................................................................... 43
7.2. Possession Booking and Release of Maintenance Windows .......... 48
8.1. Traffic Control and Line Management .......................................... 53
8.2. Real-Time Yard Management ...................................................... 58
9.1. An Overview of State-of-the-Art .................................................. 61
9.2. An Overview of State-of-the-Practice .......................................... 64
9.3. Concluding Analysis ..................................................................... 65
10.1. Introduction ............................................................................... 68
10.2. Related Work ............................................................................ 69
10.3. A Parallel Algorithm for Multi-Objective Train Rescheduling ....... 70
10.4. Experimental Study and Some Observations ............................... 71
10.5. Conclusions and Future Work ................................................... 74
1. Executive Summary

Fr8Rail II/Work-Package 3 *Real-time network management and improved methods for timetable planning* addresses the problem to improve capacity and punctuality in the railway system by developing concepts and methods for tactical planning and operational traffic. In this report the state-of-the-art has been summarised.

The aim of the project is to:

- Propose concepts and methods that improve the annual and short-term timetable planning.
- Demonstrate how the proposed timetable planning concepts improve the prerequisites for real-time network management.
- Develop methods and tools that can reduce inefficiencies in real time network management.

An important aspect is to improve the coordination between yards/terminals and the line network, and between Infrastructure Manager, Yard Managers, and freight Rail Undertakings.

We motivate our research by the current situation in Sweden, which is characterised by low on-time performance for freight trains, dense and heterogenous traffic on the major railway lines, and a rigid annual timetabling process, which is non-suitable for short-term changes. We believe that better tools for network planning and management on tactical and operational level can help to connect planning and operational processes.

Aiming for improvements of the operational traffic, there is a need for systematic development of methods applied at several planning horizons, based on both simulation and optimization techniques. Close to operation fast methods are needed, for example, based on meta-heuristics.

The maintenance planning process and improvement potential have been described. This is a new piece of the puzzle and it is important to close the gap between timetable planning and operational traffic. The different planning processes at the Infrastructure Manager, the Rail Undertakings and the Maintenance Contractors should be aligned.

When developing new approaches for computational decision-support tools for real-time network management, it is important — but very challenging — to evaluate and benchmark with existing software tools. We also observe that the research stream on computational decision-support and algorithm development for railway traffic management has not yet been sufficiently merged with the corresponding research stream focusing on aspects of human computer interaction.
## 2. Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation / Acronyms</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALNS</td>
<td>Adaptive Large Neighbourhood Search</td>
</tr>
<tr>
<td>ATO</td>
<td>Automatic Train Operation</td>
</tr>
<tr>
<td>bpM</td>
<td>before the planned month</td>
</tr>
<tr>
<td>C-DAS</td>
<td>Connected Driver Advisory System (dispatcher—driver)</td>
</tr>
<tr>
<td>DAS</td>
<td>Driver Advisory System</td>
</tr>
<tr>
<td>ETCS</td>
<td>European Train Control System</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic Algorithm</td>
</tr>
<tr>
<td>GoA</td>
<td>Grades of Automation</td>
</tr>
<tr>
<td>GPU</td>
<td>Graphics Processing Unit</td>
</tr>
<tr>
<td>ILP</td>
<td>Integer Linear Programming</td>
</tr>
<tr>
<td>IM</td>
<td>Infrastructure Manager</td>
</tr>
<tr>
<td>JNB</td>
<td>JärnvägsNätsBeskrivning – the Swedish name for Network Statement</td>
</tr>
<tr>
<td>KPI</td>
<td>Key Performance Indicator</td>
</tr>
<tr>
<td>LNS</td>
<td>Large Neighbourhood Search</td>
</tr>
<tr>
<td>LP</td>
<td>Linear Programming</td>
</tr>
<tr>
<td>MC</td>
<td>Maintenance Contractor</td>
</tr>
<tr>
<td>MILP</td>
<td>Mixed Integer Linear Programming</td>
</tr>
<tr>
<td>PaP</td>
<td>Pre-arranged (train) Path</td>
</tr>
<tr>
<td>PESP</td>
<td>Period Event Scheduling Problem</td>
</tr>
<tr>
<td>PRISM</td>
<td>Plasa Railway Interaction Simulation Model</td>
</tr>
<tr>
<td>PSB</td>
<td>Planerat Större Banarbete (about the same as a major TCR)</td>
</tr>
<tr>
<td>RNE</td>
<td>Rail Net Europe, see <a href="http://www.rne.eu/">http://www.rne.eu/</a></td>
</tr>
<tr>
<td>RU</td>
<td>Railway Undertaking, sometimes called Train Operating Company, TOC</td>
</tr>
<tr>
<td>TAD</td>
<td>Total Accumulated Delay</td>
</tr>
<tr>
<td>TCR</td>
<td>Temporary Capacity Restriction, work/track closure affecting capacity</td>
</tr>
<tr>
<td>TFD</td>
<td>Total Final Delay</td>
</tr>
<tr>
<td>TMS</td>
<td>Traffic Management System</td>
</tr>
<tr>
<td>TPD</td>
<td>Total Passenger Delay</td>
</tr>
<tr>
<td>TrV</td>
<td>Trafikverket, the Swedish Infrastructure Manager</td>
</tr>
<tr>
<td>TTR</td>
<td>Redesign of the International Timetabling Process</td>
</tr>
<tr>
<td>VNS</td>
<td>Variable Neighbourhood Search</td>
</tr>
<tr>
<td>YM</td>
<td>Yard/Terminal Manager</td>
</tr>
</tbody>
</table>
3. Introduction

The present document constitutes the Deliverable D3.1 Analysis of the gap between daily timetable and operational traffic in the framework of the Shift2Rail IP5 project Fr8Rail II/WP 3, tasks T3.1.1 and T3.2.1. The purpose of T3.1.1 and T3.2.1 is to analyse the discrepancy between the scheduled traffic and maintenance defined by the timetable, and the operational control and resulting on-time performance. The two tasks perform this analysis from the two different perspectives — the timetabling perspective as well as the operational perspective.

There is a big opportunity for railways to use current digitalisation to improve capacity and punctuality in the railway system by improved methods for tactical planning and operational traffic. Infrastructure Managers (IMs) needs to improve current process together with Railway Undertakings (RU), Yard/Terminal Managers (YM) and Maintenance Contractors (MCs). Digitalisation and shared data are enablers for automation and to solve complex problems with many actors.

The vision is timetables and operational traffic that is connected into one loop. Timetable planning and traffic control to handle minor disturbances are automated or semi-automated. For bigger disturbances that needs interaction between IMs and RUs there are efficient information sharing and integrated decision making with decision support.

Scheduling and operating railway traffic is a complex and demanding task that requires significant coordination between IMs, RUs, YMs and MCs. An efficient and transparent planning process that results in timetables of good quality, is crucial to ensure an effective capacity usage, high on-time performance of the trains and high quality of service. Significant improvements can be observed, and the digitalization is supporting new, better ways of managing the railway networks, but there are still unresolved significant issues which needs to be addressed.

One important part of the projects is to demonstrate research results. Trafikverket research in Shift2Rail in yard management and network management is focused on the freight corridor Hallsberg–Malmö–Denmark/Germany. For Hallsberg marshalling yard, capacity, traffic, processes, problems and need of decision support is described. Current focus is to study Malmö, to develop the YM role and to proceed with scenarios and demonstrations.

3.1. Related Shift2Rail Projects

Trafikverket has contacts with related Shift2Rail IP5 projects working with yard management and network management.

- ARCC (2016–2019)
- FR8Hub (2017–2020)
The SMART project (SMART, 2019) addresses automation of rail transport. The main goal of SMART is to increase the quality of rail freight, as well as its effectiveness and capacity, through the contribution to automation of railway cargo haul at European railways.

PLASA (PLASA, 2018) and PLASA-2 (PLASA-2, 2019) aim at significantly increasing system robustness in the European rail sector. On the one hand, it aims at facilitating planning activities of various stakeholders in the railway system, making the effects of planning decisions on large, complex railway networks measurable and predictable. In the projects, the macrosimulation tool PRISM is developed (see further discussions in Chapter 5.2).

The ARCC, Automated Rail Cargo Consortium, project (ARCC, 2019) aims at carrying out an initial phase of rail freight automation research activities in order to boost levels of quality, efficiency and cost effectiveness in rail freight operations of the European railway sector. In the ARCC project Hallsberg capacity and yard processes were studied. The roles of IM, YM and RU and how to handle multiple RUs in a yard were described. Deficiencies in the processes and potential for decision support and automation were described in deliverables D2.1 (ARCC, 2017) and D2.2 (ARCC, 2018a). In the ARCC deliverable D2.3 (not public), automation potential and risks with automation were described. The state-of-the-practice of freight traffic planning and operations in Germany and Sweden were also described. ARCC deliverable D 3.1 (ARCC, 2018b) research and innovation activities have been done to identify areas with a need for improved timetable planning methods and outline how new methods can be developed and implemented.

In the Fr8Hub project (Fr8Hub, 2019a) the overall aims are to increase capacity by 10 %, increase operation reliability by 10 %, reduce railway system life cycle cost by 10 %, increase energy efficiency by 10 % and to reduce noise by 5 %. In addition, emissions should be reduced, and punctuality increased. Trafikverket has been involved in WP3 “Real time network management and simulation of increasing speed for freight trains”. Deliverable 3.1 (Fr8Hub, 2018) gives a state-of-the-art description and specifies innovations, demonstrations and simulations to be done. Goal is to study the effects of 1) improved traffic management through better interaction between line and yard, and 2) increased freight speed and its effects on overall increased capacity, punctuality and reduced travel time for both passenger and freight trains. The work has continued in deliverable D3.2 (Fr8Hub, 2019b), where the proposed network management model is presented on a conceptual level.

In the OptiYard project (OptiYard, 2019a) the main objective of is to improve capacity and service reliability by focusing on yard operations, namely by providing an optimised decision support system for YMs. The OptiYard WP5 addresses real-time operations at yards and terminal, and is conceptually linked to Fr8Hub. The link is described in deliverable D5.2 (OptiYard, 2019b), which presents a network decision-support tool and an integration framework.
The yard and network management interaction concept has been developed in co-operation between Fr8Hub and OptiYard. The concept is described in Figure 1 below. In FR8Hub capacity and problems at the marshalling yard in Malmö have been studied. A demonstrator has been developed. A hybrid simulation–optimization approach for re-planning of delayed freight trains has been developed and demonstrated. A multimodal data exchange platform is described with collected data from intelligent videogates.

Both ARCC and Fr8Hub study the Swedish rail corridor Hallsberg–Malmö, and this will be continued in the upcoming project Fr8Rail III, in which WP3 will address real-time network management. The purpose is to improve the operational process by improved methods and information support and human interaction. The research will reduce the gaps between timetable planning and operational traffic, and between yard management and network management.

### 3.2. Task Descriptions

This deliverable comprises Tasks 3.1.1 and 3.2.1 in WP3 “Real-Time Network Management and Improved Methods” in the Shift2Rail project Fr8Rail II. Below follows the respective Task description, which will form the core part of the objective.
3.2.1. Task: To Analyse the Gap between Daily Timetable and Operational Control/On-Time Performance

The perspective is timetable planning. How can timetable planning better connect to operational traffic, in particular with regard to important operational constraints from railway undertakings and yard and terminal managers?

Known deficiencies and challenges are:

- Need of improved capacity planning tools and support
- Better understanding for the full process timetabling – real-time management – on-time performance and the relation between timetable planning process and actual performance
- Timetable planning interaction network – yards/terminals: Needs of improved planning process between IM, RU and YM so freight trains are operated according to timetable plans.
- Freight traffic is more difficult to plan on long-term and does often not follow the annually planned train path in the corridor. On-time performance is low.
- The daily timetable is not conflict free, both regards to train – train interaction and train – construction/maintenance work interaction
- To optimise timetable robustness with buffers between trains and time supplements
- Timetable planning of infrastructure work and maintenance

3.2.2. Task: Gap Analysis and Future

To analyse the gap between operational traffic and daily timetable and the role of operational traffic control today and in the future.

A case study will be defined. Of special interest is to conclude – from an operational viewpoint - why freight trains do not run in their assigned timetable slots. The analysis will also cover the gap between state-of-the-art in research and state-of-the-practice in real time network management to point out the most important challenges in closing of this gap and important KPIs to measure the operational performance. The study should also propose a future scenario for improved real time network management. The analysis takes as starting point other studies that have been performed in the area. Gap analysis regarding current operational setting will focus on the situation in countries contributing in Task 2.

3.3. Scope and Objectives

This project builds upon the Shift2Rail projects ARCC (2019) and Fr8Hub (2019a). In the ARCC project, the prerequisites for railway freight traffic, and in particular the processes at yards and terminals, were studied, along with potential improvements of these processes and the coordination between the yards and the line network. In ARCC WP3 (ARCC, 2018b), some key challenges and knowledge gaps were identified, which serve as the point of origin in FR8Rail II
WP3:

- The daily timetable is not conflict free, both with regards to train – train interaction and train – construction/maintenance work interaction. (This is discussed further in Chapters 4, 5 and 7.)

To construct a conflict free daily timetable, we need to better understand the complete process including the steps; timetabling of traffic and maintenance – real-time traffic management – on-time performance, and important dependencies between timetable quality and robustness, and on-time performance.

- Freight traffic is more difficult to plan on long-term and does often not follow the annually planned train path in the corridor. On-time performance is low. (This is discussed further in Chapter 6.)

We believe that there is a need to improve the planning process and co-ordination between IMs, RUs and YMs to enable freight trains to operate according to the given timetables.

The main objectives of FR8Rail II WP3 are to:

- Propose concepts and methods that improve the annual and short-term timetable planning, aiming at reducing the discrepancy between the planning perspective and the operational perspective.
- Demonstrate how the proposed timetable planning concepts improve the prerequisites for real-time network management. A demonstrator on improved short-term planning and daily planning with improved interaction IM – RU including network and yard/terminals should be developed.
- Develop methods and tools that can reduce inefficiencies in real-time network management by e.g. improving the coordination between yards/terminals and the line network, and between IM and RUs. Requirements for a real time network management demonstrator should be specified.

3.4. Outline of Deliverable D3.1

Chapter 4 starts by presenting the timetable process in Sweden. In Chapter 5 follows an introduction to timetable modelling and alternative approaches to compute, study and evaluated timetables. Chapter 6 presents an analysis of the current on-time performance for freight trains in Sweden pinpointing some challenges ahead.

In Chapter 7 we describe the Swedish maintenance planning process and how track possessions are planned together with traffic.
Chapter 8 describes real time network management and some identified challenges, followed by Chapter 9, which summaries state-of-the-art and state-of-the-practice in computational decision support for train traffic control and disturbance management. Chapter 10 presents the results from a Swedish case study, where a multi-objective parallel algorithm has been applied to solve minor disturbances on a single-tracked line.

In Chapter 11 we compile the experienced knowledge/research gap identified and propose future work and in-depth studies within FR8Rail II. Chapter 12 concludes the deliverable.
4. The Swedish Timetabling Process

Timetabling is the process of allocating track capacity for train traffic and maintenance work. In Sweden, the Infrastructure Manager (IM) is responsible for this process. The planning is split into various phases by time horizon: strategic planning, long-term planning and short-term/ad-hoc planning. Long-term planning relates to the creation of the yearly timetable, outlined in Figure 3. Ad-hoc planning includes all changes to the timetable that are taken after the yearly timetable is published, see Figure 5.

4.1. Track Capacity Allocation

The capacity of a railway is the magnitude of its ability to transport passengers and freight by train on a certain railway line. The capacity depends on the number of trains and their formation in the timetable and is often expressed in terms of number of train paths per time unit.

How the railway’s capacity is used depends on the infrastructure layout and on the frequency and distribution of traffic. Other factors influencing the capacity are the number of tracks, and the possibilities for crossings and overtaking. The design of the traffic control system, and especially the signalling system, is also an important aspect. Further, the types and number of trains using a line, as well as their speeds and stopping patterns (number of scheduled stops, and their respective length) influence the use of capacity.

Trafikverket (the Swedish Transport Administration), which is the IM of the Swedish railway network, follows up several delivery qualities, one of which is capacity, related to the national transport policy objective on accessibility. Two measures for capacity use are reported in the annual report (Trafikverket, 2018b): the capacity use per full day, and for the two-hour interval with the most intense traffic (max 2 hours). The computations are based on the UIC 406 compression technique (UIC, 2004; UIC, 2013) and is further discussed in Trafikverket (2019c).

Figure 2 shows the full day capacity use for year 2018 (Trafikverket, 2019a), divided into three levels. Very high (81–100%), medium high (61–80%) and low (≤60%) capacity use are represented by red, yellow and green, respectively. In the analysis the railway network is divided into 274 sections, out of which 7 % have very high and 23 % have medium high capacity usage. Some critical sections are the single track-line Värmlandsbanan from Kristinehamn to the Norwegian border, as well as the southern parts of the Western and Southern mainlines. Especially we note that the line Malmö–Hallsberg, which will be used as a case study in this project, partly covers such critical sections.
Figure 2: Capacity use in Sweden 2018 (Trafikverket, 2019a).
4.2. Long-Term Planning Process

Figure 3: The annual capacity planning process in Sweden (Trafikverket, 2019d).

The planning of a new annual timetable starts in January/February with an early dialogue, from February until mid-April the train operating companies (RU's) can apply for train paths. All applications for capacity for train paths and engineering works and requirements for services that were received before mid-April, e.g. April 10\textsuperscript{th}, 2017, are managed in the allocation process and result in an established Timetable, see Figure 2. That timetable consists of: the capacity for train paths, engineering works and requirements for services allocation for the entire period of the following yearly timetable, e.g. December 10\textsuperscript{th}, 2017 – December 8\textsuperscript{th}, 2018, cp. (Trafikverket, 2017).

All requests (requirement for services, applications for capacity for train paths, or applications for adjustments to capacity for train paths) that Trafikverket receives after the mid-April deadline (e.g., April 10\textsuperscript{th}, 2017) are managed within the ad-hoc process (Trafikverket, 2017), see Figure 4 and the description below for a description of that process.
This period until mid-April is followed by a consulting period from mid-April until the end of June, when a draft timetable is completed. Another consulting period based on this draft takes place during July, August and large parts of September, culminating in the publication of the fixed timetable in the end of September, e.g. on the 22nd of September in 2017, which then is used starting mid-December, e.g. from December 10th, 2017.

During April and May a strategic dialogue is also performed, which looks 2–3 years into the future. Trafikverket invites RUs and contract customers for a dialogue to discuss preliminary conditions that may affect traffic in 2–3 years’ time. The purpose is to mutually share information and to plan traffic and track work that fits the both parties as good as possible, see Trafikverket (2017). During October and November, a similar dialogue is held, looking 4–5 years into the future.

At least 11 months before the start of the timetable the pre-arranged paths (PaPs) for the Scandinavian-Mediterranean Rail Freight Corridor are published via the company website (www.scanmedfreight.eu) and in the web application Path Coordination System, PCS. These pre-arranged train paths are reserved for international freight traffic in the annual timetable.
4.3. Ad-Hoc Planning Process

All changes after the publication of the new annual timetable are considered to be part of the ad-hoc planning in Sweden. The official start for the ad hoc process in 2017 was October 17, cp. (Trafikverket, 2017). If a RU applies for a new train path until 5 days before the day of operation, Trafikverket must handle this application; all applications that arrive later must not be considered at all. No distinction is made between changes that occur within this timespan. Two other breaking points determine the ad-hoc operation: 72 hours in advance the train driver has the right to obtain his shift times, and at 15:00 the day before operation the planning department hands over the timetable to the dispatching centre. This later point constitutes the definite threshold between tactical and operational planning, see Figure 5.

![Figure 5: Deadlines in Sweden for the timetable before the day of operation (ARCC, 2018b).](image)

Trafikverket (2017) states for the ad-hoc process: “Submitted applications will be processed in the order in which they were received. If a change needs to be made to an application, the applicant
shall recall the submitted application and replace it with a new one. The replacement application will then be given a new arrival date.” The update process is described in Figure 6.

Figure 6: Timetable update process (Trafikverket, 2017).

Engineering works of an acute nature constitute an exception to this rule: they may be planned with short notice, and for safety reasons capacity must sometimes be re-allocated from what was agreed upon in the annual timetable or the ad-hoc planning (Trafikverket, 2017).

When we consider freight traffic, we also need to look at the interaction between marshalling yards and the line network. While the line network and timetable planning are under the control of the IM, a YM is responsible for the yard planning. This involves the planning of car movements and operations at marshalling yards, which is a less structured process than the timetable planning process operated by the IM.

As mentioned above, the IM is not required to handle any applications for a train path that arrives later than five days ahead of scheduled operation. The decision if a train may leave a marshalling
yard earlier than planned is taken by the dispatching centre, in Sweden this decision is usually taken by looking a few stations ahead from the marshalling yard. If this does not result in conflicts with existing trains, an earlier departure is enabled.

### 4.4. Engineering Works and Maintenance

This section provides an overview of the process for planning track access for engineering works and maintenance in Sweden. It is based on Trafikverket (2019b) and Lidén (2016). We first describe the long-term planning that foregoes the yearly timetable planning, then explain the newly introduced concept of *maintenance windows*, and finally describe the short-term planning that takes place after the annual timetable has been published.

#### 4.4.1. Long-Term Planning

Trafikverket plans their major engineering works 4-5 years ahead. Coordination, both with other countries, IMs and RUs, takes place twice a year – with focus on years 4-5 in October-November and on years 2-3 in April-May. For lines on the European rail freight corridors, a specific planning step applies in order to establish PaPs (pre-arranged paths) for the cross-border freight trains. These take precedence over the regular (national) timetables and are planned roughly one year ahead of the annual timetable planning. The PaPs are published in conjunction with the network statement.

The final decision regarding large engineering works (labelled as “Planerade Större Banarbeten (PSB)” in Swedish) for a given timetable period T, is done about 1.5 years prior to the start of T. In addition, all major possessions and maintenance windows that the RUs are expected to adhere to in their train path requests are collected and listed in the network statement (Trafikverket, 2017). The network statement is published approximately one year prior to T, which marks the start of the annual capacity allocation process. Requests for other engineering and maintenance possessions are handled together with the train path requests as described in Section 5.1.

Figure 7 shows the long-term planning process for engineering work possessions, including the preparation of the network statement and the annual timetable planning.

#### 4.4.2. Maintenance Windows and Possessions

In 2015, Trafikverket introduced “maintenance windows” as a method for giving maintenance contractors access to the infrastructure. Maintenance windows are pre-planned train-free windows, typically 2-6 hours long, where most of the recurring basic maintenance should be carried out. The intention is to dimension and construct the maintenance windows before the maintenance contracts are procured and to keep them stable during the contract period. Furthermore, they should be listed in the network statement as a prerequisite for the yearly timetable process. The goal is to improve coordination, obtain longer, less fragmented and more efficient track access times for maintenance, reduce cost and to improve robustness and
punctuality of both trains and work activities.

![Diagram of planning processes](image)

**Figure 7: Long-term planning of engineering work possessions (Trafikverket, 2019b).**

The maintenance windows provide the possibility to perform maintenance activities. The actual work must still acquire possessions in the short-term planning process but there they can be treated as minor possessions that do not affect train traffic. Window time that has not been booked by any possessions are released some time before the day-of-operation. The effects of different release time settings will be analysed in Section 7.2.

### 4.4.3. Short-Term Planning

After the annual timetable has been published, the production year is divided into four revision planning periods (roughly 3 months each). This is where the final details of all train affecting maintenance windows and work possessions are planned, with subsequent adjustments to train paths including possible cancellations, rerouting or replacements. The revision plan is then released about 10 working weeks before the revision period starts. The ambition is that all major
possessions shall be coordinated and settled with the affected contractors and RUs no later than 12–14 weeks before the actual operating day. Consequently, RUs are required to submit their train path adjustments to the IM 18 weeks ahead of operation.

The last tactical planning step is a continuous process which handles a rolling 8-week period, with weekly increments and handover. In this phase only minor possessions which do not affect train traffic should be introduced. The operative plans are then handed over to the traffic control department, which will do the final preparations before the dispatching centres take care of the operational control and dispatching.

During the operational day, unplanned possessions can be authorized using a procedure called “direct planning”. This is a manual process which is documented on paper and that largely lacks support tools. For these reasons direct planning of work tasks is being discouraged, although it provides a possibility for utilizing remaining, unused, track capacity (see Chapter 7.2).
5. Timetabling Modelling

There is a current trend that timetable planning and operation is merging. In the operational process, some traffic management systems have functions for optimization and the role of the dispatcher is changed to be more of an operational planner. In Sweden, the new national traffic control system is specified to control by planning. There is also a current automation process, where commercial software tools for timetabling and train dispatching are developed and implemented.

5.1. Introduction to Timetable Modelling

Timetable planning is to plan trains and train paths, but also to plan maintenance and infrastructure work. Railway timetabling is the process of determining time points (arrivals and departures) for events in a railway network for a set of trains, given constraints regarding travel times, waiting times, waiting patterns, performance of train units, service and quality commitments. Typically, the goal is to utilize the infrastructure as efficiently as possible (capacity utilization, often measured according to UIC 406 (UIC, 2013)). Naturally, a core part of railway timetabling is then the allocation of the resources (line and station tracks) to be used; however, the focus has traditionally been mostly on the line resources, and the track allocations at stations are typically not considered explicitly. However, it is important to realize that for a timetable to be feasible, there must exist a resource allocation such that all safety constraints are satisfied. This holds for the railway line tracks, where the temporal occupancy of any pair of trains of a track segment must respect safety constraints on headway and minimum signalling time. Worth mentioning is also the inherent trade-off between the various goals of timetabling. For example, Abril et al. (2008) states that “there is a trade-off between capacity and reliability/robustness”, which also could be interpreted as a difference between technical (theoretical) and feasible (practical) limits of capacity in terms of robustness. Originally, efficient train operation was defined by the service frequency, while performance has become more important recently, see Cacchiani & Tooth (2018).

Timetabling is a problem that has been extensively studied, in many cases a new timetable, or a larger part of it, is constructed from scratch, see, e.g., Hansen and Pachl (2014), Liebchen (2008) or Törnquist (2006) for an overview. When the timetable is assumed to have a cyclic structure, as is often the case in passenger traffic, a Period Event Scheduling Problem (PESP) can be formulated, see, e.g., Liebchen (2008).

Models applied closer to real-time are typically focused on rescheduling in case of disturbances and disruptions. The aim is then often formulated as to quickly re-obtain a feasible timetable of sufficient quality. For an overview of models of this kind, we refer to, e.g., the survey article by Cacchiani et al. (2014). Andersson et al. (2013) and Khoshniyat & Peterson (2017) show two examples of how rescheduling methods can be used on a more tactical level to redistribute available runtime margin and buffer time to increase stability.
The Swedish timetabling process was described earlier in Chapter 4. We note that, in a deregulated market, applications from several RUs must be coordinated, an aspect which is not always considered in the international literature. Further, since most of the Swedish lines have an irregular and heterogeneous mix of passenger and freight traffic, models relying on PESP are not applicable.

5.2. Timetable Simulation

Trains get delayed and the delays propagate in the network and affect other trains as well, i.e., dynamics in terms of operational perturbations occur. The ability to quickly evaluate the effects of different options is essential in railway timetable planning. Reliability is an important measure of a railway system’s performance. With help of simulation, infrastructure, trains and timetables can be modelled and their operation be estimated.

The system (infrastructure, trains, timetable, etc.) is modelled and disturbances are stochastically added. For a representative number of days, a disturbance level is picked from a distribution and applied to the system in order to evaluate the timetable’s performance under operation. Usually, the outcome is measured in punctuality or delays. Such a simulation can be performed at different levels of detail and can be time-consuming. In a microscopic simulation, exact train paths through the network are simulated, such that infrastructure disturbances can be modelled at the level of individual switches or signals and the train interactions can be evaluated in detail. Macroscopic simulation only includes certain aspects of the network, e.g., links and nodes with some attributes only. Models with a level of detail in-between are called mesoscopic.

In the reminder of Section 5.2, further descriptions of micro, macro and meso simulation models are presented as well as how simulation can be used for evaluation. Further, challenges of using simulation in real-time management are pointed out. The last parts of the chapter provide a connection to the previous and following chapters by describing possibilities to connect simulation and optimization and an overview over evaluation of simulation results.

5.2.1. Microscopic Simulation

Microscopic simulation models are commonly used. They have a high level of detail, typically including the exact station layouts, placement of switches and signals, etc., in order to represent reality well. Borndörfer et al. (2018) recommend dynamic, synchronous, microscopic, stochastic simulation to represent the system in the best way. There are several microscopic railway simulation tools available, e.g., the commercial alternatives RailSys, documented by Bendefeldt et al. (2000), Radtke & Hauptmann (2004), and LUKS (Janecek & Weymann, 2010), while OpenTrack (Nash & Hürlimann, 2004) is an open source alternative. The detail level in microscopic models results in long simulation times when larger networks are considered and increases the complexity of coding and handling of the models.
5.2.2. Macro-/Mesoscopic Simulation

Macroscopic models are much less detailed than microscopic ones. Usually, the network is modelled as a directed graph with nodes (stations) and links (lines), which contain some attributes (e.g. line speed) and use aggregated data. A macro model of a station might for instance not contain the number of tracks. Mesoscopic models have a level of detail in between microscopic and macroscopic models. The required accuracy highly depends on the task. Macroscopic simulation models with a lower degree of detail can be preferable for reducing the runtime and increase user friendliness as well as making it possible to simulate larger networks. An example of a macroscopic model for delay propagation in large networks can be found in Büker & Seybold (2012).

Another macroscopic tool is PRISM (Plasa Railway Interaction Simulation Model), see, e.g., Zinser et al. (2018; 2019), initiated in the Shift2Rail project PLASA (2018) and further developed in PLASA-2 (2019). With that Monte Carlo railway simulation tool, large networks can be simulated within a short runtime (i.e., minutes instead of hours/days) providing realistic results that represent a typical day of operation. That gives the possibility to for instance estimate the performance of a planned timetable, or the impact of a construction site on a timetable’s reliability. Cui et al. (2018) present a model which can be adjusted according to the user’s and project’s needs, with the possibility of micro-, meso- and macroscopic simulation.

5.2.3. Simulation in Real-Time Management

As described in the previous sections, simulation is usually quite slow and requires large efforts, while fast decisions are the key for real-time management. Today, the capability of decision support tools for operational planning in form of detection and resolving of potential train conflicts is limited and usually requires human input. Computer support mainly focusses on train routes and infrastructure elements. Projects developing future traffic management systems, on the other hand, have online systems in mind that assume a broader view on rail operations, where information on other resources such as rolling stock, staff, or yard and terminal operation is also included. Research in future traffic management systems and operational planning systems in this area with published results are included both in Shift2Rail IP2 (In2Rail, 2015; X2Rail-2, 2019) and IPS (ARCC, 2019; Fr8Hub, 2019) projects.

As described, conventional simulation is too slow for use in real-time management. However, simulation results would be very valuable for it: The possibility to quickly estimate the consequences of choice A or B on network-wide reliability would offer huge advantages for the traffic operation. Ongoing development of macro/meso simulation models as for example done in PLASA-2 can be promising for future traffic operation control systems.

5.3. Simulation for Evaluation of Timetable Alternatives

As simulation represents the operation of a system, evaluation usually focusses on performance indexes as delays or punctuality in general. Chapter 6 will focus on statistical analysis of real data.
Here on the other hand, it is discussed how the simulation results can be used, and which results are of interest.

The two most frequently used performance indices for railway traffic, punctuality and delays, are usually analysed in an aggregated form. Basically, analysis could be done for every specific train at every defined point in the simulation network and in time. That makes sense for very specific analyses and if distributions are of interest. However, most projects are of a larger scale and require an aggregation. Research faces the performance usually by measuring average delay (for example Lindfeldt (2015)), the number and length of occurrences (for example Nicholson et al. (2015)) or punctuality (for example Huisman & Boucherie (2001), average delay and punctuality). Büker & Seybold (2012) include mean delay and variance together with punctuality. An extended approach is to face passenger delays. Robonek et al. (2016), for example, evaluate the resulting delays per passenger instead of train and Sels et al. (2016), who include delays as expected values of stochastic variables to minimize the total estimated time for passengers. That enables for example to include missed connections due to delays. For a choice of stations, the Swedish RU SJ (2019) measures for example also passenger punctuality and compares to the punctuality of the trains, that means the percentage of passengers and trains, respectively, arriving on time.

Summarizing these approaches, an extract of parameters for evaluation of performance data is presented:

- Where to measure? E.g., at the final station of each service, at each passenger stop or each node the train passes (Example Swedish Transport Administration: Punctuality is presented for selected stations.)
- Aggregation of services:
  - All runs aggregated (e.g., in order to compare different time intervals)
  - Individual train runs (no aggregation)
  - Differentiation between freight and passenger trains, long/short distance, services using the same line, etc.
- Evaluation unit:
  - Number of delayed train departures
  - Number of passengers/amount of transported goods (weight or volume)
  - Distance-based
- Time intervals: E.g., differentiation day/night, peak/off-peak, etc.
- Limit for punctuality (Example Swedish Transport Administration: trains arriving more than six minutes late or being cancelled are counted as delayed.)
- Extended analysis with
  - Different weights, for example higher weight for prioritized services or stops
  - Statistical measurements (median, standard deviation, etc.)

The choice of parameters and aggregation level in the analysis depends on the aim of and possibilities in the specific project.
5.3.1. Combination of Simulation and Optimization

Numerous scheduling/timetabling studies for rail services in different aspects have been conducted, including optimization as well as simulation studies of rail capacity. However, combined optimization and simulation techniques to achieve optimal performance are rare in rail applications.

Methods in the literature that deals with uncertainty in data (in the context of optimization) can, according to Fischetti & Monaci (2009), be divided into stochastic programming methods and robust programming methods. Stochastic programming methods often tend to become too complicated, while robust programming algorithms often are easy to use and to solve. Cacchiani & Toth (2012) underline the efforts made to develop methods and models for producing robust timetables.

Hassannayebi et al. (2014) developed a two-stage simulation optimization approach based on Genetic Algorithms (GA) in order to minimize the expected passenger waiting times. The optimization is intended to adjust headways through simulation experiments to achieve robust timetables for operation of an urban transit rail system. A further developed methodology with robust multi-objective stochastic programming models for train timetabling is presented by Hassannayebi et al. (2017). Fischetti & Monaci (2009) have proposed a method called light robustness to solve linear programming (LP) problems with uncertainty in data. In this approach the maximum objective value deterioration is fixed and a "robustness goal" is modelled using an optimization framework. Compared to some stochastic programming models, of various complexity, they conclude that Light robustness seem to be the most suitable tool to solve large-scale real scenarios. The approach and others are applied to timetabling in Fischetti et al. (2009).

Mannino et al. (2016) consider an exact approach for train timetabling based on a microscopic-macroscopic decomposition model taking into account both operational and cyclicity constraints, for example on routing in stations. The approach is evaluated on a case study of small instances on a railway section in Norway. Högdahl et al. (2019) present a delay prediction model for a MILP timetable optimization model, which is combined with microscopic simulation. By minimizing travel times and maximizing travel time reliability delays can be significantly decreased.

5.4. Timetable Changes Close to Operation: Early and Delayed Departures

The annual timetable is constructed several months in advance, and while passenger trains aim to follow the timetable such that passengers can make their connections and arrive in time, deviations from the timetable might be desirable for freight traffic companies and other actors. For example, a train might be completed early at the marshalling yard (that is, all goods and wagons plus the locomotive have arrived and are combined for a train). In that case, the train occupies space at an often overcrowded marshalling yard. If resources, like train driver, are available, it could be beneficial — not only for the RU — to depart early. This problem of early-completed trains blocking capacity at the marshalling yards while not needing it has already been
identified in the ARCC project (ARCC, 2019). This frequently results in the wish to depart trains earlier than scheduled from marshalling yards. All applications for a train path that the Infrastructure Manager (IM) receives later than five days ahead of operation do not have to be handled, but can still be handled. The wish to depart early—a real time application for a new train path—after early completion of the freight train clearly falls into this category.

The decision if a train may leave a marshalling yard earlier than planned is taken by the dispatching centre, in Sweden, the dispatcher—in absence of a decision support system—usually takes this decision by looking a few stations ahead from the marshalling yard. According to Sköld (2017), it is doubtful that the dispatcher will check the capacity for the entire planned train path until the next marshalling yard (for example from Malmö to Hallsberg in Sweden). If looking a few stations ahead does not result in conflicts with existing trains on this stretch, an earlier departure is enabled. In addition, it is difficult to account for operational effects as for example delays. However, this short-sighted approach might result in long queues at intermediate stations, in particular, after the stretch checked by the dispatcher.

Similarly, certain parts of a train (e.g., wagons, goods) might be delayed, and depending on deadlines for other goods on the planned train, the RU might decide to depart an incomplete train in time, or to postpone departure. In all cases of a deviation from the timetable, the RU must apply for a new train path with the infrastructure manager. For an earlier, or later, departure, the IM needs to decide if a suitable train path is available, bearing in mind that the timetable must be robust against disturbances.

When an RU requests to insert a new train path into an existing timetable, and we aim for a decision support tool instead of the current manual solution, an algorithm that handles such requests must check for possible feasible solutions and use some evaluation metrics to determine which, out of several possible solutions would produce the best timetable. What is the “best” timetable could, e.g., be evaluated based on some robustness criteria describing how flexible the resulting timetable is to disturbances, and to future new requests; on shortest possible travel time; on earliest possible departure from the marshalling yard; etc. That is, also for short-term timetable rescheduling we aim to schedule conflict-free train paths.

We need to determine time points (arrivals and departures) for events in a railway network for a subset of all trains, given constraints regarding travel times, waiting times, waiting patterns, performance of train units, service and quality commitments. Instead of maximizing the capacity utilization several objectives, optimizing either the new train path’s performance or its effect on other trains, can be considered.

To ensure that the new train obtains a feasible train path to its destination is a precondition, but it is also essential to optimize the insertion of a new train in an existing timetable. Various authors considered adding a new train to an existing timetable, amongst others Burdett & Kozan (2009). In this context, Ingolotti et al. (2004) consider adding new trains to a heterogeneous, heavily loaded railway network, and aim to minimize the traversal time for each additional train. Flier et al. (2009), see also Flier (2011), present a shortest path model using a time-expanded graph, which integrates linear regression models based on extensive historical delay data, that gives Pareto
optimal train paths with respect to travel time and risk of delay. Cacchiani et al. (2010) also consider the problem of inserting a single freight train into an existing schedule of fixed passenger trains. They assume that the RU specifies an ideal timetable that the IM can modify, which also includes the use of a different path. The authors aim to add the maximum number of new freight trains, such that their timetable is as close as possible to the ideal one. To do so, they use a heuristic algorithm based on a Lagrangian relaxation of an Integer Linear Program (ILP).

Ljunggren et al. (2020) present an algorithm that inserts a train path in an existing railway timetable, with the objective of creating a resulting train path that maximizes the bottleneck robustness. All other train paths (already in the timetable) are considered as fixed. Peterson et al. (2019) presented an algorithm to insert a maximum number of train paths for a specific train type within given time windows into an existing timetable.

By improving the computational-decision support for short-term rescheduling, we want to develop a useful tool for rescheduling close to operation, which also can be used to decline changes. A key issue is to better integrate tactical and operational planning.

5.5. Need for Methodological Developments

Aiming for improvements of the operational traffic, there is a need for methodological development of methods applied at several planning horizons. For larger disturbances that make a certain part of the rail network unavailable, it is interesting to reroute trains via geographically alternative routes. Additionally, such an alternative route may be requested from YMs or RUs.

The conventional microscopic simulation approaches outlined above are applicable for long-term planning. For application in real time management, other, faster methods closer to operation are needed, for example based on meta-heuristics.

The previous sections described the importance of a proper way to predict the performance of a timetable. Methods and ways to evaluate the outcome were discussed and shortcomings shown. Within the scope of Fr8Rail II WP3, a demonstrator for timetable improvements with plug-in modules is developed. Based on the indicated needs, a simulation tool that can handle larger networks and is connectable to other modules is planned to be integrated in the demonstrator. Together with the plug-in module for short-term re-planning this offers a great opportunity for improving timetables in line with the indicated advantages of connecting optimization and simulation.

5.5.1. Simulation Module for Long-Term Timetable Planning

The simulation module in the project is planned to be based on PRISM (Zinser M., et al., 2019). Macroscopic infrastructure and timetable data enable the possibility of speeding up the simulation process and establishing a connection to optimization modules that cannot handle too detailed data. The purpose of the simulation model is to estimate the effect of randomly assigned
disturbances from a given distribution on a timetable. In the demonstrator, both original and resulting timetables can be visualized. This timetable performance analysis will be possible with disturbances applied on a train run, at stations or on-line segments with differentiation between train types, if stops are planned, etc. Visualization of statistical values for one train path could also be implemented.

The first prototype of the demonstrator is taking a macroscopic approach to the timetable planning, i.e. microscopic details such as the actual capacity limitations of the stations, are not considered. This would be a valuable additional functionality to improve the demonstrator. The aim is to develop a module that will be ready for integration with the demonstrator in the later stages of Fr8Rail II WP3, or in Fr8Rail III WP2.

5.5.2. Optimization Module for Timetable Changes Close to Operation

The previously presented approaches for timetable changes close to operation, consider all existing train paths in the timetable as fixed. By allowing some train paths to be slightly changed (e.g., moved in time), it is possible to find better solutions to the problem of inserting additional train paths. To give a simple example, see Figure 8, where we are given a timetable with three train paths, of which one (dark green) belongs to a specific RU. Only when slightly delaying that train a new train (light-green) belonging to the same RU can be accommodated with enough headways (red).

This means that if an RU requests a new train path, and no feasible solution can be found satisfying that request, it could be interesting for the RU to allow that some of its other (already scheduled) train paths can be moved in time to give room for the new request in the timetable. This could either be done to find a feasible solution (as in the example), or to improve a given feasible insertions of a certain train paths. Additionally, it is possible to have more than a single train that should be added to the timetable. Hence, we estimate that solutions to this problem will lead to a valuable part of a decision support system.

Figure 8: An additional train (light-green) is inserted (right) in an existing timetable (left).

We plan to integrate this added flexibility by creating an algorithm (meta heuristic) that, based on
the new train path(s) to be inserted, removes several existing train paths from the timetable (to make room for more flexibility) and reinsert them all (the new and the removed train paths) into the timetable in the best way. Of course, this can only be beneficial, if the RU actually has other trains running within a time window around the desired departure. Hence, we will consider only larger freight rail companies, like Green Cargo in Sweden.

The main idea of the algorithm has previously been used for several other transport related optimization problems (see e.g. Häll and Peterson (2013)) and is often referred to as “Ruin and recreate methods”, simply because one first ruins the existing solution (by removing train paths from the timetable) and then recreates the solution (by inserting them back into the timetable in a better way).

We will apply a heuristic solution approach to remove and reinsert train paths. We plan to use an adaptive large neighbourhood search (ALNS). ALNS is based on the large neighbourhood search (LNS) introduced in Shaw (1998), and in some respects it resembles the Rip-up and Reroute described in Dees & Karger (1982) and the Ruin and Recreate method presented by Schrimpf et al. (2000). The principle of the LNS is that in each iteration, an existing solution is destroyed by some operator and then repaired again by another operator. In the ALNS, several destroy operators and several repair operators can be used. The probability of choosing a specific operator (or combination of destroy and repair operators) changes over time depending on their past performance. There are several different destroy and repair operators proposed for different types of transport related problems, see e.g. Gschwind & Drexl (2019), Häll & Peterson (2013), Parragh & Schmid (2013), and Røpke & Pisinger (2006).

By using large neighbourhoods and the diversity of the neighbourhoods that are introduced when using several different operators, the ALNS algorithm can explore large parts of the solution space, and hence seldom gets trapped in local minima.

In many transport related problems, e.g., the area of vehicle routing, applying destroy and repair operators corresponds to first removing a certain number of requests from their routes and then reinserting them into vehicle route again. This can in our application of train timetabling be compared to destroying an existing timetable, by removing a number of train paths, and repairing it (re-constructing the timetable) by reinserting the, possibly adjusted, train paths. Barrena et al. (2013) use an ALNS heuristic for designing train timetables, considering a dynamic demand, with the objective of minimizing passenger waiting times.

According to Pisinger & Røpke (2007), there are three main steps in the design of an ALNS-framework:

i. Choose a number of fast construction heuristics which are able to construct a full solution given a partial solution
ii. Choose a number of destroy heuristics.
iii. Choose a local search framework (e.g. simulated annealing or tabu search) at the master level.
When choosing destroy heuristics it is important to make sure that the neighbourhoods the heuristics work with can both intensify and diversify the search.

6. Punctuality of Freight Trains in Sweden

This chapter describes the punctuality of freight trains in Sweden, and how it varies along several dimensions. It is based on data analysis and on interviews conducted at one of the country’s major marshalling yards and provides an overview of some of the issues currently affecting freight trains in Sweden. The chapter begins with a description of the background, continues with method, results, and a final section with discussion and conclusions.

6.1. Introduction and Background

In Sweden, passenger train punctuality has been quite stable at around 90%, measured as having a delay of less than six minutes at the final stop (Trafikanalys, 2016). Freight and long-distance passenger trains typically perform much worse, with between 60 and 80% punctuality (Trafikanalys, 2016). These figures can be compared to the industry target of 95% punctuality, overall, by 2020 (Gummesson, 2018). Research on passenger trains in Sweden show that their delays are mostly caused by small disturbances at stations (Palmqvist, 2019). Details of how the timetables are planned have also been shown to make a big difference for passenger trains, as has weather, particularly snow and extreme temperatures (Palmqvist, Olsson, & Hiselius, 2017).

Less is known about the performance of freight trains, and what factors contribute to their timeliness. In Sweden, it is well known that both the regularity (the extent to which scheduled trains run and arrive to their final destinations at all) and punctuality (arriving with less than six minutes of delay) are lower for freight trains (Trafikanalys, 2016). Another big difference, compared to passenger trains, is that many departures and arrivals occur significantly ahead of schedule (see Table 3 further below). Essentially, the channels that freight trains travel through are much broader than for passenger trains. Put differently, they deviate much more from the timetable than passenger trains. Of interest is the departure punctuality – the extent to which freight trains depart early or late from the marshalling yards. This varies much more than for passenger trains.

6.1.1. Swedish Railways

To give some context of the Swedish railways, about 80% are single track, and 75% of the tracks are electrified (Trafikverket, 2017). There are double tracks around and between the three largest cities: Stockholm, Gothenburg, and Malmö. This is where most trains operate. Since a series of major regulatory changes began in the late 1980’s, there has been a large increase in both services and ridership. This increase has mainly been concentrated on local and regional passenger services, which have grown by 2-3% per year since about 1990 (Trafikanalys, 2018a). Freight trains currently make up about 15% of trains (Trafikanalys, 2018a), and their volumes are affected more by macroeconomic factors than by a long-term trend (Trafikanalys, 2018b). Over time, this means
that the share of freight trains is slowly declining.

Hallsberg is the largest marshalling yard in Sweden. It connects other yards, the largest of which are in Gothenburg, which handles most of the overseas shipping, and in Malmö, which connects to the continental railways via the Øresund bridge (see see Figure 9). Most freight trains running between Hallsberg and Malmö exchange personnel in Nässjö (Gjerdrum, 2019), so that the drivers can return home during their shifts.

![Railway network in Southern Sweden](image)

**Figure 9: Railway network in Southern Sweden (Trafikverket, 2019e).**

6.1.2. **Malmö Marshalling Yard**

In this project, we put a special emphasis on Malmö marshalling yard (MGB), which constitutes the southern endpoint for freight transports on the Southern Mainline and is the portal for traffic to and from continental Europe. The northern endpoint, Hallsberg (HRBG), is thoroughly described and analysed in the ARCC-project (ARCC, 2017; ARCC, 2018a). The situation and activities at the yards on the endpoints of the line have large impact on the freight traffic along the whole line.

Malmö Marshalling yard (see Figure 10) is a one-sided hump yard with a combined arrival/departure yard (11 tracks) and a classification bowl (26 tracks). In proximity of Malmö Marshalling yard is Malmö harbour, a multi-modal terminal operated by Mertz Transport AB, a
passenger yard and service providers such as e.g. cleaning facilities and locomotive depots. Therefore, the arrival/departure yard is not only used by trains that should or have been marshalled, but also by trains that drop off or pick up cars from the multi-modal terminal and/or harbour, and by trains that only stop there for a short period for e.g. driver or locomotive changes. The roles and responsibilities at Malmö Marshalling yard are similar to those at Hallsberg Marshalling Yard. Green Cargo is the marshalling yard manager while Trafikverket is the infrastructure manager that owns the marshalling tracks and equipment. However, in Malmö some of the neighbouring infrastructure resources are owned by other infrastructure managers (see Figure 10). Further, MGB is the hub for trains to/from continental Europe that travel on the Øresund bridge. MGB is therefore affected by both the Danish/international and the Swedish systems.

![Figure 10: Malmö Marshalling Yard. The colours show the tracks of the different IMs.](image)

For example, the rules for when and how train numbers can be changed are different in Denmark and Sweden, and trains arriving from Denmark have new train numbers much more often than what is common in Sweden (TTT, 2019a; Edholm, 2019b). Trains travelling in both Denmark and Sweden may also have one train path in Sweden, and another one in Denmark. These train paths often “meet” in Malmö, which means that an arriving train (with one train number) is the same as another departing train (with another train number). However, this connection is not always made explicit (Edholm, 2019b).

The combined arrival/departure yard is considered a critical resource for MGB (TTT, 2019b; TTT, 2019a). Special IM “yard planners” are responsible for planning the MGB arrival/departure yard (Edholm, 2019b). A master plan is made for week 10 (approximately the first week in March) and is then updated throughout the year as the situation develops. The purpose of the plan is to ensure that there is enough track capacity for the planned arrivals and departures, but the plan is not intended to be precisely followed. For example, shunting movements to/from the harbour, and
to/from the multimodal terminal are not included in the plan, and the exact track allocations are expected to change depending on e.g. the length and weight of the trains on the day of operation. Also, even though this updated arrival/departure yard plan contains plenty of useful information, the plan is not (yet) always passed on to the IM dispatchers on the day of operation.

An issue that is special for one-sided yards is that roll-outs from the classification bowl to the arrival/departure yard is affected by cars being pushed over the hump into the classification bowl (a roll-in). Depending on which track a departing train is parked on in the classification bowl, it may or may not have to be pulled back over the hump in order to reach the arrival/departure yard. In Malmö, the classification bowl has three groups of tracks, colloquial named “the hollow”, “over the hump” and “the harbour hollow” or “north of the hump” (see Figure 10). Trains parked on tracks in the “hollows” can be shunted to the arrival/departure yard without having to be pulled back over the hump (although roll-in operations can still not happen at the same time as a roll-out) while trains on the other tracks have to be pulled back over the hump to reach the arrival/departure yard. Making sure that the roll-ins are scheduled together with the roll-outs are especially important when trains should depart directly from the classification bowl (TTT, 2019a).

### 6.2. Method

For this study we have considered freight trains travelling in either direction between the yards Malmö Godsbangård and Hallsbergs Rangerbangård, with scheduled stops (not merely passing by) at both yards, during the timetable years of 2013-2017. During each of these years, there are on average a little less than 5 000 such freight trains scheduled, of which two thirds actually ran. Only trains travelling along the main route, described above, have been considered. We have used train movement data with minute-level precision from the Swedish Transport Administration’s Lupp database (Trafikverket, 2018a). This data has been collected from the signalling system, with observations of arrival and departure times (both real and scheduled), at a number of stations and control points along the line.

#### 6.2.1. Interviews

To learn about the operations of Malmö Marshalling Yard, the project attended two workshops carried out by Tillsammans för tåg i tid (TTT). TTT is a long-term industrial collaboration to improve the punctuality of railway traffic. Three interviews were also carried out with IM staff working with various aspects of yard planning and dispatching. The details of the events are given in Table 1.

### 6.3. Results

In this section, we will present the results regarding (i) regularity, punctuality and precision, (ii) deviations from the scheduled channels, (iii) precision and further deviations, (iv) delay risk across trains, (v) delay risk and geography, and (vi) delay risk and time.
Table 1: Workshops and interviews, Malmö Marshalling Yard

<table>
<thead>
<tr>
<th>Description</th>
<th>Participant</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTT workshop focused on the arrival yard operations and the multi-modal terminal</td>
<td>Peter Hysing, Green Cargo</td>
<td>2019-09-24</td>
</tr>
<tr>
<td></td>
<td>Mikael Lilja, Hector Rail</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thomas Frej, Mertz</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Niclas Johansson, Sweco</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Eddie Ljungwe, Sweco</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nima Ghaviha, RISE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sara Gestrelius, RISE</td>
<td></td>
</tr>
<tr>
<td>TTT workshop focused on wagonload traffic</td>
<td>Peter Hysing, Green Cargo</td>
<td>2019-10-10</td>
</tr>
<tr>
<td></td>
<td>Benny Mårtensson, Green Cargo</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mikael Lilja, Hector Rail</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Göran Edholm, Trafikverket</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kirsten Seeger, Trafikverket</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Eddie Ljungwe, Sweco</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sara Gestrelius, RISE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nima Ghaviha, RISE</td>
<td></td>
</tr>
<tr>
<td>Interview with Göran Edholm</td>
<td>Göran Edholm, yard and timetable planner, Trafikverket.</td>
<td>2019-10-11</td>
</tr>
<tr>
<td></td>
<td>Peter Olander, strategic planner, Trafikverket.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sara Gestrelius, researcher, RISE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Martin Kjellin, researcher, RISE</td>
<td></td>
</tr>
<tr>
<td>Interview with Peter Olander</td>
<td>Peter Olander, strategic planner, Trafikverket.</td>
<td>2019-10-11</td>
</tr>
<tr>
<td></td>
<td>Sara Gestrelius, researcher, RISE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Martin Kjellin, researcher, RISE</td>
<td></td>
</tr>
<tr>
<td>Interview with Mattias Green</td>
<td>Mattias Green, instructor at Trafikverket.</td>
<td>2019-10-11</td>
</tr>
<tr>
<td></td>
<td>Sara Gestrelius, researcher, RISE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Martin Kjellin, researcher, RISE</td>
<td></td>
</tr>
</tbody>
</table>

6.3.1. Regularity, Punctuality and Precision

**Regularity** we have defined as the train completing its journey, in full, following the scheduled route. If a train has been cancelled fully or partially, it is given a regularity value of 0, while a train that fully completes its journey (regardless of any delays) is given a value of 1. If a train is rescheduled, and then completes its new schedule, the data will have one row with a value of 0 (the cancelled train) and another with a value of 1 (the newly scheduled train), for an average of 0.5. We find an average regularity of only 66%. As many as 34% of freight trains are thus cancelled or rescheduled, with a large variation (20 to 45%) from year to year.

**Punctuality.** Punctuality on the five-minute level to the yard, for the freight trains that run, is on average about 70%. This is a full 20% lower than the average across all trains, and 25% lower than the industry-wide target across all trains. From year to year it has varied from 62% to 81%, see Table 1. Assuming that freight trains make up 15% of all services (Trafikanalys, 2018a), raising
their punctuality level to the current average of 90 % would increase the total level by three percentage points.

**Precision.** As seen in Table 1, only 45 % of the freight trains fulfil one (major) operator’s operational target of deviating at most 15 minutes from the timetable – either ahead of or behind schedule (Gjerdrum, 2019). Only 13 % satisfy the Swedish Transport Administration’s tighter channel of ±3 minutes. These trains clearly deviate significantly from the timetable. This measure considers all stations and timing points along the journey, for a series of either 1, if the train arrives within the threshold, or 0, if it does not. The channel precision is then the average of all these values.

**Table 2: On-time performance statistics for freight trains between Malmö and Hallsberg.**

<table>
<thead>
<tr>
<th>Year</th>
<th>Trips</th>
<th>Regularity</th>
<th>Punctuality</th>
<th>Channel precision, narrow (± 3 min)</th>
<th>Channel precision, broad (± 15 min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>5 269</td>
<td>64%</td>
<td>81%</td>
<td>11%</td>
<td>41%</td>
</tr>
<tr>
<td>14</td>
<td>4 775</td>
<td>60%</td>
<td>71%</td>
<td>15%</td>
<td>50%</td>
</tr>
<tr>
<td>15</td>
<td>4 686</td>
<td>70%</td>
<td>62%</td>
<td>14%</td>
<td>47%</td>
</tr>
<tr>
<td>16</td>
<td>5 188</td>
<td>55%</td>
<td>62%</td>
<td>13%</td>
<td>44%</td>
</tr>
<tr>
<td>17</td>
<td>4 939</td>
<td>80%</td>
<td>73%</td>
<td>13%</td>
<td>45%</td>
</tr>
<tr>
<td>Total</td>
<td>24 857</td>
<td>66%</td>
<td>70%</td>
<td>13%</td>
<td>45%</td>
</tr>
</tbody>
</table>

**6.3.2. Departure Punctuality**

![Figure 11: Illustration of the deviations at departure from yards.](image)

A large part of the deviations occurs already at the point of departure from the yard. This is illustrated in Figure 11, where the vertical axis illustrates the cumulative share of departures, and the horizontal axis the deviation from the schedule at the originating yard. About 65 % of these departures are ahead of schedule, and more than a quarter of trains leave at least 15 minutes before schedule. Very large deviations of an hour or more happen about 8 % of the time. The
distributions are similar across directions, with the southbound trains being slightly shifted to the left – towards earlier departures.

There is no consensus on whether the early departures shown in Figure 11 are good or bad for the overall traffic situation. Unplanned, surprising interactions and conflicts are stated as reasons for why early departures might be bad for the overall traffic situation, while increased flexibility, more buffer times and freeing up later capacity channels are stated as reasons for why early departures might be good for the overall traffic situation. Several reasons for early departures were suggested in the interviews and workshops conducted with staff working with MGB (see Table 1). They are collected on the list below.

1. The driver calls and wants to depart early (because he or she wants to go home) (TTT, 2019a; Olander, 2019).
2. Force of habit - early departures are part of normal practice (Olander, 2019).
3. There are no rules stating that trains should not run early (Olander, 2019).
4. Free up a channel that can be used by some other later train (Green, 2019).
5. Flexibility - Freight trains are often put aside to solve conflicts that occur during operations (i.e. passenger trains tend to be prioritized in case of operational conflicts). If a freight train departs early these rectifying actions are less likely to make the train delayed at its final destination (Green, 2019).
6. Solve problems – e.g. if a train cannot run as fast as planned then departing early may reduce the delay to its final destination (Green, 2019).

6.1.1. Deviations from the Scheduled Channels

*Ahead of schedule.* In Table 3 we see that as many as 60% of departures occur ahead of schedule, on average by 31 minutes. This is explained in equal parts by faster than scheduled run and dwell times. 38% of run times are completed faster than scheduled, on average by 1.7 minutes. Table 3 also shows that stops are less common, but as many as 73% of these are faster than scheduled, on average by 12 minutes.

*Behind schedule.* 37% of (line segment) departures occur an average of 44 minutes behind schedule. As we see in Table 3, more than 80% of the freight train delays occur on run times: about 30% of the run times are extended, by an average of 3.4 minutes each. The number of stops
is much lower to begin with, and the proportion that are delayed is also lower, at 22%. When they are delayed, however, the size is much larger – on average 13 minutes.

*The share of observations* is simply the share of observations in the studied data that occurred on a given hour, month, line or station, etc. It offers a baseline for the amount of delays that can be expected, if they were distributed proportionately.

**Table 3: Deviations from scheduled channels for freight trains between Malmö and Hallsberg.**

<table>
<thead>
<tr>
<th>Activity</th>
<th>Observations (approximate)</th>
<th>Proportion</th>
<th>Average time or deviation (min.)</th>
<th>Share of deviations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Early/ Shorter</td>
<td>Late/ Longer</td>
<td>Scheduled</td>
</tr>
<tr>
<td>Departures</td>
<td>1 562 000</td>
<td>60%</td>
<td>37%</td>
<td>-31.1</td>
</tr>
<tr>
<td>Run time</td>
<td>1 562 000</td>
<td>38%</td>
<td>30%</td>
<td>5.2</td>
</tr>
<tr>
<td>Dwell time</td>
<td>116 500</td>
<td>73%</td>
<td>22%</td>
<td>13.5</td>
</tr>
</tbody>
</table>

6.1.2. **Precision and Further Deviations**

Table 4 illustrates the relative probability of further deviations based on current deviations. For instance, it shows that the probability of delays is 30% higher for trains that are behind schedule, and 22% higher for those that are ahead of it, compared to trains that are on time.

**Table 4: Relative probability of further deviations, based on current deviations.**

<table>
<thead>
<tr>
<th>Deviation from timetable</th>
<th>Share of observations</th>
<th>Rel. prob. of going faster</th>
<th>Rel. prob. of going slower</th>
</tr>
</thead>
<tbody>
<tr>
<td>On time</td>
<td>3%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Behind</td>
<td>37%</td>
<td>129%</td>
<td>130%</td>
</tr>
<tr>
<td>Ahead of</td>
<td>60%</td>
<td>211%</td>
<td>122%</td>
</tr>
<tr>
<td>Total or average</td>
<td>100%</td>
<td>177%</td>
<td>124%</td>
</tr>
</tbody>
</table>

The pattern is similar when considering deviations on the other end – performing run and dwell times faster than scheduled. Table 4 shows that trains that are already ahead of schedule are more than twice as likely to run faster than scheduled, than trains that are precisely on time. In both cases, the relative probability of further deviations increases as the width of the channel expands, most rapidly for trains on the early side. Unfortunately, the train is exactly on time only during about 3% of the observations – this is thus the exception rather than the rule.

6.1.3. **Delay Risk across Trains**

Naturally, the number of deviations varies across trains. Some experience only a few, while others have many of them. Regarding run times, this distribution can be interesting to study. If trains have
many such deviations, often moving either faster or slower than scheduled, this suggests a more systematic mismatch between the train and the calculated run times. Possibly because the train is lighter than expected, or because there are problems with the breaks, such that the train must run at a lower speed (Gjerdrum, 2019; Gjerdrum & Andersson, 2019).

The typical journey between the yards in Malmö and Hallsberg has 69 line segments, each with a separate run time. Figure 12 shows the distribution of deviations – both faster and slower – along this route. It illustrates that a quarter of the trains experience delays on 22 or more of these line segments. The average run time extension is 3.4 minutes, so these trains have run time delays of about 75 minutes and upwards. For these trains, we can suspect that there is some sort of mismatch regarding the run time calculation. Half of all trains have delays on more than 16 of the 69 sections, while only a quarter of them experience ten or fewer run time delays – corresponding to about a 35 minutes of run time delays or less. For shorter than scheduled run times, the distribution is more even, and shifted more to the right – these are more common than the delays, and the deviations are smaller, averaging at 1.7 minutes each. This also indicates a mismatch with regard to the calculated run times.

Figure 12: The number of run time deviations per train departure for freight trains between the yards in Malmö and Hallsberg.

*Run time and dwell time delays* have been defined as the difference between the realized and scheduled (run and dwell) process times, when these deviations are positive. In this way, we see when and where the delays occur in the first place, rather than observing the delays that have been accumulated over longer times and distances (Palmqvist 2019). All of these run and dwell time extensions have been summed up to make up the total delays. The share of delays is then the share of those total delays that occurred on a given hour, month, line or station, respectively. The same has been done for negative deviations, when run or dwell times have been performed faster than scheduled.
6.1.4. Delay Risk and Geography

The probability of deviations varies across the geography of the line. Some parts are much more likely to see delays – and the reverse – occurring than others. This is easiest to illustrate by considering run and dwell times separately.

Across sections. We begin with run times, which make up about 82% of all delays and 50% of the recovery (see Table 2). Figure 13 illustrates how the probability varies, moving from the north to the south. Deviations, and particularly delays, are larger than usual at the edges, moving in and out of the yards. We also see a spike in delays at the edge of Skåne, coming into Hässleholm from the north. Considering faster than scheduled run times, there are spikes in Nässjö, Sösdala and Hjärup. These peaks can help identify where the scheduled run times should be adjusted. Figure 13 illustrates how the risk varies, moving from the north to the south. Deviations, and particularly delays, are larger than usual at the edges, moving in and out of the yards. We also see a spike in delays at the edge of Skåne, coming into Hässleholm from the north. Considering faster than scheduled run times, there are spikes in Nässjö, Sösdala and Hjärup. These peaks can help identify where the scheduled run times should be adjusted.

Across stations. According to Figure 14, the risk is highest by far in Nässjö, where there is a change of personnel. It alone contributes almost 60% of all dwell time delays along the line, and the relative risk of delays there is about 35 times higher than the baseline. The risk is also elevated at a few other stations, such as Mjöby, Osby, Hässleholm and Vätteryd, but by a more ordinary magnitude of about two to four times the baseline.

Figure 13: Relative risk of run time deviations for freight trains between Malmö and Hallsberg.
6.1.5. Delay Risk and Time

Across the day. The risk of delays varies across the day. Figure 15 considers both the size and frequency of both run and dwell time delays. We see that the risk is elevated both between 6 and 9 in the morning, and between 14 and 21, with a peak around 16. At most, the difference in delay risk is about ±40%. The pattern closely follows that of the number of passenger trains in operation.

Figure 15: Relative risk of delays across the day for freight trains between Malmö and Hallsberg.

Across the year. The variation is smaller across different months, about ±7%. Figure 16 illustrates how the peaks occur in January, May and September, while the troughs are in March, August and December. This sine-curve pattern is more difficult to explain than the variation across the day, but may also relate to the use of capacity, including engineering works.
Figure 16: Relative risk of delays across the year for freight trains between Malmö and Hallsberg.

6.5. Discussion and Conclusions

We have seen that freight trains in Sweden do not adhere very strictly to timetables. To begin with, there is a high rate of cancellations and rescheduling, with only about two thirds of journeys being completed as scheduled. The level of punctuality, for the trains that run, averages 70% with an arrival delay of less than six minutes. The channel precision is much lower still: 13% of trains stay within ± 3 minutes of the scheduled slots, and only 45% stay within a more generous ± 15 minutes. A large part of these deviations begins right away, at the yards on either end. About 65% of departures from the yards are ahead of schedule, more than 25% by more than 15 minutes. Large deviations of an hour in either direction occur about 8% of the time. In summary, the findings reinforce the notion that freight trains do not adhere closely to timetables.

Amongst the delays that occur, over 80% of them occur when trains run between stations. These delays are relatively small, on average 3.4 minutes, but are very frequent. This differs substantially from passenger trains – where most delays occur at stations and are smaller still (Palmqvist, 2019). The probability of delays is slightly lower when a train is ahead than behind schedule, but much lower yet when it is on time. The difference in delay probability between trains running early and trains running on time might be explained by both operational conflicts that would not have happened if the train was on time and by early freight trains being “easier” to slow down or put aside to solve operational conflicts.

There are some variations across time as well, although these are relatively small. The probability of delays is higher than average in some parts of the day (± 40%) and year (± 7%), and lower in others. Most likely this has to do with the amount of other trains and maintenance works. There are much larger variations across the geography, both across lines and across stations. The station of Nässjö sticks out very clearly, with a probability of delays that is more than 35 times higher than average – likely due to a change of personnel there. The variation is smaller between line sections, but on some the probability is still up to three times as high as normal. The same goes for trains travelling faster than scheduled, with apparently systematic geographic differences. This suggests
that there are at least a few sections where the run times should be adjusted either up or down.

Another interesting pattern is that both shorter and longer run times seem to be clustered around some trains. Almost 30% of the studied trains are delayed on at least 20 (out of 70) line segments. On average, this translates to well over an hour of run time delays. Correspondingly, about 15% of trains run faster than scheduled on more than half of the line sections. Both these instances suggest that the trains are probably not operating under the same conditions that were assumed when their timetables were created. Less extreme cases following the same patterns are very common.

The overall picture is that freight trains in Sweden do not seem to adhere to timetables to any greater extent. They are released into the network from the yards with large deviations, and if anything, these deviations tend to grow along the way. Many trains run much faster than scheduled, while others are much slower.
7. Capacity Planning for Engineering Works and Maintenance

This chapter consists of two parts. First, we discuss train path adaptations due to engineering works and maintenance, the underlying reasons for these adaptations and the possible rescheduling strategies that can be used. Secondly, we take a closer look at the process of booking possessions, including possible release of maintenance windows, and discuss possible improvements and the effects of different time settings in this process. The purpose is to give a better understanding of the close interaction between traffic and maintenance planning and to pave the way for further quantitative analysis and planning support for the coordination of these activities.

7.1. Train Path Adaptations due to Engineering Works and Maintenance

Every engineering work or maintenance activity that requires a possession or access to the railway infrastructure will inflict some sort of operative restrictions for the train traffic. Consequently, there is usually a need to reschedule the timetable to adapt the train paths, both before, during and after the actual work activity takes place. In this section we will describe the causes for these adaptations and some different rescheduling strategies. Adjustments and groupings of engineering works also take place, in order to reduce the impact on the train traffic, but such possession adaptations are not discussed here.

7.1.1. Causes for Train Path Adaptations

Firstly, there are some primary capacity restrictions that are caused by engineering works, namely:

- **Closure of the work section.** The boundaries of the closed section usually coincide with signalling positions but can also be at specific track positions. Whenever the electrical power supply must be switched off, a larger area is affected, corresponding to the boundaries of the power feeding system. The closed area is normally not possible to use for train operations. Furthermore, the closed section cannot be used as safety area for adjoining track sections, which means that the usage of connecting tracks may become impaired or impossible. Particularly for closures of switches there may be a cascading effect that affects several connecting tracks.

- **Single track operation.** On double- or multi-track lines, one may choose to let one track be open for traffic, to not completely close the line during the work activity. Thus, opposing train flows will interact which requires some trains to wait for meeting ones to pass.

- **Speed restrictions.** If the work activity only closes a subset of the parallel tracks and trains can pass on the adjoining tracks, there will usually be a speed restriction for such train passages. Speed restrictions may also apply before or after the work activity (both spatially and temporally), for example due to track degradations, for ballast settling after the work or if the work is not fully completed.

- **Occupancy by engineering vehicles and work train movements.** The engineering works may involve some work trains or engineering vehicles that must be moved to the work site.
These movements can be treated as normal trains if they fulfil normal safety regulations, but may also need extra safety precautions, e.g. when it’s not guaranteed that the track circuits will detect the vehicles, or if they can change running direction. Before and after the actual work activity such engineering vehicles may need to be parked on side-tracks close to the work site, which will reduce the track capacity for regular trains.

Secondly, there are some secondary capacity restrictions caused by engineering works, such as:
- **Signalling limitations.** Some tracks may have a preferred running direction, with reduced signalling in the other direction. Furthermore, it is customary to have shorter signal blocks on double track lines than on single track lines – in order to enable shorter headways between following trains. A side effect of this is that when double track lines are operated as single track, the turning of the line direction may take longer (since they involve more signal blocks) – which should be considered in the rescheduled timetable as an increased separation between opposing trains.
- **More crossing and diverging switch movements.** The allowed speed is normally lower through the diverging direction of a switch. Track closures usually result in more diverging train paths and thus reduced speeds. Note that these speed reductions apply for the full extent of the train length, which means that longer (freight) trains will be affected more than shorter (passenger) trains. Furthermore, the location of meeting or overtaking tracks can suit one train running direction better than the other, which means that the closure of one (out of two or more) line tracks may impact one train direction more than the other.

Finally, the engineering works can have some indirect effects on the traffic, such as:
- **Increased sensitivity to disturbances.** Depending on how the traffic is rescheduled, the capacity utilization may increase around the work possession, which will make the planning less robust. To counter this effect, it might be necessary to increase margins or impose additional timetable construction rules while the work is carried out, in order to compensate for this (i.e. to reduce the extra capacity usage) and to obtain an appropriate robustness.
- **Network and chaining effects.** A work possession will, together with the rescheduled traffic, impact the capacity usage both in the upstream and downstream traffic direction. This can give opportunities for synchronization with other work tasks (in the shadow of the main possession) but can also cause unwanted chain reactions which require further adjustments of the train traffic.

### 7.1.2. Adaptation Approaches

There are various approaches for adapting the traffic to engineering works, which we can divide into two groups. In the first group, the geographical route of the train paths remains untouched. These approaches include strategies ranging from simple addition of margins, over platooning of trains, to detailed rescheduling of the timetable. In the second group, the geographical routes are changed. These strategies include rerouting, short turning, train coupling etc. Finally, train paths can be cancelled or replaced with other transportation modes (buses, trucking, shipping).

---

1. In Sweden, the standard is that all tracks have bi-directional signaling.
In the following we will discuss these approaches in more detail, but first we give a quick survey of related publications. We group the papers into tactical and operational planning and present them in chronological order. The papers are summarized in Table 5 and Table 6.

Table 5: Tactical scheduling publications

<table>
<thead>
<tr>
<th>Publication</th>
<th>Infrastructure type</th>
<th>Possession type</th>
<th>Adaption approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vansteenwegen et al. (2015)</td>
<td>Station</td>
<td>Multi-day closure</td>
<td>Detailed rescheduling, rerouting</td>
</tr>
<tr>
<td>Lidén &amp; Joborn (2017)</td>
<td>Aggregate network</td>
<td>Short maintenance windows</td>
<td>Rescheduling, rerouting, cancellations</td>
</tr>
<tr>
<td>Van Aken et al. (2017a),</td>
<td>Network</td>
<td>Multi-day closure</td>
<td>Rescheduling, short-turning, cancellations</td>
</tr>
<tr>
<td>(2017b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arenas et al. (2018)</td>
<td>Detailed network</td>
<td>Short engineering work</td>
<td>Detailed rescheduling and train movements</td>
</tr>
<tr>
<td>Backman et al. (2019),</td>
<td>Line</td>
<td>Short maintenance window</td>
<td>General margins, platooning, rescheduling</td>
</tr>
<tr>
<td>Backman &amp; Solinen (2019)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang et al. (2019)</td>
<td>Parallel lines</td>
<td>Night-time maintenance</td>
<td>Rescheduling, rerouting</td>
</tr>
</tbody>
</table>

Tactical (re)scheduling has been treated in several papers. Vansteenwegen et al. (2015) study rerouting and retiming of trains at a large station where some tracks become blocked during several days for engineering works. Lidén & Joborn (2017) develop a scheduling model for integrated planning of maintenance windows and train services, which is capable of handling train cancellations, rerouting, rescheduling, capacity reductions, speed restrictions and increasing number of meetings during single track operations. An aggregated approach is used, which is suitable for long-term planning, but does not produce a detailed and conflict free timetable. Van Aken et al. (2017a), (2017b) study adjustments of periodic timetables during multi-day infrastructure closures. Retiming, short-turnings and cancellations are considered, while pre-processing techniques are used for reducing the problem size. Arenas et al. (2018) study the detailed rescheduling of trains past engineering works, where speed restrictions are considered as well as maintenance trains to and from the work site. A microscopic model is used which handles block sections, signal separation and a detailed train movement scheduling. Backman et al. (2019)/ Backman & Solinen (2019) perform a simulation study of different approaches for handling single track traffic past a maintenance work of varying length on a double track line. The strategies of general margins, platooning and rescheduling are studied and compared. Based on the simulated traffic outcome timetable construction rules are devised for when the different approaches are suitable and which headway margins etc. that should be used. The simulation model they use is based on the work presented in Haster (2016). Zhang et al. (2019) schedule overnight trains together with night-time maintenance (3-4 hour long) where the trains can either wait for the completion of the engineering work or be rerouted via the parallel normal speed line network.
heuristic Lagrangian method is used for fitting the overnight trains within an existing train timetable.

Table 6: Operational rescheduling publications

<table>
<thead>
<tr>
<th>Publication</th>
<th>Infrastructure type</th>
<th>Possession type</th>
<th>Adaption approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Louwerse &amp; Huisman (2014)</td>
<td>Line</td>
<td>Partial or full blockage</td>
<td>Rescheduling, cancellation</td>
</tr>
<tr>
<td>Veelenturf et al. (2015)</td>
<td>Network</td>
<td>Major disruption</td>
<td>Transition process, including rerouting</td>
</tr>
<tr>
<td>Zhu et al. (2018)</td>
<td>Network</td>
<td>Multiple blockages</td>
<td>Rescheduling, short-turning, cancellation</td>
</tr>
<tr>
<td>Wang et al. (2019)</td>
<td>Line</td>
<td>Segment blockage</td>
<td>Rescheduling, train coupling</td>
</tr>
<tr>
<td>Van Aken et al. (2019)</td>
<td>Typical structure patterns</td>
<td>Segment blockage</td>
<td>Rescheduling, stop-skipping, cancellation</td>
</tr>
</tbody>
</table>

Train traffic management and operational rescheduling are fields with many publications. Here we only mention a few ones, which address rescheduling due to engineering works or track blockages. Louwerse & Huisman (2014) study a double track line with a partial or full blockage and develop methods for computing a temporary timetable where trains can be rescheduled or cancelled, while considering the rolling stock constraints. Veelenturf et al. (2015) extend this model by considering the whole transition process from the start of the disruption, over the running of a reduced timetable and to the restoration of the normal timetable after the disruption has been cleared. In this model rerouting is considered as a method for reducing the number of cancelled train services. Zhu et al. (2018) treat a similar problem but consider multiple disruptions and use short turnings during the disruption. Wang et al. (2019) study a high-speed line and consider the possibility of coupling trains in order to reduce the traffic load after the disruption as well as the total delays. Van Aken et al. (2019) study segment blockages at three typical infrastructure patterns (double track lines without stops, with stops, and multi-track corridors with stops). For each situation effective dispatching rules are developed mathematically and shown to outperform heuristic rule-based approaches similar to the ones adopted manually by train dispatchers.

We now discuss the different adaptation approaches in more detail.

General margins. A simplistic approach to account for future track blockages is to add scheduling margins to the trains. This will reduce the rescheduling efforts, but the dimensioning of these margins and where to add them is not easily determined. As showed by Backman and Solinen (2019), train traffic simulation can be used as one possible method. General margins will work less well than platooning and detailed rescheduling, especially when the capacity restrictions (speed restrictions etc.) are large. This adaptation method is therefore probably best suited for handling secondary or indirect effects on the traffic.
Platooning. This approach means that trains will be grouped together and run as a platoon past a single-track stretch, before changing the direction and running another platoon in the opposite direction – resembling the common approach with traffic signals at road works. The purpose is to reduce the amount of clearance times between direction changes and to utilize the fact that trains in the same running direction can run closer to each other. However, the approach requires trains to be stacked up on either side of the blockage, which will require further track capacity and therefore affects a larger part of the network, as well as the time required to accumulate trains on either side. Determination of the best number of trains to group together is not easily determined and is further discussed in Van Aken et al. (2019).

Train coupling. This approach is similar to platooning but means that one or more trains are joined together and run as one train, either past a working site or after the engineering work has been completed – see Wang et al. (2019). The purpose is to reduce the number of train paths. However, the coupling procedure may take a substantial time, unless the trains are specifically designed for this (such as the Danish IC3 trains\(^2\)). Also, it must be considered whether to split the trains later on or if they should be kept together for the remainder of their trip (which constrains which trains that can be coupled).

Detailed rescheduling. This approach considers one or more capacity restrictions and the traffic is rescheduled – often using the normal timetable as a solution reference. The rescheduling can either be treated as a tactical problem and solved beforehand (Table 5) or as an operational problem and solved in the real-time dispatching situation (Table 6), which requires decision support in the traffic management system(s). The strength of these methods is that the full extent of the blockage and all primary and secondary capacity restrictions can be considered and that several different rescheduling strategies can be utilized (retiming, skip-stop, re-ordering, etc). The drawback is that an extensive planning effort might be necessary, especially if manual scheduling methods are used. Furthermore, most methods assume that the timing and length of the track blockage is known beforehand, which might not be the case for some urgent repair activities.

Rerouting. By rerouting we mean that trains will pass other lines and stations than used in their original train path in order to reach their destination. This approach means that some of the original stations will not be visited, which affects passengers bound for, or boarding at, these places (if there are commercial stops at these stations). Hence, this method works better for freight trains (when the train weight etc. is not limiting) and for passenger trains with few passengers being affected. Rerouting also means that a new train path is constructed, which involves some administrative handling (new or revised train numbers, information sharing etc.), that might be cumbersome during day of operation. Consequently, this approach tends to work better in the tactical planning situation.

Short-turning. This approach means that the train does not complete the full train path but will stop at some station and either return towards its origin (on an existing or new train path) or be moved to/parked at some yard before being rescheduled for some other service. As with rerouting this will affect passengers who want to travel beyond the short-turning station, which might be

\(^2\) https://en.wikipedia.org/wiki/IC3
resolved by waiting for a later train or by taking another travel mode (bus, taxi). Short-turning is less common for freight traffic, since wagons or cargo must be parked at specific yards and later picked up by other train sets or carriers.

Cancellations. To cancel a train means that no part of its train path will be utilized. This is a drastic measure that will be necessary when major disruptions or large engineering works take place. On the other hand, it can sometimes be beneficial to avoid having cargo or passengers begin their trips – if they are better catered for at their origin.

Replacements with other transport modes. This approach may become necessary because of rerouting, short-turning and cancelling of trains. For passengers, the alternate mode is usually buses, public transport (subway, commuter trains), taxi or cars. For freight it may be possible to switch to trucking (or in rare cases shipping or air cargo), but in most cases the only feasible or reasonable option is to delay the transports.

With this listing we have tried to cover all the traffic adaption approaches that may be considered when engineering works or maintenance activities are needed on a railway infrastructure. As discussed, the specific network, traffic situation and transport requirements will determine which of these that will be appropriate to use.

7.2. Possession Booking and Release of Maintenance Windows

In this section the actual booking of possessions within given maintenance windows will be discussed and how the non-booked parts are handled. Specifically, the effects of having an early or late release time for non-booked maintenance windows will be discussed. The material is based on interviews with representatives from the Swedish Transport Authority (Edholm, 2019a) and the major freight RU (Gjerdrum & Andersson, 2019) in Sweden.

First, we describe the process from the IM’s and maintenance contractors’ point of view. Then we study the planning process of a freight RU. Finally, some conclusions are drawn, followed by a discussion of these results and aspects to consider for future continued research.

7.2.1. Possession Booking Process

There are two types of maintenance windows currently in use at Trafikverket (TrV), which we label as “regular” and “dedicated”:
- Regular windows are typically 2 – 4 hours long and recur daily or weekly over the whole timetable period. Booking of possessions within regular windows should be done 12 – 4 weeks before the activity takes place. Unused window time could be released 4 weeks before day of operation, but the current practice is to keep them until 1 week before, to enable late bookings, handling of exceptional situations etc.
- Dedicated windows are usually longer (typically 4 – 6 hours), grouped together, and scheduled into concentrated track work campaigns. These windows are handled in the
same way as large engineering works and major possessions, which means that the final decision is taken in the revision planning (4 times per year) and that the necessary traffic adaptations should be done 18 weeks before day of operation, and that the contractor should have done their final possession booking 4 weeks ahead of time. Even so, the release time for dedicated windows is also set to 1 week before day of operation.

Unfortunately, there is a lack of reliable statistics regarding how early or late the possession bookings take place. Furthermore, there are several different systems involved in the planning and reporting of work tasks which hampers the data quality. Specifically, the mapping between maintenance windows and possession bookings is not working well (Johansson & Edholm, 2019). Hence it has not been possible to establish how often possession bookings within maintenance windows are made in the time frame 4 – 1 weeks before execution. Planners at TrV estimate that most bookings within regular windows take place 6 – 4 weeks before the activity date. As for dedicated windows, it should be possible to book the possessions 18 – 12 weeks before execution and to release them at that time.

From the maintenance contractors’ point of view there are roughly two types of work tasks – those that are plannable, and those that arises with a short notice (urgent tasks):

- Plannable tasks can be handled either with regular windows with a booking time of 18 – 4 weeks (short activities), or dedicated windows (long activities).
- Urgent tasks are usually triggered by inspection remarks, track problems detected by measurements (such as track alignment, cracks found by ultra-sonic testing etc.) or weather conditions (snow, frozen ground, heat, wet leaves etc.). These must be handled with a one to two week notice or in the operational phase. For short tasks there is usually sufficient spare capacity but occasionally urgent train path adjustments are needed.

Based on this, we can conclude that the benefit of maintenance windows is high 18 – 4 weeks before execution, small or none 4 – 2 weeks ahead, and that some spare or reserve capacity is necessary to uphold for urgent tasks 2 – 0 weeks before day of operation. TrV is currently investigating the possibility to adjust the planning procedures to release unused regular windows and dedicated windows 4 weeks and 18 – 12 weeks before execution, respectively (project SF2.0). In addition, the possibilities for direct work is being investigated, as a method for increasing flexibility and reducing the administrative burden.

### 7.2.2. Freight RU Planning Process

Green Cargo is the main freight train operator company in Sweden. Their planning process is representative for a large company\(^3\), while smaller RUs may have a less structured and perhaps quicker planning approach. Green Cargo operates roughly 400 train paths during a normal production day, of which about 100 are long-distance wagon load trains, 100 are dedicated system trains and 200 are terminal trains (moving wagons between junctions and industry terminals). A small amount of repositioning movements is also done.

About 3% of the train paths are cancelled after the monthly operative plan has been published. Primarily this is due to terminal trains that do not need to run (no wagons to transport a specific day). There is also a small amount (~1% in total) of additional wagon load trains, which are planned 1 – 2 weeks before operation. The reasons can be that some yards become saturated with too many wagons (which must be moved elsewhere to make room for new arrivals) or that transport order volumes exceed the available train capacity (which requires additional trains to satisfy the delivery requirements). During 2019, between 0 and 30 such additional trains have been necessary per week.

A monthly planning cycle is used for train paths, rolling stock circulations and crew schedules. Train path modifications due to operative reasons (new, changed or cancelled trains) are submitted to TrV 10 – 9 weeks before the planned month (bpM), so they can be approved and decided about 7 weeks bpM. Then the rolling stock circulations are optimized and handed over to the crew scheduling phase 5 – 4 weeks bpM. Finally, the crew schedules are constructed, reviewed and approved, to be published 2 weeks bpM. All the resource plans (train paths, rolling stock, crew) are maintained and adjusted continuously after these handover/publication times, which requires a certain resource buffer (locomotives, wagons and crew).

Train path adaptions due to engineering works foregoes the normal monthly process, since TrV needs these adaptations 18 weeks before the revision period. If applied strictly, this would require the RUs to revise their train paths for the first month of the revision period about two months earlier than they would do their operative train path adjustments. In practice, TrV holds a revision planning meeting four times per year, where the planned changes in maintenance windows and major possession are presented, after which the RUs will decide on the consequences and when to submit the resulting train path adjustments.

From the RU’s point of view, maintenance windows need to be released about two months (9 weeks) before the planned month in order to be useful in the normal planning process and enable resource efficient train paths. Windows that are released 2 weeks before day of operation could be useful for the small amount of additional wagon load trains that has been mentioned above. However, given the small amount of such trains and that window releases will probably happen relatively seldom, it is unlikely that they will be possible to match geographically and temporally in a way that will have a substantial impact on train operations.

Unfortunately, it has not been possible to obtain an estimate of how much the resource efficiency or operative cost differs between normally planned and resource efficient trains and lately adjusted / additional trains. Research literature indicates that savings in the order of at least 10% are possible when comparing manual and optimized rolling stock plans for passenger trains (Fioole, Kroon, Maróti, & Schrijver, 2006). This gives a hint of the value for releasing unused maintenance windows before the normal train path modifications start at the RUs.

7.2.3. Conclusions
The first observation is that the planning processes for maintenance and trains differ regarding planning cycles and in how the deadlines and handover times are constructed. The maintenance
planning uses a quarterly process (revision handling), followed by a rolling 8-week period and with deadlines at a certain number of weeks prior to the planned activity. The train planning (at the major RUs) uses a monthly process, followed by continuous adjustments, and use deadlines that are measured as the number of weeks before the start of the planned month of train operations.

The obvious question is why these processes are not better aligned. It seems natural to transform the quarterly revision process into a monthly planning cycle with deadlines measured towards the start of each month also for the infrastructure maintenance. That would make it easier to understand how the processes connect to each other, enable better data exchange procedures between the different planning tools and hopefully ease the burden on planners. In such a setting it is still perfectly possible to handle quarterly updates if wanted, or to use longer planning horizons if needed (for example for passenger trains).

The second observation is that the changes that TrV is considering in the handling of maintenance windows (releasing regular and dedicated windows 4 and 18 – 12 weeks respectively before execution) seem appropriate. However, it would be even better if a monthly planning was used. Windows that are released more than 10 weeks before the planned month would then become beneficial for normal trains, while those that are released 4 weeks or later can primarily be useful for technical transports, locomotive & wagon repositioning, and engineering vehicles & work transports.

Finally, it can be noted that there are several incentives and fees regarding train path reservations and cancellations, while there are fewer such incentives regarding maintenance windows and possession bookings, and their respective releases and cancellations. For example, it seems natural to consider fees for making possession bookings outside maintenance windows, especially if they affect (or are close to) train paths, or to charge the contractors for unused possessions – possibly depending on their size and/or path adjustment costs. (One such refund regarding cancelled track works is already regulated today but is not used in practice due to administrative reasons.)

7.2.4. Discussion

As shown above there are three planning processes going on at the same time (at the IM, the contractors and the RUs), which are coordinated and synchronized at certain points of time. Thus, the value of booked/released window time will be strongly connected to these synchronization times and will likely exhibit a stepwise behaviour rather than vary continuously over time. This property must be considered if a quantitative evaluation model for booking and release of maintenance windows is to be devised.

During this study we have considered the possibility to quantify the benefits of different release time settings – both for RUs, contractors and train dispatching, based on actual capacity planning data and statistics. Methods from transport economics could be suitable for establishing metrics and measuring the (socio-economic) value of track capacity made available at different time horizons. It has however not been possible to collect the necessary statistical data for doing such a study, neither from the IM nor the RU side. The possibility for doing a stated preference study
with affected RUs and/or maintenance contractors has also been considered. This could be an interesting research topic for the future but has been judged as too ambitious within the given scope of the project.

Finally, it is not clear which measure to use when evaluating and comparing the effects for the different stake holders. Changes in operation costs, track quality and delay statistics are obviously the most interesting end results, but the causal effects will likely be hard to establish. Thus, it might be necessary to instead use some proxy or linked secondary effect. A closer analysis of such possible metrics, and their cause-effect relationship, has not yet been performed but is an interesting subject for future research. The willingness among different companies to release data needs to be considered in such work. From a process and planning efficiency point of view it could also be of interest to use the amount of planning rework and adaptions as a metric when comparing the effect of different time settings and planning deadlines.
8. Real-Time Network Management

Real-time railway network management encompasses a large range of processes and activities, and involves multiple systems. It includes e.g. monitoring the power grid, the signals and the interaction on the tracks, allocating train paths and dispatching individual trains, coordinating train connections and informing passengers as well as ensuring that the maintenance of the infrastructure can be carried out safely. How these activities are carried out depends on the context and the national organization of railway traffic and services, which varies a lot. Furthermore, according to the Rail Net Europe glossary (RNE, 2017), the terminology used in the national network statements also varies between countries.

In this document, real-time network management refers to the management and operational control of the train traffic (1) within and (2) between the different network hubs. We do not consider the constraints and activities associated with the management of the power grid, nor do we explicitly consider maintenance activities other than that we acknowledge that (un)scheduled track maintenance reduces the network capacity in a specific way during a certain time period.

We also distinguish between two main activities:

- Real-time line management
- Real-time yard management

**Real-time line management** includes here the management and operational control of the train traffic in the overall network, but it does not include handling the scheduling and dispatching of trains into and inside the marshalling/shunting yards. **The yard management** is often in practice, as well as in research literature, considered as a separate, challenging activity, but the available capacity in the yard(s) is directly affecting the incoming and outgoing traffic flows on the lines, and vice versa. Hence, if the yard is a capacity bottleneck it may hamper the throughput and on-time performance of trains on the associated and adjacent railway lines. So, when the capacity on either side is limited, the yard management and the line management need to be coordinated somehow to achieve a desirable system performance. For reasons of convenience we do, however, start by describing these two activities separately.

8.1. Traffic Control and Line Management

As previously mentioned, the management and control of train traffic depends on how the railway sector and market in each country, or region, is organized. Since there have been several larger railway traffic-oriented EU-projects in the past, e.g. AMORE, ARRIVAL, ON-TIME and ARCC - which have surveyed the train traffic control processes and tasks time and again - we refer to those deliverables for excellent descriptions and overviews, see e.g. ON-TIME (2011) and ARCC (2019). The organization and processes concerning the management and control of train traffic have not changed that much since the mentioned surveys. However, the on-going digitalization and desired
incorporation of new efficient integrated, support systems, forces the national rail administrations to define and evaluate alternative future processes for traffic management, where the technology can support and enhance the skills of the human experts.

Several countries are now preparing to deploy new railway Traffic Management Systems (TMS), which integrate, unify and automate a large part of the traffic management. One example is the Swiss initiative Smartrail 4.0 and its six sub-programs aiming for a test phase during 2020-2026 and industrialized rollout during 2027-2038. Another example is the Swedish digitalization program, referred to as DAT, which consists of the infrastructure data management project ANDA, the traffic planning system project MPK and the traffic control system project NTL as well as the integration of these sub-systems and relevant legacy systems. The NTL concept is expected to be deployed fully by 2023.

One main advantage of these new systems is the possibility to increase the level of automation for (1) operating trains (ATO, Automatic Train Operation) as well as (2) monitoring and controlling the traffic during normal conditions and during disruptions. For ATO, there are according to UITP (International Association of Public Transport) five grades of Automation (GoA), see Figure 17 and a survey by Wang et al. (2016) concerning the main benefits and expected challenges associated with the deployment of ATO in urban rail transit systems.

<table>
<thead>
<tr>
<th>Grade of Automation</th>
<th>Type of train operation</th>
<th>Setting train in motion</th>
<th>Stopping train</th>
<th>Door closure</th>
<th>Operation in event of disruption</th>
</tr>
</thead>
<tbody>
<tr>
<td>GoA1</td>
<td>ATP* with driver</td>
<td>Driver</td>
<td>Driver</td>
<td>Driver</td>
<td>Driver</td>
</tr>
<tr>
<td>GoA2</td>
<td>ATP and ATO* with driver</td>
<td>Automatic</td>
<td>Automatic</td>
<td>Driver</td>
<td>Driver</td>
</tr>
<tr>
<td>GoA3</td>
<td>Driverless</td>
<td>Automatic</td>
<td>Automatic</td>
<td>Train attendant</td>
<td>Train attendant</td>
</tr>
<tr>
<td>GoA4</td>
<td>UTO</td>
<td>Automatic</td>
<td>Automatic</td>
<td>Automatic</td>
<td>Automatic</td>
</tr>
</tbody>
</table>

*ATP - Automatic Train Protection; ATO - Automatic Train Operation

Figure 17: An illustration of the five grades of automation (GoA) concerning train operations, where UTO refers to Unattended Train Operation (UITP, 2018).

The development and deployment of ATO has historically mainly been for metro systems and for dedicated freight lines, which both are rather closed systems with homogenous traffic and often only one operator. Examples are the AutoHaul system (automation level GoA4) for iron ore railway
transports by the Rio Tinto Group in the Pilbara, Western Australia and the Sydney Metro line (GoA4, since 2019).

Since national public railway networks often are more complex and larger than metro systems and dedicated freight lines, and they involve many different actors, the deployment of automation has not been as rapid in this area. Currently there are a few examples of where ATO has been commercially deployed over ETCS (European Train Control System) Level 2 on a main railway line, e.g. London’s Thameslink, which corresponds to GoA2. Apart from the very large investment that is associated with introducing increased automation on a large scale, there are several other significant challenges which are addressed and investigated in particular in the Shift2Rail Innovation Programme 2 (IP2), where the aim is to “develop and validate a standard ATO up to GoA3/4 over ETCS, where applicable, for all railway market segments (mainline/high speed, urban/suburban, regional and freight lines).”

A related development that has focused more on the introduction of computer-based support systems for traditional railway lines, is the deployment of driver advisory systems (DAS). DAS in combination with a computer-aided train dispatching (referred to as C-DAS, connected driver advisory system) is where the real benefits can be found, see (Rao, Montigel, & Weidmann, 2016). This requires that the TMS incorporates intelligent, flexible and semi-automated dispatching systems that are able to proactively identify potential conflicts and assist the dispatchers to resolve them.

In the remaining part of this section we will focus on computational decision support functionality for conflict detection and conflict resolution in railway networks. The associated review of state-of-the-art and state-of-the-practice in Chapter 11 will also have this focus.

For simplicity, we distinguish between three levels of automation and decision-support for real-time railway line management here:

- Automation and decision-support during “normal” conditions.
- Automation and decision-support during “smaller disturbances”.
- Automation and decision-support during “larger disruptions”.

This simple distinction is made in order to initiate a discussion on what kind of support and automation that is motivated in different situations. The distinction requires that we define what constitutes (1) “normal” conditions, (2) smaller disturbances and (3) larger disruptions, including discussing what kind of deviations are within the tolerance interval of each level.

The distinction between “normal conditions” and “smaller disturbances” is particularly interesting. That is, what kind of adjustments of the scheduled traffic in real-time is considered as expected and “normal” and can these adjustments be (partially) automated?
Figure 18 and Figure 19 illustrates the time deviation of different trains on the Swedish Iron Ore Line during May 2014 and January-May 2015. Figure 18 shows that the heavy iron ore train is often both significantly ahead or behind schedule leading to that out of 31 days in May 2014, the scheduled meeting is only realized 16 times.

Figure 19 depicts the on-time performance for four different trains at the station “Torneträsk” on the single-tracked Iron Ore Line in Sweden during the period January-May 2015. Also here, it can be observed that some of the trains deviate frequently from the initial schedule.

The analysis of the data presented in the figures above shows that it is quite frequent that the trains on this line do not follow the initial timetable. Hence the dispatchers spend a lot of their time monitoring the trains and on manually re-arranging meetings, changing route allocation, etc. There are multiple reasons for why some of the trains do not follow their initial schedule and perhaps there is a need for a more dynamic timetable for these freight trains. However, in order to proactively and effectively manage such dynamics on congested, single-tracked lines with heterogeneous traffic, computer-based decision support for dispatchers would be beneficial. The responsible dispatching centre in Boden, Sweden, is equipped with a visualization tool, STEG. STEG allows for some basic conflict detection, but no support for conflict resolution yet. There are, however, plans to improve and enhance the STEG environment which is under discussion currently. The discussion on what deviations that are expected to occur in normal conditions is therefore interesting.
Figure 19: Visualization of the time deviation of four different trains at the station “Torneträsk” on the Iron Ore Line in Sweden during the period January-May 2015. “Pax train” refers to passenger train (Törnquist Krasemann, 2016).

Similarly, it is relevant to define what constitutes a minor disturbance and what constitutes a larger disruption. A disturbance in a railway network can occur due to a smaller incident such as an overcrowded platform and unexpectedly long boarding times causing minor delays. Disturbances can also be more significant and occur due to e.g. weather, rolling-stock breakdowns, power shortages, or signaling system failures.

Larger disturbances are in the context of railway traffic management sometimes referred to as disruptions. The distinction between smaller and larger disturbances has been discussed in Cacchiani et al. (2014), where the following definition is used:

“…disturbances are relatively small perturbations of the railway system that can be handled by modifying the timetable, but without modifying the duties for rolling stock and crew. Disruptions are relatively large incidents, requiring both the timetable and the duties for rolling-stock and crew to be modified.”

Hence, the distinction primarily is based on what type of actions that may be needed to cope with the incident rather than the initial sources of disruption. Furthermore, for larger disruptions the situation can be illustrated by the bathtub metaphor (Ghaemi & Goverde, 2015), see Figure 20.

The illustration depicts that during the first phase, the incident is identified, and an assessment of the required capacity reduction is carried out. During the second phase, a reduced level of traffic is maintained until the system goes back to normal again via a transition plan. Traditionally, the handling of disruptions is mainly done manually. How automation and computer-based decision-support could help in these more complex situations in the future – rather than hamper – is an
important question.

Figure 20: An illustration of the railway system state evolution subject to a disruption, depicted using the bathtub model (Ghaemi & Goverde, 2015).

The research efforts dedicated the last 20-30 years to develop different types of decision support functionalities for railway traffic management is significant. New knowledge and insight on what potential improvements that can be made and what methods there are, have been transferred to practitioners in various ways. The implementation rate of proposed concepts and methods is still modest (Lamorgese, Mannino, Pacciarelli, & Törnquist Krasemann, 2018). Interesting questions to reflect on are therefore:

- To what extent is the industry and the stakeholders able to absorb and make use of the knowledge and innovations generated by the research community?
- Is the focus of the research community addressing the actual needs experienced in practice?
- What is needed to bridge between research and practice?

In order to attempt answering the second question, an analysis of state-of-the-art and state-of-the-practice regarding computational decision-support for real-time railway traffic management has been conducted and can be found in Chapter 9.

### 8.2. Real-Time Yard Management

Real-time yard management includes the operational planning and coordination of the resources and activities within a major marshalling yard in the rail freight network. The marshalling yards are resource intensive entities in the railway network and an efficient operation is crucial both for cost control and in a freight transportation perspective. In particular, the single wagon load transport system is very dependent on the marshalling yards, as wagons have to be transferred from one train to another at the yards to fulfil the transport needs. The marshalling yard resources to be controlled includes both infrastructure, vehicles and personnel. Table 7 summarizes the most important resources at the yards that need to be coordinated.
Table 7: Some resources of the marshalling yards that need to be coordinated in the real-time yard management.

<table>
<thead>
<tr>
<th>Resource</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arrival yard</td>
<td>Infrastructure</td>
</tr>
<tr>
<td>Departure yard</td>
<td>Infrastructure</td>
</tr>
<tr>
<td>Classification bowl</td>
<td>Infrastructure</td>
</tr>
<tr>
<td>Hump</td>
<td>Infrastructure</td>
</tr>
<tr>
<td>Surrounding tracks</td>
<td>Infrastructure</td>
</tr>
<tr>
<td>Wagons</td>
<td>Vehicle</td>
</tr>
<tr>
<td>Incoming trains</td>
<td>Vehicle</td>
</tr>
<tr>
<td>Outgoing trains</td>
<td>Vehicle</td>
</tr>
<tr>
<td>Shunting locomotives</td>
<td>Vehicle</td>
</tr>
<tr>
<td>Personnel</td>
<td>Personnel</td>
</tr>
</tbody>
</table>

The management of the yard have implications on the line operations and vice versa. For example, delayed arrival of some trains may cause a lack of tracks in the arrival yard, which in turn gives implications of the operation of the yard. And vice versa, delayed operations on the yard may lead to delayed departures from the yard. The boarder between line management and yard management is the arrival and departure yards, which is a shared resource, utilized by both line operations and yard operations. Two major roles are identified in the management of line and yard in this setting: The Infrastructure Manager (IM) and the Yard Manager (YM), and these two persons may belong to completely different organizations. Normally, the IM is in charge of both the arrival and departure yard, and the YM makes requests for the utilization tracks at arrival and departure yards.

Freight trains very often run outside their planned timetable slot, they both run before and after the planned slot. Actually, already at the departure from the yard, freight trains are very often early or late. Thus, it is obvious that the yard operations have implications on the traffic on the lines. However, the underlying reasons both to delayed and early departures from the yard are not always obvious. In this project, we will investigate more into the correlation between the yard operations and trains departure times, we will both make a qualitative interview study and a quantitative data analysis. In Section 10, both the qualitative and quantitative studies are outlined.

The real-time yard management has been thoroughly investigated in previous projects. In particular, the in ARCC-project the situation at both Swedish and German marshalling yard were investigated. A special focus was the interaction between the IM and the YM and how a decision-support system could improve the coordination and improve to operation on both the line and the yard (ARCC, 2018a; ARCC, 2018b). The Fr8Rail II WP3 work regarding yard management and coordination with lines and builds upon the results from the ARCC-project.

Rail yard management is also studied in other Shift2Rail-projects. In the project SMART (SMART, 2019), a high-level optimization tool is developed to coordinate many of the most important resources at the yard. In Fr8Hub project (Fr8Hub, 2019a), a predication model for yard delay is
developed. In Optiyard (OptiYard, 2019a), both an optimization model and a simulation model for yard operations are developed, where the aim of the simulation is to be used for verification of the optimization model results. The mentioned projects also have many other activities in addition to the mentioned, please see reports from each respective project. In addition, separate detailed optimization models for yard activities are also developed, see e.g. (Gestrelius, Aronsson, & Peterson, 2017).
9. Computational Decision-Support for Real-Time Train Traffic Management

As mentioned Chapters 3 and 8, the railway industry faces many challenges today. The perhaps biggest challenge concerns the organisational aspects of the complex inter-dependent processes of planning and controlling traffic and maintenance, as well as systematically conducting system performance assessments to act on the lessons learnt. However, the use of computational decision-support provides significant potential to facilitate the management of these processes. In this chapter, an overview of the current state-of-the art and state-of-the-practice in the development and application of computational decision-support for real-time railway network management is presented, followed by some concluding remarks.

9.1. An Overview of State-of-the-Art

As an attempt to survey and analyse the different research streams and aspects on computational decision-support for real-time railway traffic management, we have categorized the research topics based on the four main “components” (see Figure 21) that a computational decision-support system relies on:

i. Information flow requirements and system integration

ii. Problem formulation

iii. Solution method

iv. Visualization and user interaction

9.1.1. Information Flow Requirements and System Integration

Access to sufficient and reliable data is a corner stone for any computational-decision support system. The topic “Information flow requirements and system integration” includes here aspects on e.g. information flows during traffic management, see e.g. (Shipper, Gerrits, & Koppenjan, 2015), system integration and architectures that enable a closed-loop traffic management, see e.g. (Quaglietta, et al., 2016), as well as methods to predict train running times, see (Kecman & Goverde, 2015) and train delays (Oneto, et al., 2016). The interest in aspects on information flows and data mining has grown significantly the past years, in line with the general growing interest in Big Data analytics. See more recent surveys on data analytics for train traffic management, e.g. (Ghofrani, He, Goverde, & Liu, 2018) and (Wen, et al., 2019).

9.1.2. Problem Formulations and Solution Methods

Problem formulation and choice of solution method are strongly connected. Depending on the properties of the specific problem and how it is mathematically formulated, certain solution approaches are more suitable than others. Problem formulation refers here to alternative mathematical models of train traffic and its dynamics as well as the associated railway network
and its temporal constraints - in normal conditions as well as during disturbances and disruptions. An introduction to train rescheduling and train dispatching can be found in (Jacobs, 2008) and (Lamorgese, Mannino, Pacciarelli, & Törnquist Krasemann, 2018). As briefly mentioned in Chapter 8.1, the problems to be modelled look different depending on the context and situation. There are several ways of classifying models and solutions approaches, but generally one can mention some general properties of such mathematical models in this domain:

- Level of granularity {Macroscopic, Mesoscopic, Microscopic}
- Time representation {Discrete, Continuous}
- Objective(s) {Minimization of timetable deviations, Minimization of train delays.}

Figure 21: Illustration of four different thematic topics on computational decision-support for real-time railway traffic management and associated aspects.

Level of granularity refers here to how the infrastructure and its resources are modelled, see e.g. (Radteke A. , 2008), including which details that are relevant to include for the particular application and context. There exists no exact definition of what details each level includes, but typically both macroscopic models and mesoscopic models disregard the railway signalling system and simplify the capacity constraints. Macroscopic models typically consider stations as nodes with infinite or limited capacity, see e.g. (Törnquist & Persson, 2005). Mesoscopic models e.g. the ones proposed in (Törnquist & Persson, 2007), (Meng & Zhou, 2014) and (Bach, Mannino, & Sartor, 2019) typically consider a station as a set of unique tracks and where the lines consists of a set of parallel line tracks, each composed of one or several consecutive block sections, where the trains are required to operate with a minimum time distance between. Microscopic models, such as the one proposed by (Pellegrini, Marlière, & Rodriguez, 2014), typically represent also the track circuits and dynamic
time distance between consecutive trains. Since the selected level of granularity depends both on the access to associated input data and what level of detail that is reasonable to use in the specific case, there is a wide range of different models used by the research community.

Time representation refers here to if the train slots are modelled as a set of consecutive discrete, fixed time windows (e.g. 30 s each), or as a time window defined by a continuous start time variable and a continuous end time variable. The two different time representations have both strengths and weaknesses. Generally, a continuous time representation enables a more detailed representation of the traffic and infrastructure and it is also more appropriate for longer time frames, since the problem size does not grow as fast as with a discrete time representation.

The use of a continuous time representation is therefore most common and examples are presented in e.g. (Mascis & Pacciarelli, 2002), (Törnquist & Persson, 2007), (Sato, Tamura, & Tomi, 2013), (Pellegrini, Marlière, & Rodriguez, 2014) and (Lamorgese & Mannino, 2015). Examples on approaches that apply discrete time representation are presented in e.g. (Harrod, 2011) and (Meng & Zhou, 2014) while surveys and benchmark studies are presented in e.g. (Lusby, Larsen, Ehrgott, & Ryan, 2011), (Harrod & Schlechte, 2013) and (Fang, Yang, & Yao, 2015). With an effective solution method, also the discrete models can be solved within reasonable times for some instances.

Objective(s) refers here to what aspects that are minimized and/or maximized during the real-time rescheduling. During normal conditions and in sparse networks it may be reasonable to minimize deviations from the actual timetable with extra penalties for delays above a certain threshold, while in congested public rail transit networks it may be more effective to minimize secondary delays and passenger delays. More recent surveys and discussions on the different objectives and quality indicators for railway rescheduling in different contexts can be found in e.g. (Samà, Meloni, D’Ariano, & Corman, 2015), (Törnquist Krasemann, Computational decision-support for railway traffic management and associated configuration challenges: an experimental study, 2015), (Corman, Quaglietta, & Goverde, 2018) and (Josyula, 2019).

Solution methods refer here to different approaches used to solve the mentioned problem formulations, i.e. to resolve (potential) resource conflicts during train operations and rescheduling. The types of methods range from the application of different train prioritization schemes see e.g. (Mu & Dessouky, 2013), to the deployment of commercial optimization software (Pellegrini, Marlière, & Rodriguez, 2014), or the application of advanced, tailored algorithms enhanced by parallel computing, see (Bettinelli, Santini, & Vigo, 2017) and (Josyula, Parallel algorithms for real-time railway rescheduling, 2019). There are pros and cons with all mentioned approaches. The priority schemes and rule-based approaches are transparent and easy to understand, and the input data requirement is relatively low, but they tend to be myopic, i.e. short-sighted, and cannot easily account for the system-wide effects and delay propagation. The approaches that use a mathematical model and solve it by commercial software have the advantage that they are quite easy to extend and adjust by adding, removing or re-formulating some constraints and objectives. However, the general-purpose optimization solvers may be too slow to handle the defined problem and hence not always appropriate for some larger instances. The tailored algorithms are on the other hand designed and implemented for a particular type of
problem which makes them fast. It is also easier to enable parallel computations and handle distributed decision-making. However, the algorithms may not be as easy to extend and adapt, and how they work may appear as a “black box” to the user unless the assumptions and suggested decisions are clearly communicated to the human experts.

For a more extended overview of solution approaches, we refer to the recent survey in (Van Thielen, 2019).

9.1.3. Visualization and User Interaction

For future decision-support functionalities to be useful for and accepted by the human train traffic control experts, it is critical that the interaction between the human experts and the computational support system is effective and transparent. That is, a system perspective needs to be adopted (Andreasson, Jansson, & Lindblom, 2019) to ensure that the automation supports the experts rather than hampering their work, as have been observed in earlier studies, see (Golightly, et al., 2013).

What can be observed is that the research stream on computational decision-support and algorithm development for railway traffic management, has not yet been sufficiently merged with the corresponding research stream focusing on aspects of human computer interaction. Several research studies analyse potential objectives that could dictate, or guide, the train rescheduling, and propose quality metrics that may be relevant to assess and used to present the proposed solutions to the dispatchers. Meanwhile there are some projects which analyse the complexity of the mentioned decision-making process and what automation can/cannot do, see e.g. (Jansson, 2017). However, no published study which aim to compare and discuss those objectives, quality metrics and visualization aspects with potential decision-support users, could be found so far. Such studies may very well be conducted within industrial research collaborations related to the development mentioned in the next section, but may then not be available to the public.

9.2. An Overview of State-of-the-Practice

There exist a few examples today of more advanced systems for train traffic control and optimized disturbance management. An overview of state-of-the-practice can be found in (Lamorgese, Mannino, Pacciarelli, & Törnquist Krasemann, 2018) and (Van Thielen, 2019) and a summary will be presented below. The TMS that today deploy, or have deployed, some form of optimization-based real-time train traffic control on main railway lines are the following, according to the best of our knowledge:

i. Regional lines such as Trento–Bassano and Orte–Terontola–Falconara, Italy (2011-ongoing, system provided by Bombardier Transportation).

ii. The regional line Stavanger–Moi, Norway (February-December 2014, system provided by a research team from SINTEF).
iii. Automated railway freight traffic on certain lines, Latvia (2017-ongoing, system provided by Bombardier Transportation).

iv. ICONIS real-time conflict solution algorithm (to be deployed, system provided by ALSTOM).

A discussion of the first three mentioned implementations is found in (Borndörfer R., et al., 2017). The underlying optimization approach for the regional Italian lines is a heuristic incorporating the operative rules of the infrastructure manager. The system is semi-automated, meaning that the dispatchers decide which of the – if any – proposed and ranked solutions are to be implemented. The solution quality is primarily measured in terms of punctuality, i.e. the share of trains that are on-time or within the different delay intervals 5–10, 10–15 or larger than 15 minutes.

The approach deployed on the regional line Stavanger-Moi in Norway uses an exact optimization algorithm incorporating integer programming techniques including Benders’ decomposition and delayed row/column generation. This approach is also semi-automated.

The application in Latvia uses a heuristic which is based on local operative rules. The dispatching decisions concerning the freight trains are automatically implemented while the human dispatcher needs to approve the proposed decisions when solving conflicts involving passenger trains. Since few passenger trains operate on those lines, a large part of the dispatching is fully automated.

Regarding the fourth mentioned system, ALSTOM has together with a team of Italian researchers, developed a real-time conflict solution algorithm for the train rescheduling, to further enhance the optimization functionality in their TMS ICONIS. The algorithm is, as far as can be seen, not put into operation yet but has been experimentally evaluated showing promising results (Bettinelli, Santini, & Vigo, 2017). The algorithm is a greedy heuristic which prioritizes to find a first feasible solution quickly by performing an initial intelligent sorting of the trains. The trains are then reordered iteratively based on several alternative strategies. What is different with this algorithm, compared to most other similar algorithms, is that it during the iterative solution improvement phase does not strictly forbid the violation of hard capacity constraints. Instead such violations are given a very high penalty so that the algorithm promotes solutions that do not include such violations. However, it means that the algorithm may propose infeasible solutions, but then the violations are clearly presented to the dispatchers. The algorithm tries to minimize the costs associated with several aspects such as time deviations, missed connections and resource conflicts, and the pool of solutions are ranked accordingly. The approach also makes use of parallel computing to decrease the computation times. The algorithm provides solutions of good quality (Bettinelli, Santini, & Vigo, 2017), and delivers solutions within 2 seconds for an instance of 151 trains and a time frame of two hours.

9.3. Concluding Analysis

The overview of state-of-the-art and state-of-the-practice indicates that there have been significant advances in the domain during the past ten years. The opportunities provided by innovations in computing, have resulted in that several approaches can tackle quite large problem instances, including a detailed representation of the infrastructure and train traffic dynamics,
within a few seconds. Since the prerequisites and operating rules are different in each country and network, the objectives and quality indicators adopted are naturally different, but the need to enable the use of multiple objectives has received more attention lately.

Results related to more practical aspects are unfortunately rarely reported, so far, limited by the modest number of field tests and practical implementations.

It is also challenging to compare and benchmark alternative approaches, partly due to that the contexts and case studies in focus shift, but also because the objectives adopted in the suggested approaches are different and no common evaluation framework for quality assessment of such systems and their rescheduling solutions exists yet.

There are, however, a few exceptions where approaches have been applied to (partly) the same rail corridor.

In (Törnquist Krasemann, Computational decision-support for railway traffic management and associated configuration challenges: an experimental study, 2015) a Mixed Integer Linear Programming (MILP) approach is applied on a set of 20 disturbance scenarios on the Swedish part of the Iron Ore Line, i.e. the stretch Björnfjell–Riksgränsen–Kiruna–Boden–Luleå. The approach uses a mesoscopic model of the network. Two main alternative objectives functions were adopted: 1) Minimizing the total final delay exceeding three minutes when trains arrive at their final destination (within the problem instance) and 2) minimizing the delays exceeding three minutes at intermediary, scheduled commercial stops as well as the final destinations. There are some obvious pros and cons with these two objective functions. The first one does not consider what happens to the trains “en-route” although it is often reflected partly at the final destination, but connections and so forth may then be overlooked. The second objective, which attempts to include the en-route punctuality by minimizing the delay at commercial stops, can be interpreted as if trains with several stops are given priority. Time penalties were associated with unscheduled stops by the freight trains, which were intended to represent the run time extensions associated with a heavy and long train needed to come to a full stop and then accelerate to full speed again. Furthermore, trains were also prohibited from making a stop at a station if no track with sufficient length was available. 20 different disturbances scenarios were simulated. The initial source of disturbance included both early and delayed departures of up to 45 minutes as well as temporary large running time extensions to simulate temporary signalling failures on the line. The scheduling time frame was four hours and the instances included 113 trains, of which 46 trains ran on a longer part of the line. Four different solution quality metrics were used including the two train delay measures, the number of delayed trains and the number of extra stops by freight trains. The MILP model was solved by commercial software, CPLEX 12.5, and it took 1.2 s – 45 seconds to reach optimal solutions. However, part of the computation time was used to prove optimality rather than to find the corresponding optimal solution.

In a more recent study by Bach et al. (2019), the Bender’s like decomposition approach presented in (Lamorgese, Mannino, & Piacentini, Optimal train dispatching by Benders’-like reformulation, 2016) is applied to the mentioned Iron Ore Line and specifically to the stretch from Narvik in Norway to Kiruna in Sweden. The model used in (Bach, Mannino, & Sartor, 2019) has the same
level of detail as the model in (Törnquist Krasemann, Computational decision-support for railway traffic management and associated configuration challenges: an experimental study, 2015). The rescheduling time frame is 2 h and the approach delivers feasible solutions within 2 seconds, which appears to be the goal. The quality of the delivered solutions, and how they are evaluated by the Norwegian dispatchers in the study, is not discussed in the paper.

Furthermore, in the ON-TIME project, the optimization modules ROMA (D’Ariano & Pranzo, 2008) and RECIFE (Pellegrini, Marlière, & Rodriguez, 2014) were applied on three different test cases (Quaglietta, et al., 2016); the East Coast Main Line in the UK, the Utrecht-Eindhoven-Tilburg-Nijmegen corridor in the Netherlands and the Northern part of the Iron Ore Line in Sweden and Norway. In the case of the study on the Iron Ore Line, the two optimization-based rescheduling modules were set to handle the traffic on the stretch Narvik–Kiruna–Svappavaara during a speed reduction (90 km/h -> 20 km/h) between Rensjön and Bergfors lasting seven hours, from midnight until 07:00 am. 15 trains in total, of which three passenger trains, were to be dispatched and the optimization time frame was set to 1 hour. No computation times for that study are reported in (Quaglietta, et al., 2016). Since ROMA minimizes the maximum consecutive delay while RECIFE minimizes the total delay, the modules suggested different rescheduling solutions, which were compared and discussed. Both performed best w.r.t. their selected objective. The ambition to integrate several modules and attempt such a comparative study of algorithms is rare and an achievement in itself. Experience from that should be used when future demonstrator platforms are to be defined.
10. Real-Time Network Management: a Swedish Case Study

In Chapter 9, the potential benefits of computational decision-support for real-time railway network management is discussed. Following that discussion, we present in this chapter the results from a Swedish case study where a multi-objective parallel algorithm has been applied and evaluated. This chapter begins with a brief introduction to the problem of rescheduling railway traffic during disturbances, from a computational perspective. A summary of related work on parallel algorithms is then presented, followed by a presentation of the applied algorithmic approach. Then a summary of the conducted experiments and associated results are presented, followed by a discussion of the conclusions and some pointers for future work.

10.1. Introduction

In fully deregulated networks such as the Swedish national railway network, the control of the infrastructure and traffic management lies on a neutral national transport authority (IM), while the trains and associated transport services are operated by several different private companies (RUs). The decision-making during disturbances and disruptions is then depending on two, or more, different organizations. The decision-making includes different types of rescheduling actions, which can be divided as follows:

(a) Re-timing of trains by allocating new arrival and departures times, including modification of speed profiles and halting schedules.
(b) Re-ordering of trains by adjusting the meet-pass plans.
(c) Local re-routing, by allocating alternative tracks on the line between two stations, or within the stations (i.e. platform re-assignment).
(d) Global re-routing by allocating alternative paths in the network.
(e) Management of train service connections and passenger transfers.
(f) Train service cancellations (partially or fully).

The tactics (a)-(c) can normally be employed by the IM without consulting the RUs, while the tactics (d)-(f) require consultation with the affected RUs. The rescheduling tactics are naturally associated with certain objectives and those tend to be different depending on which stakeholder perspective that is considered and how the traffic and transport system is organized. In practice, there may be regulations stating how the train dispatchers (i.e. the responsible part of the IM) shall, or should, prioritize between trains when delays occur. In Sweden, there is a general rule stating that trains on time have priority over trains that are delayed. However, experienced dispatchers know that this rule is not always practically relevant to apply and may instead take decisions that are better from a system perspective. Furthermore, RUs have internal priorities between their own trains and passenger transfers associated with trains of different RUs also need to be considered by the train dispatchers. So, in practice there is a palette of soft constraints and objectives which the train traffic controllers and dispatchers work with. Hence, also computational decision support should be able to take those aspects into account in one way or the other.
Most common is that final train delays are used as a primary quality indicator, although it becomes more frequent to include also more passenger train-specific and freight train-specific aspects either in the objective function, or as an evaluation metric used to assess and rank the proposed solutions.

As already mentioned, solving large railway traffic rescheduling problems in real-time is computationally demanding. The rescheduled timetables need to be of good quality, both from an operational as well as a passenger-oriented perspective and the solutions need to be provided sufficiently fast, i.e., within the allowed computational time limit.

Balancing this trade-off between speed and solution quality is a well-known challenge faced by current train rescheduling algorithms. Hence, there is a need to investigate faster solution approaches to train rescheduling that consider different perspectives.

10.2. Related Work

Multi-objective train (re-)scheduling has been a topic of considerable research interest for many years. Zhou and Zhong (2005) developed a branch and bound algorithm with effective dominance rules to generate pareto solutions to solve a bi-objective railway scheduling problem. Binder et al. (2017) solved a tri-objective railway rescheduling problem with special emphasis on minimizing passenger inconvenience. The problem is formulated as an integer linear program (ILP) that includes $\epsilon$-constraints for two of the three objectives.

More recently, Shakibayifar et al. (2018) proposed a multi-objective version of the variable neighbourhood search (VNS) method to solve the real-time traffic rescheduling problem and generate pareto frontiers (see Chapter 11.3). According to them, their approach supports the decision-maker to find a trade-off between both passengers’ and RUs’ viewpoints.

Research studies that employ concepts of parallel computing in railway research have been recently reviewed by Wu et al. (2018). Though the use of parallel computing for railway rescheduling has been investigated in recent years, e.g. by Bettinelli et al. (2017) and Josyula (2019), and has shown promising results, research on parallel computing for multi-objective train rescheduling is still rather scarce.

Recently, Josyula et al. (2018) reported significant speedup in train rescheduling as a result of parallel exploration of multiple branches of the search tree. Further research (Josyula, Törnquist Krasemann, & Lundberg, 2019) on parallel algorithms explores the potential of GPUs in train rescheduling. However, these parallel algorithms have been employed in the context of a single-objective train rescheduling problem, the objective being minimization of final delays of trains. Hence, it is relevant to also investigate the potential of parallel algorithms for multi-objective train rescheduling.
10.3. A Parallel Algorithm for Multi-Objective Train Rescheduling.

The sequential algorithm, and its parallel version outlined in Josyula et al. (2018), aims to re-schedule the railway traffic when (potential) conflicts are identified. The conflicts are identified and resolved iteratively by this greedy algorithm. A conflict is here defined as when a pair of trains require a particular resource (block section on the line, or in a station) during overlapping time periods, while also considering the required time separation. The conflicts are resolved by rescheduling the trains via re-timing, re-ordering and/or local re-routing decisions. The algorithm operates according to the Depth-First-Search principle in order to quickly get a first feasible solution which then also is used as a reference for enabling effective branching and cutting in the search tree.

To parallelize a sequential algorithm means in this case that multiple copies of the sequential algorithm is launched by parallel threads, traversing different parts of the search tree. The parallel algorithms continuously share information on their progress, in order to avoid exploring partial solutions that will lead to equally good, or worse solutions.

The algorithm is based on an effective representation of the search tree. Each node in the binary search tree corresponds to a partially infeasible timetable, where the next identified conflict between two particular trains on a particular track resource is identified. Each branch corresponds to the prioritization of one of the two trains, and the conflict is resolved by re-timing, re-ordering and/or local re-routing. Each partial solution, i.e. node, is evaluated by an optimistic assessment of the minimal cost for the resulting solutions (i.e. leaf nodes) as in a classical branch and cut. The evaluation metrics used to compute the cost are:

- **TFD** (Total Final Delay), where TFD refers to the sum of all trains’ delay at their final destination. ‘Final destination’ refers here to the last station within the instance.
- **TAD_{2}** (Total Accumulated Delay > 2), refers to the sum of trains’ delays exceeding two minutes at commercial stops.
- **TPD_{2}** (Total Passenger Delay > 2), refers to the sum of all passengers’ delays exceeding two minutes at their final stop.
- **#D_{2}pax** refers to the number of passengers that arrive at their final stop with a delay larger than two minutes.
- **#D_{2}trains** refers to the number of trains that arrive at their final destination with a delay.
- **#D_{2}sectr** refers to the number of secondary delayed trains, i.e. trains that are affected by knock-on delays, and it excludes here the trains that are considered as also suffering from a primary delay.
- **#tracks** refers to the number of re-arranged track and platform allocations, compared to the initial timetable.

There are two main ways of incorporating a multi-objective approach when iteratively searching for improved solutions. One way is to weigh all the different objectives in relation to each other and then you only need to use the accumulated objective value when comparing the alternative solutions, which simplifies the solution process significantly. It is, however, challenging to find
appropriate weights when you have multiple components, and their relation may not be linear.

The other way is to assess the solutions in each separate dimension, but the search tree then becomes significantly larger and computationally demanding to traverse. The result is a pool of several solutions with different properties, referred to as the Pareto frontier, see Figure 22 for an illustration. For a more detailed over of the algorithmic approach, see Chapter 9 in (Josyula, 2019).

![Figure 22: Illustration of a Pareto frontier of feasible solutions after rescheduling with a minimization of TFD and TPD (Josyula, 2019).](image)

**10.4. Experimental Study and Some Observations**

The aim of the study was to investigate which properties and factors that may be relevant to consider when evaluating the final solutions as well as when pruning branches in the binary search tree. We applied the first six metrics for pruning while the last metric, #tracks, were used only as a stability measure. That is, to assess if the rescheduling solutions required significant changes, which hence could affect the robustness of the solution.

Since we wanted to analyse how much the different metrics affect the solution procedure related to (a) the quality, as well as (b) the computational effort required when the search tree is expanded, we had six different pruning configurations. In P₁ only TFD was used for pruning and in P₂ both TFD and TAD₂ were used for pruning etc., see Table 8 below.

We applied the parallel algorithmic approach using the six different pruning configurations on 40 disturbance scenarios occurring on the Blekinge railway line, including all stations from Karlskrona to Tjörnarps, see Figure 21.
Table 8: Overview of the different configuration used for the pruning the search tree when the algorithm performs the “branch and cut”.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Pruning metrics used in the criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{P}_1$</td>
<td>TFD</td>
</tr>
<tr>
<td>$\mathcal{P}_2$</td>
<td>TFD, TAD$_2$</td>
</tr>
<tr>
<td>$\mathcal{P}_3$</td>
<td>TFD, TAD$_2$, TPD$_2$</td>
</tr>
<tr>
<td>$\mathcal{P}_4$</td>
<td>TFD, TAD$_2$, TPD$_2$, #D$_2$pax</td>
</tr>
<tr>
<td>$\mathcal{P}_5$</td>
<td>TFD, TAD$_2$, TPD$_2$, #D$_2$pax, #Dtrains</td>
</tr>
<tr>
<td>$\mathcal{P}_6$</td>
<td>TFD, TAD$_2$, TPD$_2$, #D$_2$pax, #Dtrains, #D$_2$sectr</td>
</tr>
</tbody>
</table>

Figure 23: Overview of the studied railway stretch between Karlskrona-Kristianstad-Hässleholm (Trafikverket, 2018b).

The infrastructure consists of a single-track line with 59 sections (including stations), and all tracks are bi-directional. The original timetable is from 15:50 to 21:10 (5 hours 20 minutes). The disturbance scenarios 1–8, 9–16, 17–24, 25–32, and 33–40 correspond to induced primary delays of 5 minutes, 13 minutes, 17 minutes, 21 minutes and 25 minutes respectively. The time windows for the scenarios vary between 2 hours and 4.9 hours. In every considered disturbance scenario
of the case study, we have a single initial source of disturbance that then is spread to other trains in most scenarios, depending on how the algorithm resolves the identified conflicts while avoiding deadlocks.

In Table 9, we can observe how the pool of solutions increase with an increasing number of pruning metrics from one to four in scenario 3. The last two criteria do not add any value in this case. In scenario 3, train 1250 becomes minimum five minutes late between Vinslöv and Önnestad and the traffic between Karlskrona and Hässleholm within a time period of 3.5h is then re-scheduled. See Figure 24 on how the trains are prioritized in the solution found when \( P_1 \) was applied.

**Table 9: Overview of the resulting pareto frontiers in scenario 3.**

<table>
<thead>
<tr>
<th>( P_i )</th>
<th>TFD</th>
<th>TAD2</th>
<th>TPĐ2</th>
<th>#D2pax</th>
<th>#Dtrains</th>
<th>#D2sectr</th>
<th>#tracks</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_1 )</td>
<td>12.7 min</td>
<td>24.3 min</td>
<td>9.8 hr</td>
<td>135</td>
<td>4</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>( P_2 )</td>
<td>15.7 min</td>
<td>33.3 min</td>
<td>9.1 hr</td>
<td>152</td>
<td>4</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>( P_4 )</td>
<td>20 min</td>
<td>34.5 min</td>
<td>9.4 hr</td>
<td>135</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>( P_4 )</td>
<td>22 min</td>
<td>32.5 min</td>
<td>12.9 hr</td>
<td>117</td>
<td>3</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>( P_4 )</td>
<td>26 min</td>
<td>62.5 min</td>
<td>15.1 hr</td>
<td>107</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>( P_4 )</td>
<td>21.3 min</td>
<td>51 min</td>
<td>13.5 hr</td>
<td>129</td>
<td>3</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

**Figure 24: A snapshot of a part of the re-scheduled solution in scenario 3 when pruning configuration \( P_1 \) was applied (i.e. the objective TFD).**

In brief, the study shows that it is important to consider more than one objective in this context, in particular when the objectives may be in conflict. A detailed presentation of the experiments and the result can be found in Chapter 9 of (Josyula, 2019).
10.5. Conclusions and Future Work

The results from this study indicate that it is beneficial to explore the search space in more than one dimension, but which set-up of objectives that is most beneficial to use needs to be investigated further. For obvious reasons, the computational time increases in some cases as a result of a growing search space, but an efficient parallel algorithm should be able to tackle this.

In the experiments, we have used artificial passenger flow data to compute the passenger-oriented metrics. Since almost all trains are passenger trains in the used data set, and we wanted to achieve a similar prioritization of the trains in this study, we treated all trains as passenger trains. Other passenger flow data, and using relevant weights for the freight trains, would probably have generated other solutions and the impact of TPD$_2$ and #D$_2$pax on the pruning could be different, which needs to be investigated further in our future studies. In those studies it would be relevant to also use data sets where more freight traffic is present, and also to incorporate freight train oriented metrics, inspired by the work presented in (Törnquist Krasemann, Computational decision-support for railway traffic management and associated configuration challenges: an experimental study, 2015) .
11. Research Gap and Challenges

Research gives new knowledge about the railway system. An overall objective in the Fr8Rail II project is to improve capacity and punctuality in the railway system by improved concepts and methods for tactical planning and operational traffic. In Fr8Rail II, a demonstrator on improved short-term planning and daily planning with improved interaction IM – RU, including network and yards/terminals, will be developed. The project will develop methods and tools that can reduce inefficiencies in real time network management by, e.g., improving the coordination between yards/terminals and the line network, and between IM and RUs. In addition, requirements for a real-time network management demonstrator will be specified.

Research and demonstration are good methods to solve complex problems and to initiate cooperation and discussion between involved actors and problem owners.

The main challenges we address in the Fr8Rail II project are:

- To reduce and connect the gap between planning and operational traffic
- Improved decision support planning, re-planning and punctuality prediction by data analytics, AI, simulation and optimisation
- Improved short term planning for freight trains and maintenance
- More precise and competitive freight trains - improved planning and control

To get more precise and competitive freight trains we need improved interaction between yards and network. For this we need to develop decision support for IMs, YMs and (freight) RUs.

In this report, we have pointed out open research tasks related to several aspects of the gap between daily timetable and operational traffic: analysing timetable reliability; improving the process for reserving (and releasing) maintenance track possessions; and analysing correlations between train departure times and operational conditions at the yard.

We focused on the Swedish system and showed that a majority of the freight trains depart before or after their scheduled departure time or are cancelled. For the trains that run, punctuality is low. When observing the current process of allocating a train path for an early or late departure for a freight train, we can clearly identify the need to automatize (and optimize) the process. A previously proposed method for inserting such a train relies on that all other trains are fixed. In case this gives a conflict-free (feasible) solution, and when indeed only one train needs to be inserted, the method gives a good decision support. However, a research gap remains: how can we have a more flexible rescheduling, which allows for train paths to be inserted even in time periods with high capacity usage, and how can we add train paths that allow flexibility for future insertions? This is the focus of the short-term rescheduling in this project.

Another key point that was pointed out is the importance of the reliability of a timetable, especially when punctuality is known to be low. Given simulation methods are often time-consuming, making them difficult to apply for processes that require fast decisions, e.g. rescheduling in case
of disturbances or short-term planning close to operation. Furthermore, the simulation of larger networks is often difficult to perform with common simulation tools. In addition, the way to evaluate the performance is an important research question, as well as the question regarding how detailed a model must be in order to provide the right results.

Moreover, capacity allocation for maintenance work leads to restrictions for the operating train traffic. Hence, research into several aspects of how maintenance windows should be booked, and when previously booked track capacity should be made available again, are important questions.

The identified research gaps in the area of capacity planning for engineering works and maintenance can be summarized as follows:

- Quantitative studies regarding the suitability of different traffic adaptation approaches due to engineering work closures are interesting both for planning and real time traffic control.
- The value of track capacity made available at different time horizons might be possible to establish via a stated preference study with affected RUs and maintenance contractors.
- Incentives and fees regarding maintenance windows and possession bookings need to be further studied.
- Cause–effect relationships regarding how planning and execution of infrastructure maintenance affects different stakeholders and aspects like operation costs, track quality and delay statistics are largely missing. In such studies, the willingness by different companies to release data needs to be considered.

Evaluation and benchmark of alternative approaches for computational decision-support for real-time network management are important but currently very challenging. One reason is that the contexts and case studies in focus shift, another that the objectives adopted in the suggested approaches are different and, hence, that no common evaluation framework for quality assessment of such systems and their rescheduling solutions yet exists.

Furthermore, results related to more practical aspects are unfortunately rarely reported, and, so far, limited by the modest number of field tests and practical implementations. What can be observed is that the research stream on developing computational decision-support systems and solution algorithms for railway traffic management has not yet been sufficiently merged with the corresponding research stream focusing on aspects of human-computer interaction. Several research studies analyse potential objectives that could dictate, or guide, the train rescheduling, and propose quality metrics that may be relevant to assess and used to present the proposed solutions to the dispatchers. Meanwhile there are some projects which analyse the complexity of the mentioned decision-making process and what automation can/cannot do, but no published study which aims to compare and discuss those objectives, quality metrics or visualization aspects with potential decision-support users, could be found so far.

Finally, while most of the research gaps discussed above mainly refer to line management, the
yards and terminals are important entities of the freight railway network. The operations at the yards lead to large deviations from the planned timetable, where a majority of freight trains depart outside the timetable slots. An interview study indicated several causes to variance in the handling time at the yard, which can result in departure time deviations. The correlation between the status and operations of the yard, and the actual departure time of freight trains should be further investigated, and we propose a quantitative study to determine the correlation between train departure deviation and operational conditions at the yard. Especially the following questions are of interest:

- How can freight trains be scheduled better, and to what extent should they be scheduled in detail at all, knowing that in practice they deviate widely from the timetable?
- What are the impacts on other trains from freight trains deviating so much from their train paths?
- How can yard management and dispatching be made to respect the timetable to a greater extent?

We believe that optimisation and simulation models for railway timetabling are key components to answer these questions.
12. Conclusions

Fr8Rail II addresses the problem of how to improve capacity and punctuality in the railway system by developing concepts and methods for tactical planning and operational traffic management. An objective is to develop methods and tools that can support and improve real time network management by, e.g., improving the coordination between yards/terminals and the line network, and between IM, YMs and freight RUs. In this report the state-of-art has been summarised.

The scope is most relevant for Sweden, where the freight trains currently and historically adhere very poorly to the timetable. For the line Malmö – Hallsberg years 2011 – 2017, the narrow channel precision (±3 min) is between 11% and 15% and the broad channel precision (±15 min) is between 41% and 50%. Further, 65% of freight train departures from marshalling yards happens ahead of schedule. The probability of deviation varies substantially in both time and space, and it increases with the distance from the scheduled train path. Better tools for network planning and management on tactical and operational level can help manage the deviations and bring the planning and operational processes closer together.

Aiming for improvements of the operational traffic, there is a need for methodological development of methods for several planning horizons. The conventional microscopic simulation approaches outlined in this report are applicable for use during long-term planning. For application in real time management, other, faster methods are needed.

We distinguish between timetable analysis and generation methods that are based on simulation, and those based on optimisation. Within the scope of Fr8Rail II, a demonstrator for timetable improvements with plug-in modules will be developed. Based on the indicated needs, a simulation tool that can handle larger networks and is connectable to other modules is suitable. For short-term scheduling and rescheduling, we believe an optimisation model based on meta-heuristics is a good strategy. The two modules will be connected, and there will also be possibilities to combine the simulation and the optimisation techniques.

The maintenance planning process and improvement potential have been described. This is a new piece of the puzzle and is important to close the gap between timetable planning and operational traffic. For maintenance planning we have following conclusions:

- The different planning processes at the IM, the RUs and the MCs should be aligned.
- A common monthly planning cycle with deadlines measured towards the start of each planning period seems natural to adopt.
- To release maintenance windows earlier seems to be appropriate, especially if monthly planning is adopted.
- The amount of planning rework and adaptions is an important metric to consider when redesigning the various planning processes and when comparing the effect of different time settings and planning deadlines.
When developing new approaches for computational decision-support tools for real-time network management, it is important, but very challenging to evaluate and benchmark with existing software tools. One reason is that the contexts and case studies in focus shift, another that the objectives adopted in the suggested approaches are different and no common evaluation framework for quality assessment of such systems and their rescheduling solutions exists yet. Furthermore, results related to more practical aspects are unfortunately rarely reported, so far, limited by the modest number of field tests and practical implementations.

Finally, we also observe, that the research stream on computational decision-support and algorithm development for railway traffic management, has not yet been sufficiently merged with the corresponding research stream focusing on aspects of human computer interaction. Several research studies analyse potential objectives that could dictate, or guide, the train rescheduling, and propose quality metrics that may be relevant to assess and present the proposed solutions to the dispatchers. Meanwhile there are some projects which analyse the complexity of the mentioned decision-making process and what automation can/cannot do. However, no published study which aim to compare and discuss those objectives, quality metrics and visualization aspects with potential decision-support users, could be found so far.
13. References


ARCC. (2017). *D2.1 - Description of automation/optimisation requirements and capabilities of decision making process in marshalling yards and terminals*. DB Cargo AG.

ARCC. (2018a). *D2.2 - Description of business processes of a network management system and the interactions/interfaces with a real-time yard management system*. Trafikverket.


ARCC. (2019). Retrieved from Automated Rail Cargo Consortium: Rail freight automation research activities to boost levels of quality, efficiency and cost effectiveness in all areas of rail freight operations: https://projects.shift2rail.org/s2r_ip5_n.aspx?p=ARCC


Edholm, G. (2019a). Personal communications Trafikverket 07/05/209. (T. Lidén, Interviewer)


Fr8Hub. (2018). D3.1 State-of-the-art and specification of innovations, demonstrations and simulations. Real time information applications and energy efficient solutions for rail


Haster, E. (2016). *Mikrosimuleringar av servicefönster på Södra stambanan – Kapacitetsanalys*


relaxation. *Transportation Research Part B (submitted).*


