Deliverable D 8.1
The need for future development of methods and models for capacity simulations and feedback loops between planning and operations

Project acronym:	MOTIONAL
Starting date:	01/01/2023
Duration (in months):	46
Call (part) identifier:	HORIZON-ER-JU-2022-01
Grant agreement no:	101101973
Due date of deliverable:	Month 13
Actual submission date:	22/12/2023
Code	FP1-WP08-D-TRV-001-03
Responsible/Author:	Per Köhler (TRV)
Dissemination level:	PU
Status:	External review version

Reviewed: (yes)

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101101973.

Document history		
Revision	*Date*	*Description*
1	2023-02-23	First draft
2	2023-09-01	Change to new template
3	2023-10-15	Internal WP8 review version
4	2023-11-17	Internal review version
5	2023-12-18	External review version
6	2024-06-14	Modified version after JU feedback

Report contributors		
Name	Beneficiary Short Name	Details of contribution
Emma Dyrssen	TRV	Subtask leader, coordinating writing Section 6
Kristian Persson	TRV	Subtask leader, coordinating writing Section 6
Gerben Scheepmaker	NS	Subtask leader, coordinating writing Section 7
Emma Solinen	TRV	Subtask leader, coordinating writing Section 8
Alwin Pot	ProRail	Subtask leader, coordinating writing Section 9
Blanca Delgado	INECO	WP8 reviewer
Isabel Meseguer Hijos	CAF	WP8 reviewer
Henri Olink	ProRail	WP8 reviewer
All beneficiaries in WP8	Several	Contributed with texts in all sections
Carmen Ramos Prieto	Indra	Internal reviewer
Giovanni Luca Giacco	FS	Internal reviewer

Table of Contents

List of Figures

List of Tables

1. Executive Summary

Deliverable D8.1 mainly focuses on methods and models for capacity simulation including feedback loops between planning and operation. The aim is to improve current practice and to extend the scope of capacity studies with the application of models which enable capacity-impact studies of, in ER FP2 R2DATO, new developed technologies specifically: ETCS HTD (previous HL3), ATO/C-DAS linkage to ETCS and next generation brakes. Also, the aim is to identify best practices and needs for further developments of these methods and the modelling configuration of the abovementioned innovations. This document sets the baseline for the development of methods and models to test these in a simulation environment. After being tested on feasibility they will be prepared for capacity studies to identify the future potential of the above-mentioned innovations, which will be executed by WP 9. Results of WP9 will be disseminated to R2DATO, where they are part of the new technology's impact assessment.

A conclusion from the mapping of existing tools among partners is that there are several capacity simulation tools available with developed functionality for simulation of a transport plan and also for simulation of ETCS L2. However, developments are needed to simulate the capacity effect of new digital technologies like ETCS HTD, next generation brakes, C-DAS/ATO, driver behaviour and TMS functionality, including the train path envelope and the concept of TMS steering of ATO. This also applies for improved feedback loops including crew scheduling and large networks.

A general methodology is derived for the verification, calibration and validation of railway simulation models using literature review and practical examples. However, an integral description with application of verification, calibration and validation processes for railway simulation tools is missing.

In the deliverable, an overview of feedback loops between operations and planning is given and development needs are defined. It can be concluded that feedback loops between operations and planning are essential to improve railway planning and that timetable analysis and simulation can give useful outputs, as a complement to operational performance. In order to achieve more solid and reliable models for planning and simulation; data improvement by continuous feedback of historical information available for analysis is needed. There are also other areas where methodological development is needed, typically related to specific cases where there are missing functionalities today, i.e. simulation with TMS.

The next part of the report describes per partner the status of current research on capacity effects of system developments such as ETCS level 2 and hybrid train detection, ATO and TMS. Since capacity becomes scarce, solutions are being sought in these new technologies, also European CCS will gradually transfer to ETCS. However, a lot is still unknown due to a lack of operational situations, so simulation data becomes more valuable. The biggest uncertainty is train driver's behaviour on ETCS equipped lines, with L2 or the newly developed HTD[1].

The results from this report will be used for capacity studies with improved simulation methods, also capable of designing the capacity impact of new technologies developed in FP1 and FP2.

[1] Please note that the ETCS scope is limited to operations deployable functionalities within the scope of Europe's Rail. As a consequence of this, ETCS L3 capacity studies are excluded.

2. Abbreviations and acronyms

Abbreviation / Acronym	Description
ANS	Automatic Numbering System
ARS	Automatic Route Setting
ATC	Automatic train operation
ATO	Automatic Train Operation
ATP	Automatic Train Protection
CBTC	Communications-based train control
CCS	Control, Command and Signalling
C-DAS	Connected Driver Advisory System
DMI	Driver Machine Interface
EBD	Emergency Brake Deceleration
EBI	Emergency Brake Intervention supervision limit
EMU	Electric Multiple Unit
EOA	End Of Authority
EPA	ERTMS Protocol Analyzer. Tool for reading ETCS protocol data (RBC-OBU)
ERA	European Union Agency for Railways
ERTMS	European Rail Traffic Management System
ETCS	European Train Control System
GoA	Grade of Automation
GPS	Global Positioning System
GSM-R	Global System for Mobile communications for Railways
GUI	Guidance curve
HABD	Hot Axle Box Detector
HESE	Headway and signal locations in ETCS L2 (simulation tool)
HL3	Hybrid Level 3
I	Indication speed supervision limit
ICNG	Intercity Nieuwe Generatie (Dutch trainset “Intercity New Generation”)
IM	Infrastructure Manager
IXL	Interlocking
KPI	Key Performance Index
L2	Level 2
MA	Movement Authority
MRSP	Most Restrictive Speed Profile
Nexteo	Nouveau Système d’EXploitation des Trains
OB	On board
OBU	On Board Unit
OSRD	Open-Source Railway Designer
OT	OpenTrack
P	Permitted speed supervision limit

PAB	Performance Analysis Office (Dutch: Prestatieanalsyebureau)
P20, P50, P80	20th, 50th and 80th Percentile
RBC	Radio Block Centre
RTC	Rail traffic controller
RTO	Remote train operation
RU	Railway Undertaking
SAAL-corridor	Dutch corridor between Schiphol Airport – Amsterdam – Almere – Lelystad
SBD	Service Brake Deceleration
SBI	Service Brake Intervention supervision limit
SIPH	Système Industriel de Production Horaire
SMS	Stochastic Microscopic Simulation
SNG	Sprinter Nieuwe Generatie (Dutch regional trainset “Sprinter New Generation”)
SSEM	Supervised Speed Envelope Management
SFERA	Smart Communications for Efficient Rail Activities
STIPT	Methodology to automatically derive the cause of the delay in realization data using multiple data sources
STM	Specific Transmission Modules
SvL	Supervised Location
TMS	Traffic Management System
TOC	Train Operating Company
TPS	Train Planning System
TRP	Train Regulation Point
TS	Trackside
TSI	Technical Specifications for Interoperability
TTCMS	Timetabling, Traffic Control and Management System
TTSM	Time Table Speed Management
VSTP	Very-Short Term Planning
W	Warning supervision limit
WP	Work Package

Explanation of some capacity related terms

Capacity utilisation	Percentage of time that is occupied by trains on a certain stretch. The occupation is built by headway, route setting, train driving, route releasing, buffer times and capacity loss due to speed differences and merging and diverting traffic.
Headway	Time between the front of two trains on a certain location. Ideally the second train can run without hindrance which gives the technical minimum headway. On top of the technical minimum headway there are buffer times added to create a stable timetable.
Punctuality	Percentage of trains arriving within a given delay threshold (usually 1, 3, 5 or 6 minutes) from their original planned time. Times can be measured only at the last station or also at a (limited) set of intermediate stations.
Robustness	The timetable's possibility to recover from delays and prevent delays from spreading.
Running time	Time to run between two timing points measured at the train head. Usually this is measured in the unhindered situation.
Macroscopic infrastructure model	"Node-link-models that contain aggregate information on nodes and links." (Hansen & Pachl, 2014, p. 315).
Mesoscopic infrastructure model	"Node-link-models as syntheses of microscopic and macroscopic infrastructure models. Signal blocks and headways are modelled only in stations and interlocking areas, not on links." (Hansen & Pachl, 2014, p. 315)
Microscopic infrastructure model	"Node link models that contain, depending on their purpose, the highest possible level of detail on nodes and links" (Hansen & Pachl, 2014, p. 316)
Static simulation	Timetable analysis without simulation by analysing the running times and conflicts.
Deterministic simulation	Simulation without any randomness, the model gives the same result every simulation run.
Stochastic simulation	Simulation with randomness added by including probability distributions for different processes, the model gives different result every time and has to be run several times to give reliable result.
Feasible timetable	Timetable where all trains have feasible running times (not faster than they can run in reality) and where no occupation conflicts between trains exist.

3. Background

The present document constitutes the Deliverable D8.1 report *"The need for future development of methods and models for capacity simulations and feedback loops between planning and operations"* in the framework of the Flagship Project FP1 – [MOTIONAL] as described in the EU - RAIL multi-annual working program and contributes as well to the Flagship Project FP2 – [R2DATO].

This is the first delivery of WP8 *"Development of simulation and operational feedback for improved planning"* based on input of task 8.2 *"Identify the need for future development of methods and models for capacity simulations and feedback loops between planning and operations"*. The leader of the task is TRV and the other beneficiaries are NSR, PR, SNCF, ADIF, INDRA, CAF and CEIT. The task has been performed during month 2-13 of the project (January – December 2023). December 2023 was also set as deadline also set for deliverable D8.1.

Rail traffic simulation is a powerful method for analysing and optimizing rail transportation systems. It involves methods and models through the use of software programs that simulate the simultaneous movement of trains and rolling stock through a network, considering various factors such as train characteristics, train schedules, track capacities, signalling systems, traffic control and other variables. These tools can be used to visualize and analyse certain processes, train and dispatcher behaviour, subsystems or a complete use case.
Simulation models are essential to be able to estimate the capacity of a given infrastructure and the feasibility and robustness of a timetable. By using simulation models to test different scenarios and strategies, infrastructure managers, railway operators and planners can improve the reliability and efficiency of daily or future operations.

The European railway network has many lines with high-capacity utilisation and bottlenecks. New infrastructure investments are in general costly and take normally very long time to complete. New digital systems such as C-DAS, ATO, next generation brakes and ETCS Hybrid Level 3 may be a way to improve the capacity, but a lot is still unknown about the capacity effects due to a lack of real-world implementations. Therefore, capacity simulation will be a relevant tool to evaluate and indicate the capacity effect of different developments planned within Europe's Rail.

4. Objective & Scope

The overall objective of WP8 is to develop and improve railway traffic simulation methods, models and knowledge. This will enable more reliable and effective capacity and punctuality evaluations as well as predictions of the railway network. The developed methods and models will be used to improve feedback loops between operations and planning and to increase knowledge about the potential capacity impact of new technologies such as C-DAS, ATO, next generation brakes and ETCS HTD.

This document is a base for the future development work within this work package. By describing the current state of the art regarding capacity simulation it will be easier to identify necessary development needs within the field of capacity simulation.

Scope

This WP focuses on improving methods and models used by partners within the WP8/WP9 project group. In specific cases existing software adjustment is needed to facilitate the research objectives. Generic software development or delivery is not in scope. Regarding research on new innovative technologies the project's scope is based on requirements coming out from FP2 on technologies expected to be ready for deployment around 2030. Given this ETCS HTD is considered being more technologic mature for operations than ETCS L3 for deployment in this timeframe.

5. Outline

The structure of this document is as follows. Section 6 discusses research on relevant simulators and also discusses how they connect to use cases. Afterwards, the processes of verification, calibration and validation are discussed in Section 7. Feedback-loops between planning and operations are discussed in Section 8. Next, research on the capacity effects of TMS combined with C-DAS or ATO is discussed in Section 9. Section 10 explains the interaction with other Flagship Projects. Finally, the conclusions of this deliverable are discussed in Section 11.

Terms being used in this document are being explained in Section 2.

6. Research on relevant simulators and how they connect to use cases

In this section, research on the application of capacity simulations and methodologies of simulations is conducted, linking those to specific purposes and use cases. Current development among partners in WP8/WP9 is described and different tools are evaluated. Finally, the need of future development is defined. The main focus is on micro and macro simulations.

6.1 Methodologies for capacity simulation

Capacity simulators can either be microscopic or macroscopic, or a mix in-between (called mesoscopic). Microscopic simulation is highly detailed with positions of each signal, track and switch, as well as detailed modelling of signalling and the train protection system. Its drawback is that it requires extensive data pre-processing prior to simulation. The simulation computation times, considering stochastic simulations, are often long and, especially with single track lines and/or complex stations, there are often problems with deadlocks complicating the analyses. It may even be impossible to use microsimulation in analyses if the number of deadlocks is too high. The long simulation times can be compensated by providing more computing power. However, this does not solve the deadlock issues.

Macroscopic simulation is normally significantly faster and requires less preparatory work prior to simulation. The drawback is that the higher level of aggregation in terms of both the infrastructure description and modelling of train movements means that macroscopic simulation may not be appropriate in certain contexts, for instance when checking the robustness of timetables. Whereas simulating a network with many trains can take several minutes per cycle in a microscopic tool, corresponding macroscopic simulation takes a few seconds per cycle. The fast computation speed provides a practical opportunity to simulate significantly more scenarios in applications such as timetable construction, track work planning or infrastructure investment planning. One way to ensure feasible timetables after macroscopic simulations is to make microscopic checks when the best macroscopic timetable suggestion is found, for example potential track occupation conflicts at stations, headway conflicts or conflicts at switches. The research question (i.e. the required level of detail in the analysis to answer the research question) should determine if macroscopic, mesoscopic or microscopic simulation is to be used, or a combination of them. Regardless of whether microscopic or macroscopic simulation is used, the timetable needs to be feasible enough, i.e., the number and magnitude of built-in conflicts should not be too high.

Capacity simulations can either be static, deterministic or stochastic. Each simulator tool could provide support for one or several of these. A static simulation is not a real simulation, but rather timetabling, since only direct interaction between different capacities are studied. For example, static models can identify capacity allocation conflicts between two trains (i.e., two trains are using the same track at the same time), but it does not consider how that conflict is solved, which train that will run first or if the delay caused will lead to new conflicts with other capacities. A static model is therefore more of a conflict detection system and a good way to investigate feasibility of a timetable. Other simulations (deterministic and stochastic) also consider secondary delays and conflicts. Simulation without any stochastic perturbations is called a deterministic simulation. If there are conflicts to begin with, this can give an idea of the magnitude of built-in delays in the timetable. Introducing stochastic (random) delays to model initial/primary delays, requires that a simulation is run for multiple cycles to provide a stable output for the measure(s) of outcome that is analysed. The required number of cycles depends for instance on the model network, properties of the different stochastic distributions, aggregation in output processing and measure(s) of outcome.

When choosing a simulation tool for investigating a certain topic it is important to first define a clear research question. Thereafter, a model type that suits the research question should be chosen. For example, if the research question requires large networks and long time periods, a macroscopic model might be preferred over a microscopic model. The same choice should be done for the type of simulation. Then a tool should be chosen with the preferred model and simulation characteristics. **Fehler! Verweisquelle konnte nicht gefunden werden.** shows an overview of this p rocess.

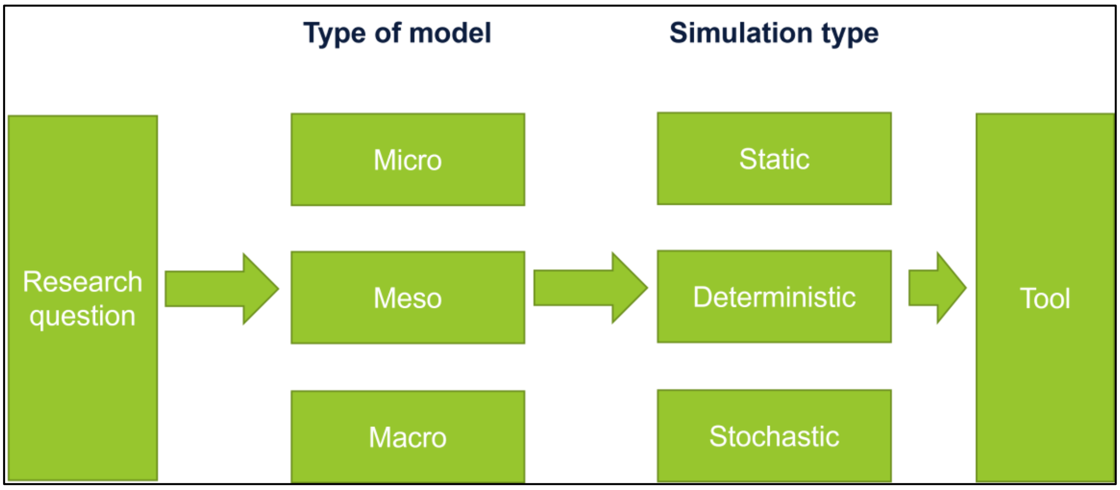


Figure 1. Criteria for choosing a suitable simulation tool. The basis for the selection is always the research question.

6.2 Mapping of current simulation usage among partners

In Europe, many different capacity simulation tools are used. Examples of microscopic railway simulation tools are Railsys (Radtke & Hauptmann, 2004), OpenTrack (Nash & Huerlimann, 2004), Trenissimo (De Fabris et al., 2018) and LUKS (Janecek & Weymann, 2010). RTC (Rail Traffic Controller) is a railway simulation tool used mainly in North America. PLASA (2017) covers microscopic and macroscopic techniques for railway simulation, also mentioning advantages and disadvantages of these approaches and some of the mentioned simulation tools. Some are used for ordinary timetable design and assessment while others are only used in research.

For the purposes of WP8/WP9, with capacity simulations concerning ETCS, ATO and C-DAS and feedback loops between planning and operation, 15 different capacity simulation tools have been identified that are used in a European context by the different partners involved in this work package. The different tools, their properties and which partner in WP8/WP9 that uses the tool are presented in **Fehler! Verweisquelle konnte nicht gefunden werden.**. A "X" in the table means that the simulator has the required capability, a D" indicates that development is needed before functionality exists and a "(X)" indicates that only limited development is needed. If a cell is blank, the simulation tool has not the described functionality at all or the tool is not suitable for that type of functionality.

Table 1. Capacity simulation tools and properties, per involved partner in Motional WP8/WP9

Model name	**Partner**	**Tool type**	**Static**	**Deterministic**	**Stochastic**	**To be used in WP8/WP9**
AnyCrew	NS	Macro		X	X	Yes
CAF Tool	CAF	Micro		X		Yes
EPA	TRV-VTI	Micro		X		Yes
ERTMS Traffic Simulation Laboratory	CEDEX (Adif)	Micro		X		No
FRISO	NS, ProRail	Micro	X	X	X	No
HESE	TRV-KTH	Micro	X	X	D	Yes
INTROOS	NS	Macro		X	X	No
OpenTrack	ProRail	Micro	X	X	X	No
OPTICON	CEIT	Macro	X	X		Yes
OSRD (Open Source Railway Designer)	SNCF	Macro/Micro	D	D	D	No
Planif	Adif	Macro/Meso	X			No
PROTON	TRV-KTH	Macro		(X)	X	Yes
RailSys	NS, ProRail, SNCF, TRV	Micro	X	X	X	Yes
TMS_CAP	Indra	Macro/Micro		X		Yes
VTI Train-Driver Simulator	TRV-VTI	Micro			X	Yes

The following sections contain practical examples of usage of the different capacity simulation tools and the future development needs identified by the partners using them.

6.3 Simulation usage at TRV

At TRV, RailSys (Radtke & Hauptmann, 2004) is the main tool for capacity analysis and evaluation, mainly used for analysis of future timetable and infrastructure scenarios. The extent to which different functions in the tool are used varies depending on the scope and purpose of different projects. Some studies can mainly be seen as feasibility studies involving one or several future timetable scenarios and/or infrastructure variants. A typical question is whether a proposed traffic can be accommodated, with which headways (or headway margins) and if additional infrastructure resources are needed in relation to the current situation or those already proposed in the study. In some cases, this type of study is supplemented with stochastic simulation in order primarily to be able to compare expected outcomes between different alternatives. Another relatively typical question is to examine how a planned track work that involves restrictions in the available infrastructure affects both the realizable timetable and punctuality. Several consultant companies in Sweden use RailSys as well, both in connection to projects commissioned by TRV and in projects for other clients.

TRV maintains a national Swedish network infrastructure model and provides it on request for Swedish RailSys users. They provide two main models: the current infrastructure and a future network model for 2040. By having both models, it is possible to measure capacity effects of upcoming infrastructure investments, local effects as well as regional or national. The upcoming year's annual timetable is imported into RailSys annually, usually a few months before it is taken into use. ETCS L2 and HL3 have been implemented in RailSys and capacity has been evaluated for a couple of railway lines.

The macro tool Proton is also available at TRV, but right now the forms and methods of use of the tool are being developed in a research environment to get better knowledge about how it could and should be used at TRV (see section 6.4). The objective is to gain experience in the use of macroscopic simulation in a Swedish context and with the intention of using simulation to a greater extent in the future for various analyses.

6.3.1 Development needs at TRV

At TRV, there is a need to develop simulation tools to be able to analyse the capacity effects of new digital developments such as different ETCS improvements, C-DAS and ATO.

6.4 Simulation usage at TRV A.E. Lund/KTH

Lund/KTH uses Railsys, PROTON and HESE as simulation tools.

Within the Shift2Rail project DB developed a macroscopic simulation tool, formerly named PRISM and now named PROTON (Zinser et al., 2019). One of the main goals was that the tool should be able to simulate large networks in short computation time. KTH has, together with LU and TRV, used PROTON in various projects, for example within Shift2Rail projects Plasa2 and FR8Rail II and III, see PLASA-2 (2021) and FR8RAIL II (2021).

PROTON started out as a macroscopic tool focused on the simulation of large railway networks. Development is done by DB Analytics and efforts have been made to be able to import infrastructure, timetable, and other data from DB internal systems. The Swedish applications where PROTON has been used have, so far, been research oriented. Specific conversion scripts have been developed for importing and generating input data from Swedish sources for PROTON. TRV aims to be able to use the tool in more applications in the future and to have the opportunity to choose between micro and macro simulation depending on the application. DB has further developed PROTON into a microscopic simulation tool, and it is still under development, however, no microscopic application with PROTON has yet been performed or tested in Sweden.

TRV has used and is using the microscopic simulation tool RailSys (Radtke & Hauptmann, 2004) to varying extents and maintains a national RailSys model. Currently, RailSys is the main tool for capacity analysis and evaluations at TRV. RailSys is also used by academia in Sweden, mainly at KTH. How and for what RailSys is used for at KTH varies. Examples of projects where RailSys has been used is as a simulation engine in a timetable optimization framework where timetables are adjusted based on their operational performance simulated in RailSys. Some projects are more focused on signalling, whereas in other projects data from the national RailSys model is mainly exported and used as input into other models.

HESE is a tool intended for calculating technical headways in ETCS L2 setups. It was developed within a research project and the aim was to get a tool specifically for calculating technical headway for different setups regarding marker board locations. One of the main aims was that it would be easy to vary track, signal and other parameters. Calculation of the different warning, intervention and braking curves is implemented according to ETCS SRS 3.6.0 described in ERA (2016). During development the calculations for different curves were validated by comparing output for a set of different cases between this tool and ERA Braking curves simulation tool (ERA, 2018). Technical headway can also be calculated with, e.g., RailSys which has ETCS L2 implemented. Verification of the headway calculations is therefore done against RailSys (see section 7.2.2).

Infrastructure input in HESE tool is given in Excel-sheets with line speed profiles, gradients, marker board (signal information point) locations and balise locations Multiple sets of marker board and/or balise locations can be given to facilitate visual comparisons between cases regarding technical headway. ETCS national values and other system parameters are also imported from an Excel-sheet. Train data is imported from RailSys train model files, this includes traction force, braking, and other parameters.

Notably, HESE lacks a graphical interface, relying on scripts for functionality. Despite this, the primary output is graphical. The tool handles cases where two trains run either entirely on the same track or an overtaking situation where the first train vacates the main track at a turnout and the following train continues the main track. Figure 2 provides an example of output from the tool.

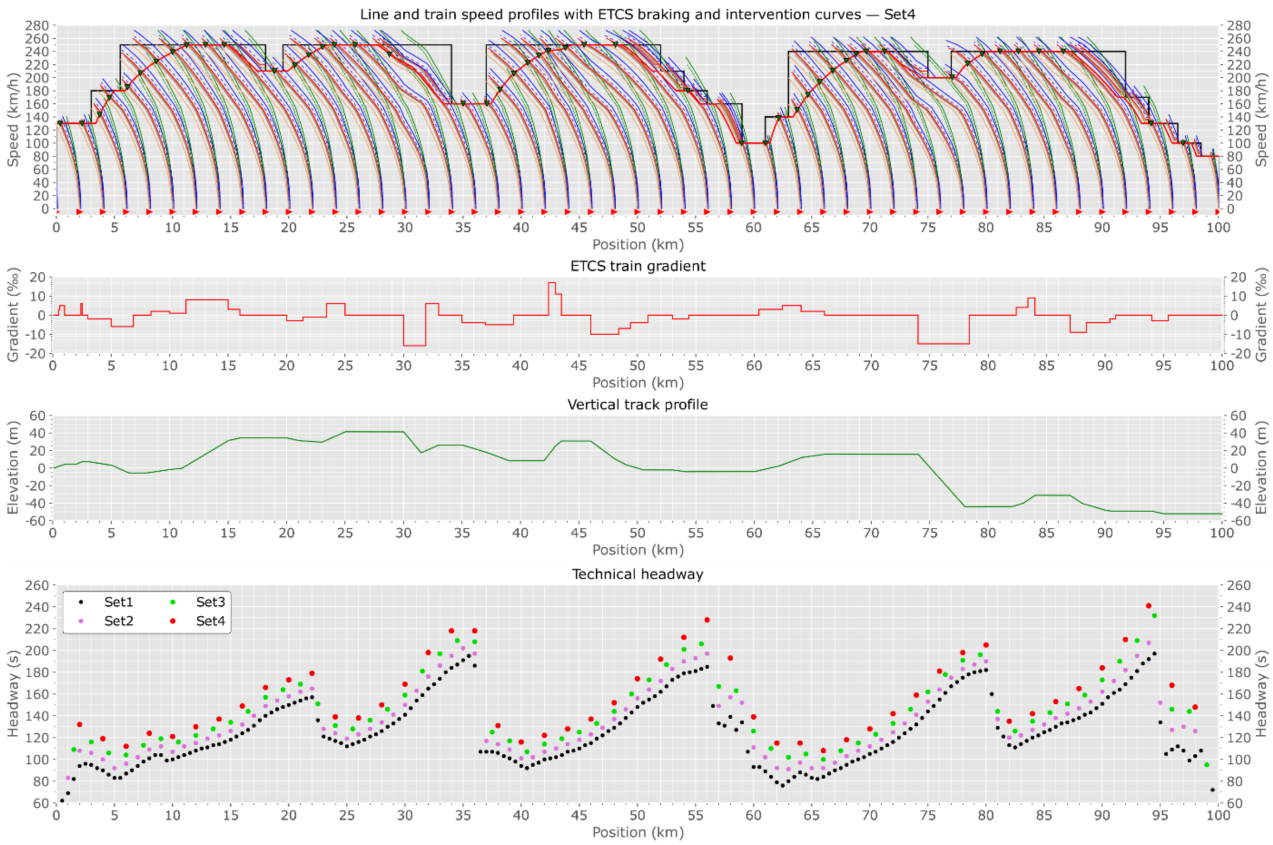


Figure 2: Example of output from the HESE tool. Line and train speed profiles, train gradients, vertical track profile and technical headway, in this example for four sets of marker board (block section) distances (500, 1000, 1500 and 2000 m).

6.4.1 Development needs at TRV A.E. Lund/KTH

Simulation of large full networks or large areas within networks is an interesting topic as it enables the capture of network effects that may be challenging to assess through limited area simulations alone. In this project, we aim to primarily simulate large geographical areas of the Swedish national rail network for one or multiple historical timetables.

We also plan to improve the modelling of primary input delays. For instance, we plan to explore the relationship between scheduled turnaround times (vehicle chaining) and departure delays, aiming to enhance the precision of initial distributions used in simulations. Furthermore, we will categorize input data into days with lower and higher levels of delays to better capture the entire range of punctuality across days/simulation cycles and investigate the correlations between primary delays. This can give a better understanding of different risk factors, focusing on better capturing the range of punctuality of trains within days/simulation cycles. To achieve these goals, emphasis will be placed on refining the calibration and validation of simulation models. This includes, among other things, considering impacts of vehicle chaining that were deemed necessary, using different scaling factors for entry, run, and dwell time delays and more finely differentiate the delay distributions in terms of train groups/types and locations.

A broader aim is to improve the conditions for being able to simulate one or preferably multiple timetable proposals in the future, for large areas and up to the whole Swedish network and thereby contribute to a feedback loop between planning and, through simulation, estimated operational outcome. This could primarily be used to compare (rank) different timetables and, e.g., indicate areas with high delay growth.

6.5 Simulation usage at TRV A.E. VTI

VTI will use the train driving simulator (Figure 3) and the EPA tool (Figure 3) in WP8/WP9.

The train driving simulator is available in more than 60 copies in Sweden distributed among 12 train operating companies (TOCs) and is developed by VTI. Different versions of hardware meet different needs for users, some prefer portable desktop simulators (flexible use) while others have simulators that reflect a specific train model (The train in Figure 3 is a replica of a well-known EMU in Sweden, called Regina). It is mainly used by the TOCs for training and examinations, and at VTI for research on human behaviour. A selection of this research is presented below:

- Effect of ERTMS Baseline 3 Release 2 (SRS version 3.6.0) (compared to Baseline 3 Maintenance Release 1 (SRS version 3.4.0)) on driving and braking behaviour.
- Effect of different level of ERTMS filtering on attention (eye-movement study) and compared to lineside signalling (Kircher, et al., 2022).
- Effect of train driving simulator practice in ERTMS on driver performance (Olsson, et al., 2022).
- Validation study of the simulator at the user level where driver performance in the simulator environment was compared with performance in real train driving.

Figure 3: One hardware version of the VTI Train driving simulator.

The EPA (ERTMS Protocol Analyser) tool is an VTI inhouse developed train dynamics measurement tool. The tool method is based on train and track information sent between RBC and ETCS onboard equipment, enabling studies on train routes over time (Rosberg, et al., 2022). It is cost-efficient as a considerable time saving process can be applied for all ETCS data collections, meaning that data collections can be performed remotely and independent of the railway operator. The main research objectives with this tool have been:

- Train driving behaviour – compared to standardized ERA braking curves.
- Train driving behaviour – impact on different rail planning strategies, for example impacts of different speed profiles, release speeds, etc.
- Train driving behaviour – impact on different ETCS Levels/versions.

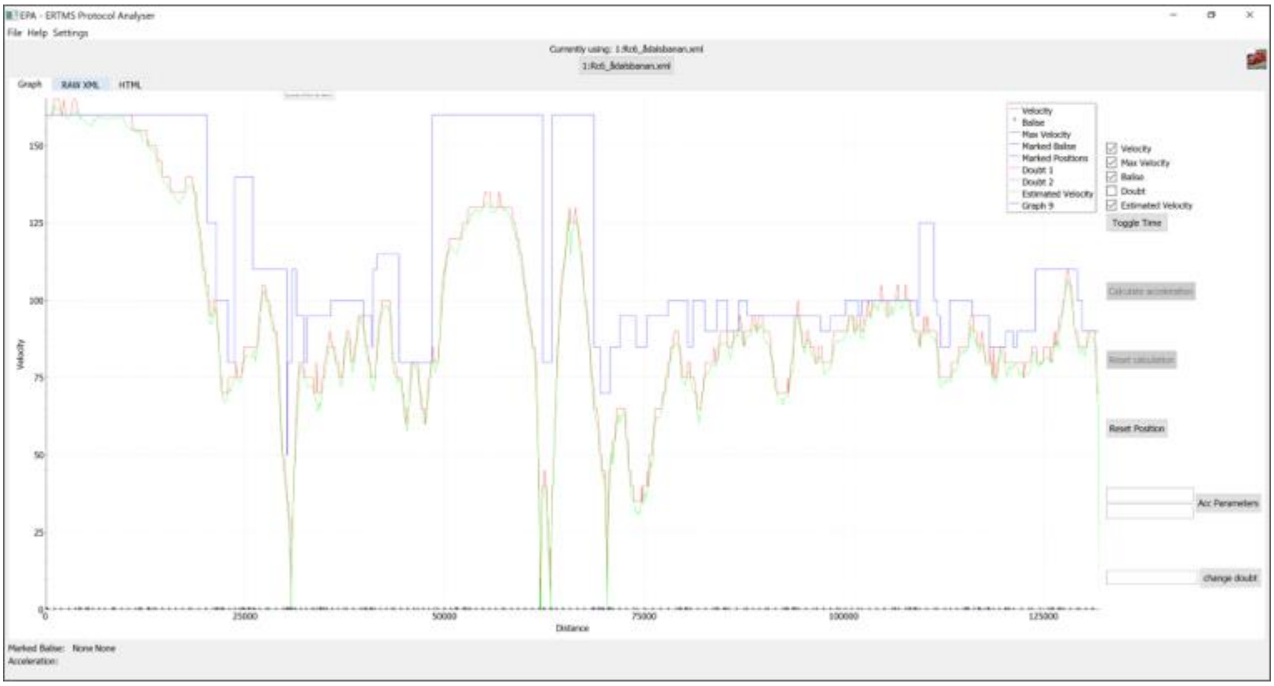


Figure 4: EPA graphical user interface

6.5.1 Development needs at TRV A.E. VTI

Within WP8/WP9, the VTI Train Driving Simulator will be used to study the effect of a new ERTMS implementation strategy aiming to reduce traffic interruptions at the time of implementation. VTI will engage train drivers to study the effects of the new strategy in terms of capacity, driveability, and safety. In order to carry out the study, a development of the simulator environment is needed in accordance with the planned implementation strategy. There is also a need for development in the area of C-DAS and ATO functions within the existing simulator software.

The EPA-tool is dependent on the ETCS Baseline version, and today up to Level 2 SRS version 3.4.0 is supported. Support for higher SRS versions (i.e., 3.6.0) needs to be added to enable reading of radio logs from tracks with this baseline (for example the Ådal line, Sweden). There is also a need for reading freight logs. Due to their slightly different structure compared to passenger logs, the EPA tool have to be developed to handle both types of logs.

6.6 Simulation usage at ProRail

Timetables are designed within the planning systems using a static timetable check on running times and conflicts. This is based on best-practices learned in Japan and the European ON-TIME project (ON-TIME, 2013). Medium to large infrastructure projects and yearly timetables are checked using a deterministic simulation, without disruptions or stochastic behaviour. The method was originally developed using the simulation tool OpenTrack. The static simulation check of yearly timetables and some building phases has been taken over by the simulation tool FRISO. FRISO is also used for human in the loop simulations. RailSys is used for more in-depth (microscopic) analysis of technological development projects such as ETCS/ERTMS braking behaviour, Hybrid Level 3 and ATO. Recent research reports regarding ERTMS HL3, written in cooperation with Delft University of Technology can be found in Jansen (2019) and Vergroesen (2020).

It has taken a huge effort to identify the differences between different planning and simulation tools and develop confidence about verification and validation with realisation data. This has led to the insight that a timetable planning or infrastructure study should not rely on only a single planning or simulation tool.

6.6.1 Development needs at ProRail

Focus of simulations will be related to work packages of FP1 and FP2 for the development of ERTMS, TMS and ATO. Regarding ERTMS, the simulation methods are already quite well developed, clear and implemented in RailSys. Questions are instead ERTMS design related, for example: How to design the layout of tracks with a mix of physical and virtual block ideally under HL3? When it comes to TMS and ATO, we know how to model trains according to general principles or braking curves, but TMS and ATO add more dynamic behaviour and interference of trains. It is not well developed how to include these effects in static planning tools and deterministic simulations. These insights have to be developed here.

6.7 Simulation usage at NS

RailSys usage at NS

RailSys is used at NS since 2015 and it was a research result of the RailwayLAB (collaboration ProRail and NS innovation). RailwayLAB did research regarding potential tooling for microscopic (considering a detailed modelling of the infrastructure including functionality of the interlocking, safety and block system) timetable design and simulation due to the scarce available infra capacity in the Netherlands (high-capacity utilization). NS applies RailSys for the following research areas:

- Microscopic timetable design: investigate the potential of using blocking times instead of headway norms during timetable design
- Integral timetable design: investigate the potential of considering both network and shunting timetable design at a microscopic level
- Stochastic microscopic simulation: investigate the relationship between norms (running, buffer and dwell time) and operational performance

- Timetable stability and robustness analysis: RailSys is used in production (replacing OpenTrack) since 2022 to evaluate the (annual) timetable stability and robustness, and to compare the results with realization data. Timetable feasibility is defined as realistic running times and conflict-free timetable based on blocking time theory (Pachl, 2009, 2014). Timetable robustness is defined as that the timetable can cope with design errors, variations of parameters and changing operational conditions (Goverde, et al., 2016).
- ERTMS and ATO: ProRail investigates the capacity effects of ERTMS (Hybrid Level 3/HL3) and ATO (Automatic Train Operation). NS assists ProRail in this research area, which is also part of this work package. Recommended improvements for RailSys in this area are: fine tuning ETCS braking, ETCS HL3, and ATO parameters.

FRISO usage at NS

FRISO is a microscopic timetable simulation tool developed by Incontrol for ProRail, purposely built for the Dutch railway system. NS uses FRISO for the following cases:

- Chain simulator ("ketensimulator"): connection of FRISO with traffic control and train driver simulators. It is part of the Railway Gaming Suite, where FRISO simulates the train traffic. Trains can be controlled by a driver or by FRISO, and traffic control can be done by with a traffic control simulator.
- Introduction of ATO: determine the potential conflicts with different driving strategies.
- Introduction of ERTMS: timetable feasibility and stability with ETCS (generally level 2) compared to legacy train protection system (ATB)
- Stochastic microscopic simulation: investigate the relationship between norms (running, buffer and dwell time) and operational performance.
- Timetable stability and robustness analysis: FRISO is used in production (replacing OpenTrack) to evaluate timetable stability and robustness for mid-term and long-term timetables.

Stochastic microscopic simulation

Since 2022, NS has been conducting research on stochastic microscopic simulation with the following aims:

- Gain insight in relationship between norm times (running time supplements, buffer time and dwell time) & operational performance.
- Achieve more in-house experience with (stochastic) simulation
- Compare the differences between microscopic simulation tools (FRISO & RailSys)

For this research, NS considered the following main assumptions:

1. No traffic control measures are included: NS only would like to see the effects of the timetable.
2. Other train operating companies: NS is not allowed to have information regarding their performance, so NS does not have their delay distributions.
3. NS basically considers the complete Dutch railway network, except for one of the reference scenarios. Note that NS uses RailSys only for reference scenario with regional network due to a high amount of manual work (i.e., no automatic timetable or delay distribution import).

External consultancy companies with experience in stochastic microscopic simulation are involved to help NS in developing the method and to evaluate the results. A general description for setting up a simulation model can be found in Medeossi and De Fabris (2018). In the research project of NS, the following methodology for stochastic simulation is considered:

- Analyse the static timetable by checking the feasibility of the running times and conflicts based on block occupation.
- Apply a deterministic (single) simulation run without minimum buffer times to see if the model is behaving as expected
- Apply multiple stochastic simulation runs to evaluate the timetable robustness (iterations):
    a. Determine the different delay distributions for entry, dwell, departure and running time using realization data (model calibration). Also determine the driving strategy which is a balance between realistic behaviour, and the possibilities of the software and the main assumptions. At least consider peak hour(s) in the simulation (filtering realization data).
    b. Determine the minimum number of stochastic simulation runs (dependent on the size of the network)
    c. Step-by-step build up the complexity of the stochastic simulation runs:
        - Include dwell time disturbances
        - Include running time disturbances
        - Include departure time disturbances
        - Include entry delays
    d. Run multiple stochastic simulation runs and check the number of deadlock situations (at maximum 10%)
    e. Evaluate results of stochastic simulation run using the 20th, 50th, and 80th percentile of arrival times of trains (NS does not include punctuality, because big delays are limited simulated since NS excludes traffic control measures)
        - Network level
        - Regional level
        - Specific trains.
    f. Validation for the simulation results with realization data (if available) or compare simulation results with a reference scenario
    g. Iterations of steps a—f to improve the stochastic simulation runs
    h. Stop the iterations when either of following events occur:
        - The 20th, 50th, and 80th percentiles are in line with realization data;
        - The main differences between simulation and realization can be explained, and are accepted by the experts;
        - There is insufficient time for the project.

Based on the research it can be concluded that stochastic microscopic simulation might be very interesting for light disturbances in the timetable, see **Fehler! Verweisquelle konnte nicht gefunden werden.**. For undisturbed scenarios of the timetable a deterministic simulation would be better. For larger disturbances stochastic microscopic simulation is not very suitable on a network scale due to limitations in dispatching in the software. Macroscopic simulation is a solution for situations where a lot of dispatching is done.

Table 2. Relationship between delay and type of type of microscopic simulation

Scenario	Delay	Example	Deterministic micro simulation	Stochastic micro simulation	Stochastic macro simulation
Undisturbed	< 1 min	On time running	++	+	--
Light disturbances	1—3 min	Peak hour traffic	+	++	-
Big disturbances	> 5 min	Stopped trains	--	+-	+
Major disruptions	>20 min	Track blocked	--	--	++

Macroscopic modelling: usage of INTROOS at NS

Within NS macroscopic modelling is used for estimating punctuality effects of new developments. For timetable simulations, the model INTROOS is used. Example cases are:

- Punctuality loss due to the introduction of ERTMS L2 during training period with or without fallback option
- Introduction of new rolling stock (e.g., ICNG and SNG) and estimating punctuality effects in different scenarios
- High-level assessment of the benefits of extra infrastructure in regular and disturbed situations

These simulation models are developed within NS using generic simulation software (AnyLogic, Java-based). NS derives probability distributions from realisation data to model different disturbances, for example late starting, running time delays, dwell time delays.

The value is provided in two ways: the simulation visualises potential critical situations, thus providing insights. The second way is by simulating a large number of days (usual more than 1,000) using Monte Carlo Simulation, which gives a range of the expected punctuality effects.

Additionally, macroscopic models are used for non-timetable related questions. NS investigates the optimal number of maintenance facilities and where they should be placed. The shunting process and rolling stock dispatching are simulated using purpose built macroscopic models.

Energy-efficient train driving

The driving strategy of a train has a big influence on the performance of the timetable, such as the punctuality and capacity utilization, and on the energy consumption of trains. Therefore, train operating companies investigate energy-efficient train driving to minimize their traction energy consumption, while running on-time (i.e., not too early and not too late). The reader is referred to Scheepmaker et al. (2017) and Su et al. (2023) for more background information regarding the topic of energy-efficient train operation. Scheepmaker and Goverde (2021) showed that there is a balancing relationship between the available running time and energy consumption, without influencing the infrastructure occupation and timetable robustness. Furthermore, the development of C-DAS and ATO enables big potential for energy-efficient train driving, especially in the compliance rate and the usage of advanced advice. Since timetable simulation results should correspond with realistic behaviour, energy-efficient train driving should be considered as a driving strategy.

Therefore, NS investigates to incorporate energy-efficient train driving in timetable design. This means that also during timetable simulation (to evaluate the developed timetable) energy-efficient train driving should be incorporated. However, the majority of available timetable simulation tools on the market do not currently integrate energy-efficient train driving, but are mainly focussed on cruising (reduced maximum speed driving strategy, see Figure 6-5). Therefore, there is a need to incorporate energy-efficient train driving in simulation tools.

Current application of energy-efficient train driving at NS focussed on coasting, which involves putting off traction and allowing the train to roll out (see Figure 5 maximal coasting driving strategy), where research shows that for the Dutch situation this driving strategy approaches the energy-optimal driving strategy (see Figure 6-5 energy-efficient train control) (Scheepmaker, et al., 2020). Furthermore, coasting advice is consistent with the experience of the train drivers and has a high acceptance rate among them. This energy-efficient driving strategy is incorporated in FRISO but is missing in RailSys.

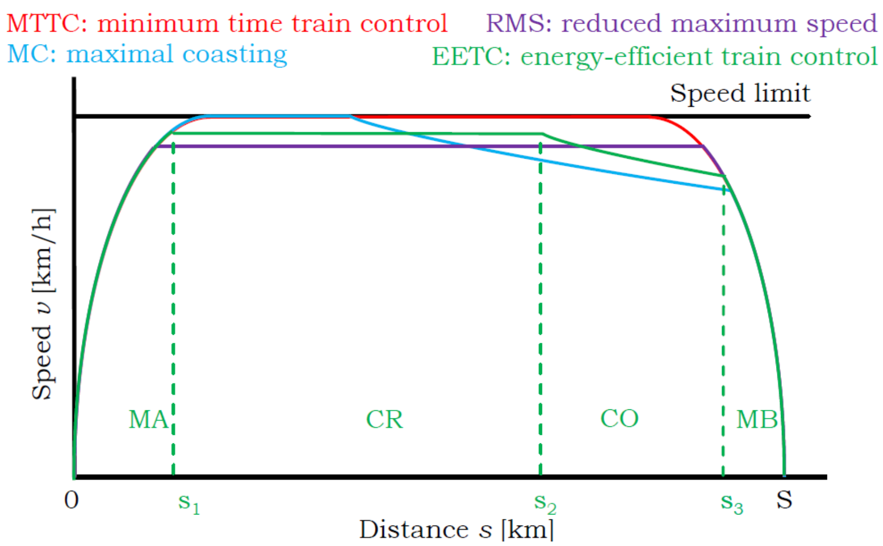


Figure 5: Different (energy-efficient) train driving strategies with different driving regimes (MA = maximum acceleration, CR = cruising, CO = coasting, MB = maximum braking)

6.7.1 Development needs at NS

Development needs of capacity simulators

Although NS can use current simulation tools for timetable evaluation and capacity analysis, there is still space for further improvement. First, energy-efficient train driving by coasting should be incorporated in all microscopic simulation tools, because this is the driving strategy normally applied by NS train drivers in operation.

A first version of coasting behaviour is already available in FRISO, but this driving behaviour will be improved to also consider critical passing points in the timetable as a constraint for the algorithm. RailSys does not consider the coasting driving strategy, so here the specification of the coasting driving strategy is required, followed by development and testing.

Second, during the stochastic simulation research, NS found that RailSys is not able to maintain the order of trains according to the timetable. This is needed for timetable evaluation at NS, where NS does not consider dispatching for small delays (e.g., less than 5 minutes), because else the results of the simulation runs do not represent the operational performance. The Dutch NS '54 signalling and train protection ATB system includes automatic route settings (in Dutch called "automatische rijweg instelling" or ARI) for trains given by the traffic control, that maintains the order of the timetable even if they are delayed. During delays the system informs the dispatcher who might interfere the automatic route setting (i.e., the dispatcher might manually change the order of trains then). RailSys has some features to maintain the order of trains, but they require a lot of manual work, and they cannot be applied on junctions where two trains interact and do not have a scheduled stop. Not keeping the order of trains as scheduled leads to bigger delays in the simulation compared to realization. In addition, deadlocks occur when the train order is fixed at specific stations and the order is changed at a junction before this station.

Third, another improvement for RailSys is to run before the scheduled timetable. RailSys does not enable negative values for delays (i.e., being too early) or to apply the minimum running time driving strategy if trains are running on-time. However, in practice a lot of trains In the Netherlands run before schedule and thus have more running time. As a result, the delays in operation are in general lower compared to the simulation output of RailSys. To have a representative behaviour, NS needs to model them in the simulation tool as well.

Development needs of simulator of transport plan (AnyCrew)
NS together with SISCOG is developing a new simulation model, called AnyCrew, for simulating a transport plan with focus on crew. This model is being built from scratch. NS can reuse some components (data structures, agent state charts) from other macroscopic models, notably INTROOS. The development consists of:

1. Nation-wide macroscopic infrastructure representation
2. Timetable import (from timetable design system DONS)
3. Rolling stock plan import (from rolling stock system TAM)
4. Crew plan import (from Crews)
5. Matching the timetable, rolling stock and crew plan
6. Introduce a representative distribution of delays
7. Develop performance indicators to evaluate the robustness of the crew plan

All aspects of this development are being carried out in-house by NS and SISCOG.

6.8 Simulation usage at SNCF

Microscopic simulations using DENFERT
A customized version of RailSys called DENFERT has been implemented at SNCF and is used both on the IM and RU side. It is used at SNCF Réseau in order to carry out operational studies with multiple use cases:

- To monitor the performance of the network in case of modification of infrastructure or signalling mode or rolling stock (e.g., ERTMS implementation and block modification)

- To meet a need for a new offer: to match the current and projected infrastructure with the timetable studies in order to best meet the needs of the customer while maintaining the performance of the network
- To define future infrastructure developments (new line etc.)
- To assess the robustness of a transportation plan
- To help projects with their work phasing
- To evaluate the feasibility of a new commercial offer (stops at certain stations) compared to the current infrastructure
- To measure the impacts of the track works on the transport plans in case of e.g., accessibility works, traffic interruptions and speed limits
- To look at the impacts and the capacity of a station in case of works on the line or in the station

These studies consider both the infrastructure, signalling, rolling stock and expected service in order to produce an output schedule. For each new study, the geographical zone considered is modelled in DENFERT with the target infrastructure, models can be reused if multiple studies are carried out on the same zone. Once the situation is modelled, different scenarios can be explored, with modifications on the infrastructure, and recommendations can be made based on relevant criteria.

Stochastic methods

In order to assess the robustness of the target transport plan, stochastic methods are used in DENFERT. These methods incorporate delay distributions, driver behaviour and regulation decisions in the process in order to produce robustness indicator that evaluate the resiliency of the planned schedule.

For these studies, historical data of incidents and delays are analysed in order to calibrate the distribution of delays inserted used in the study.

OSRD Development

In parallel to the current development activities, SNCF is developing an Open-Source Railway Designer (OSRD), which is centred around a microscopic simulation engine that can do timetabling design. It targets multiple key moments in the operational timeline and includes an infrastructure editor in order to carry out exploitation studies. The following diagram summarizes the targets of OSRD (see Figure 6).

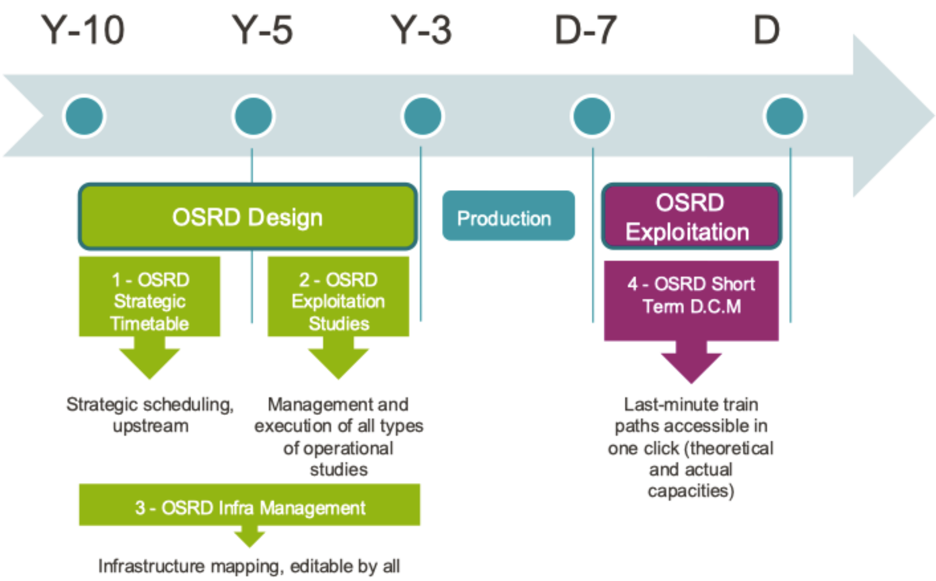


Figure 6: Diagram that summarizes the targets of OSRD

One of the ambitions of OSRD is to be interoperable at the EU Scale through the creation of the Open Rail foundation, that will be the governing body of the tool.

6.8.1 Development needs at SNCF

SNCF Reseau is interested into developing a better and easier way to carry out stochastic studies. This involves improving:

- The input data and how to obtain and analyse them
- Importing input data into RailSys interface
- Routing rules applied by RailSys, making them clearer and more efficient
- The output data from RailSys, to help analysing the results and improve the calibration

These developments involve RailSys, but we also would like to develop some apps and interface (e.g., Excel and Power BI) to improve data processing.

SNCF Reseau is also interested in participating in exchanges on RailSys development needs on ETCS-related topics. We have identified development needs, in particular for the simulation of disrupted situations in ETCS (e.g., On Sight mode, Staff Responsible mode and arrival on occupied track).

6.9 Simulation usage at ADIF/RENFE/CEDEX/INECO

ERTMS Traffic Simulation Laboratory

The ERTMS Traffic Simulation Laboratory at CEDEX is basically a test environment where it is possible to place industrial ETCS components (both on-board and trackside), following the Hardware in the Loop (HiL) approach. Up to now, the laboratory has been used to evaluate the interoperability between ETCS Trackside and On-board, as well as checking the Trackside ETCS engineering respects the Spanish Operational Principles. The laboratory has been extensively used in the Spanish ERTMS deployment, although it has also participated in some projects abroad, thanks to the confidence built with the railways industry.

Besides the simulation core which enables the integration of industrial ETCS components, the laboratory also comprises ETCS equipped train simulators, Interlocking simulators, RBC simulators and an Automatic Route Setting module, for a fully automated test environment. Figure 7 includes the ETCS components integrated in the ERTMS Traffic Simulation Laboratory environment.

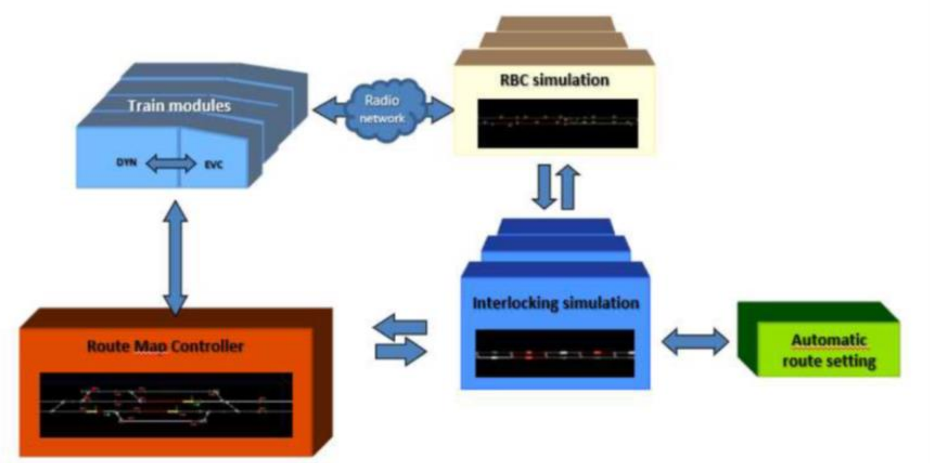


Figure 7: Current schema of the ERTMS Traffic Simulation Laboratory environment

The test environment relies on the precise characterization of track topology, primarily determined by track lengths and gradients. The train motion simulation is also highly configurable with a model parametrization based on actual dynamic train parameters.

Planif

In the Spanish Adif network, there are already some simulations. Adif owns an IT tool for planning (Planif). The aim of this tool is just to manage the planning activity within the rail network of ADIF and ADIF high speed rail network, but it is not a simulation tool.
This tool includes different modules:

1. Topology module for planning, which includes drawings and physical characteristics in the macro and medium level, for example, locations, speed limitations, track gauge or voltage of the track. This module does not include micro level details, such as interlockings
2. Rolling stock module, which includes the characteristics of current rolling stock which is allowed to run through the network, such as, speed, acceleration and deceleration curves, traction or track gauges
3. Management of trains module, which is used by the capacity staff to create timetables based on the points and distances of the topology module and the rolling stock module detailed above, taking into account tracks and stations

The topology module supports the functionality to model different sets of topological data corresponding to different situations (real-time, future, analysis), for example, including new points or changing some characteristics of the current topological information.

In the module of management of trains, the users could work in a workspace called "project" (which can be shared with different people or just for the creator). In these workspaces, the users can create trains using the active topologies or any of the different set available in the topology module. When a train project is shared it is open for modification, but the basic parameters such as composition, braking, etc. remain the same. In other words, in an inherited or shared project the variables related to the time slot, origin and destination can be modified, but not those that determine the running calculation. In this way, some simulations could be made, trying to identify some trains that could have a better performance or how to make some changes to identify how to improve the capacity of the network. The simulations and different attempts to get better results must be made manually, since there is not an automatic tool to make that or to collect the results.

The information of these projects remains in the planning system, and it is not exported elsewhere, mainly because the TMS or any other planning data consumer have not a similar environment to take advantage of this information.

6.9.1 Development needs at ADIF/RENFE/CEDEX/INECO

ERTMS Traffic Simulation Laboratory environment

The ATO technology in the Control, Command and Signalling TSI (ERA, 2023), brings some challenges into ERTMS Traffic Simulation Laboratory environment. In the first place, it is the CEDEX intention to provide our test environment of all the modules necessary to test, at least the ATO onboard. However, the ATO comes with additional technological advances:

- Train positioning, taking advantage of accurate Galileo satellites geolocation services
- Train-Track communication, to cope with the ATO exchange, expected to be "heavier" than the ETCS exchange
- Accurate and Harmonized Track Description, in order to build the ATO Journey and Segment Profiles

From CEDEX point of view, the addition of these features in our test laboratory in the following years will expand our capabilities to enter into some aspects of the operational world: the timetable planning and the traffic control. As onboard equipment is only one of the two sides of the whole ATO system, the test results in one side will increase the reliability to tests done in the other.

The CEDEX Laboratory will also be involved in several WPs in FA2 and FA6, as well as some critical WPs from the Transversal Topics (or Digital Enablers) in MOTIONAL.

In this context we are strongly collaborating with the SCNF LEF and DLR RailSite laboratories in FA2 and expect to find synergies with other laboratories/universities/partners in MOTIONAL.

6.10 Simulation usage at Indra

Indra´s simulation environment is a functionality that belongs to the TTCMS, the global system that provides functions to plan, monitor, manage and control the rail traffic of the whole network.

The simulation environment allows operating with the TTCMS in a normal way by connecting to a virtual environment that simulates the complete evolution, in a coherent and synchronized way, of all the field elements involved in railway operation.

The responsibility of the simulator is to enable the user to replicate situations arising in the commercial operation of a railway line, without being connected to the real environment, thus being able to interact with the systems without interfering with the operation of the line. This allows users to train and learn from the different situations that can occur while operating without causing perturbations to the current real operation.

Indra's simulation environment allows testing different timetables and different topologies. The simulator uses different inputs to create the most accurate environment. It should be highlighted:

- Topology network data including:
    - Set of interlockings present in the system:
        - Information of each interlocking: the information of all the elements of the interlocking and their relationships
        - Track circuits
        - Points
        - Signals
        - Line sides
        - Routes available in the interlocking:
            - Start and end of route
            - Associated routes
            - Sub overlaps
        - Level crossings
        - Signalling rules: rules necessary for signal status to reflect track conditions
    - Connections between elements: information about the connection between elements to be able to move the train along the track
    - TRP: (Train Regulation Points) these points are needed to move the scheduled trains so that they comply with their schedule
    - Slopes: needed for the calculation of the train movement
    - Speed limits by infrastructure: needed to adjust the train movement
    - Detectors: hot boxes (HABD)
    - Energy: list of catenary elements

- Timetable with two additional parameters:
    - Period: range of dates in which the plan is active
    - Circulations: list of circulations for each of the days in which the plan is active:
        - Service number: service identifier. A service represents a train of a certain type that performs a route on a certain date and with a defined schedule.

- Train type: one of the types defined in the "Rolling stock" database. The necessary values will be used to calculate the movement of the train.
- Route: list of points ordered according to the route the train is going to take. The first point will mark the place and time of departure and in the last point the place of arrival and target time for the end of the service. The following information is required: identifier of the waypoint, arrival time and departure time. At those points where the arrival and departure times are the same, it means that the train will not stop.

- Start time: initial time of the simulation

This simulation tool is useful in order to identify initial problems with the timetable and the topology, but with a limited scope. The Figure 8 below summarizes the process for a simulation exercise including the involved modules.

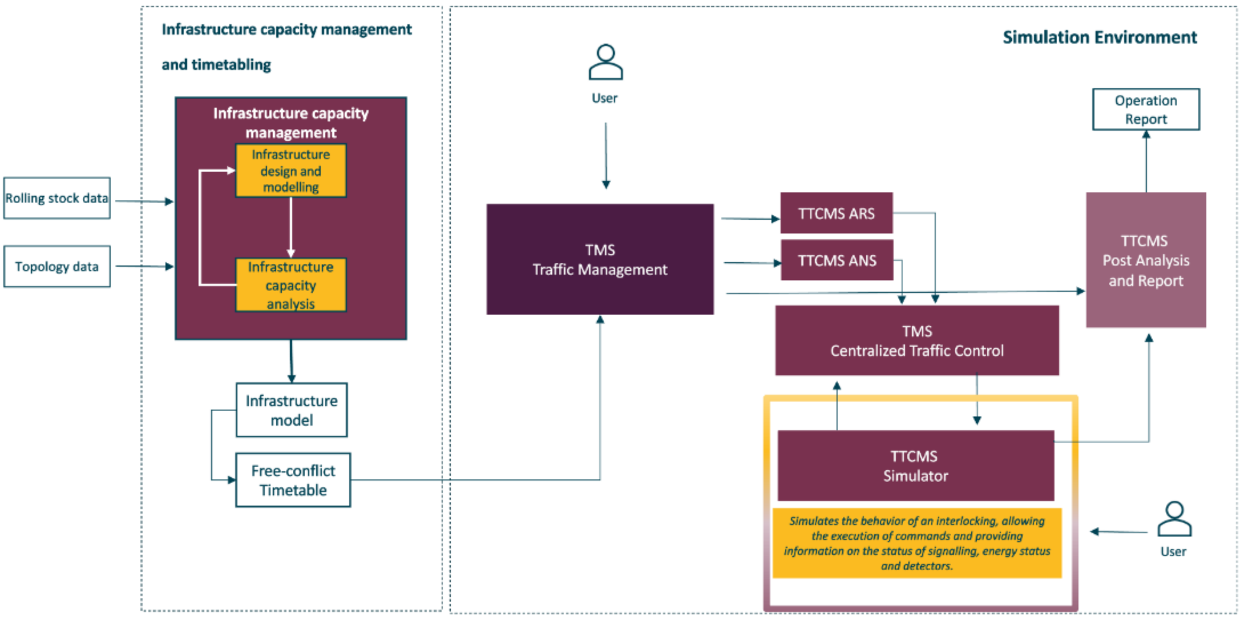


Figure 8: Architecture of Indra's simulation environment

First, a simulation is performed using the simulation environment with a specific topology and planning, specific network topology and timetable (as a source for the simulation environment). As a result of this simulation, poor planning behaviour is obtained with that topology. The Infrastructure Capacity management tool identifies bottlenecks and necessary infrastructure changes. With new infrastructure model and after timetabling we again use the simulation environment. Finally, it is identified if the simulation improves/worsens with the new topology.

The Infrastructure Capacity Analysis tool will perform the analysis of theoretical timetables defined by the user with different versions of topology and rolling stock. It informs about train movement, block times, critical block sections information and the details of the slopes, speed restrictions and conflicts found in the timetable.

Besides it is studied how to calculate the optimal capacity between stations applying the criteria defined by the UIC405 standard (UIC, 1996) and using the different ATP technologies (including ERTMS L1 / ERTMS L2).

The aim is to evaluate the infrastructure layout by developing the Infrastructure Capacity Analysis (TMS_CAP) and demonstrate how it helps to timetabling. TMS_CAP will support most of ATP systems, applying their behaviours to the calculation and signalling procedures:

- ERTMS Level 1
- ERTMS Level 2
- Different driver modes
- C-DAS

It ensures a high level of adaptability to the actual operating conditions of the railway infrastructure.

To sum up, Figure 9 shows the elements involved.

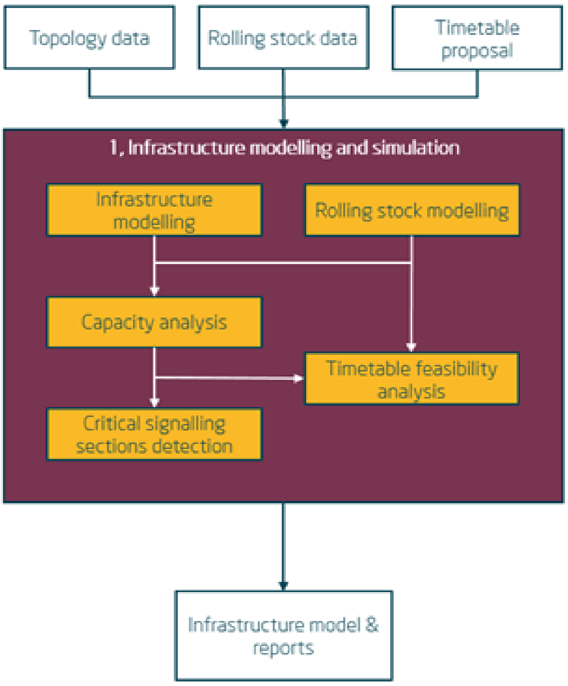


Figure 9: Infrastructure modelling, simulation, and deployment process overview

6.10.1 Development needs at Indra

The main use case in which Indra will be involved focuses on demonstrating that different timetables can be optimised by updating topology networks, increasing consequently capacity. The objective of Indra is to develop a capacity analysis tool and test different models and scenarios. In this use case we will demonstrate the performance of a capacity analysis tool with several topologies and timetables. In case of poor performance by simulating with a specific topology, the capacity tool is used to accurate identify bottlenecks that can be solve by topology network changes. Iterative simulations with different topologies facilitate timetabling.

6.11 Simulation usage at CAF

CAF simulation tool does not have a commercial name because it is only used internally; it is a capacity analysis tool through simulation. This tool is designed to serve two different purposes:

- On the one hand, to have it internally as a testing tool that allows us to validate the advances, we accomplish in TMS. In this case, the focus will be on the automatic system regulator, which will be in charge of managing disturbances.
- On the other hand, it will serve as a tool with commercial value to assess the feasibility of the needs presented by a customer or in the pre-project phase. Here, the focus will be on the planning capacity to see whether or not satisfies the pre-set conditions.

This simulation tool will give us some valuable information that could be used as an input for the planning tool that we are going to develop in this WP. This new planning tool will have a particular functionality that will allow us to plan in a mixed way, considering the train position (space) or the service hour (time).

The functionality will consist in the following:

- In the case of time, there are two different cases, when it is a rush hour, traffic will be planned and regulated by headway. While if it is an off-peak hour, it will be managed by timetable.
- Considering the space, we must take into account the area through which the train runs. It may be the case that the train changes from an area regulated by headway to another that does so by timetable. This can be found on a long line that runs through central areas of the city and through more distant areas or branch lines.

6.11.1 Development needs at CAF

Some planned improvements of this tool are, among others:

1. Blocks with a pre-defined size, which enables results similar to Moving Block (MB), when simulating it offline.
2. Calculation of the minimum interval in rush hour of the complete network (generating the timetable in rush hour).
3. Calculation of energy consumption and thus be able to use it as input to obtain the planning.

All these improvements are going to be developed in this WP.

Artificial Intelligence will not be used for the development of this simulation, instead deterministic algorithms will be used.

The TRL with which we start is very low, but with the improvements we have mentioned our goal is to increase it. In order to test it, we will develop some use cases. Once we have the capacity simulation tool and the planification tool develop, we will generate a simulation environment to validate the plannings that we are creating.

6.12 Simulation usage at CEIT

OPTICON is an energy optimization tool developed by CEIT under the Spanish national research program Retos Colaboracion. This software is designed to assist decision-making in optimizing energy consumption in railway systems in a holistic manner, making it highly valuable for railway and infrastructure operators looking to minimize energy usage in their systems.

The software is built using Matlab® and is divided into three main modules: vehicle, grid, and energy bill (see Figure 10), each with a distinct focus.

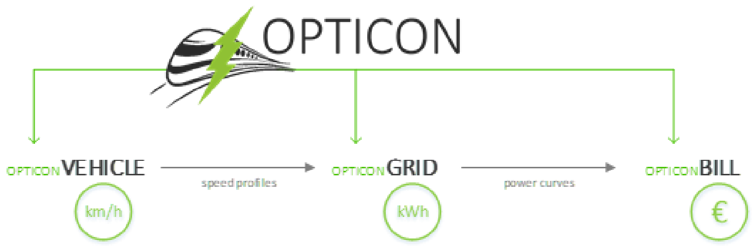


Figure 10: Overview of OPTICON main modules

OPTICON Vehicle: This module is focused on train dynamics and the generation of energy-efficient speed profiles, taking into account vehicle traction equipment characteristics and track profiles. It requires input on the dynamic and electric characteristics of the vehicle and the macroscopic data of the route, such as gradients, curves, tunnels, and speed limits.

The module calculates minimum running times and energy-efficient driving profiles for a specific running time, using different optimization algorithms and strategies. Smart traction-coasting cycles are used for large interstation distances, considering traction efficiency map, while an intelligent use of traction-cruising-coasting-braking modes is employed for shorter interstation services, taking into account track gradients. A lot of work has been applied into developing a fast and computationally-efficient algorithms for its use in onboard DAS systems for driving assistance.

Figure 11: OPTICON_Vehicle module main interface and energy-efficient speed profile generation

OPTICON_Grid: This module is focused on the movement of all vehicles involved in the system and power flow calculation, taking into account the energy demand of all the trains and calculating the energy consumption at energy supply points. It enables a static analysis of timetables in terms of running times and conflicts, such as the minimum distance between adjacent trains.

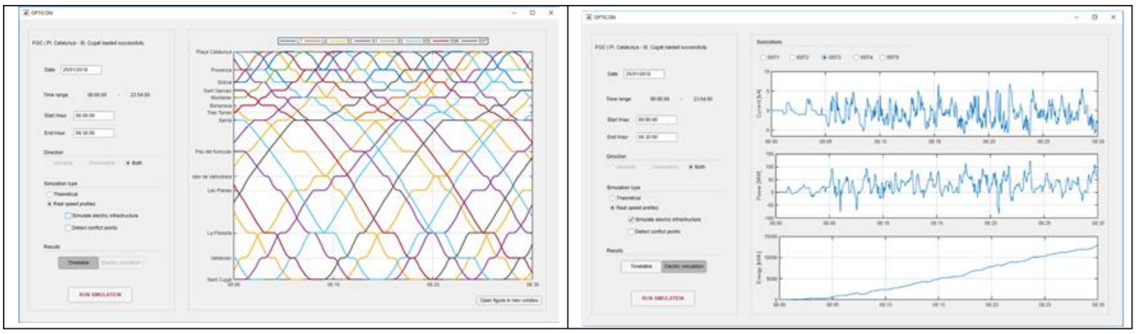


Figure 12: OPTICON_Grid module main interface with simple timetable analysis and power flow resolution

OPTICON_Bill is a module that provides a small but essential tool to optimize the electric bill by analysing the real power demand at power substations. This tool helps to find the correct balance between contracted power and power excesses, which in turn optimizes the final energy bill. Additionally, the module takes into account the tariff system used in each country to ensure the most efficient and cost-effective solution.

6.12.1 Development needs at CEIT

Current and future work focuses on improving the planning capabilities of the tool, taking into account the interaction with C-DAS systems and shifting or extending the current energy focus of the tool towards capacity analysis.

6.13 Summarised usage, functionality and development needs

From section 6.2-6-12, we can see that several different tools are used. For microscopic simulation, several partners are using RailSys (NS, ProRail, TRV) or a modified version of RailSys (SNCF: DENFERT) as the (main) simulation tool. This is a microscopic simulation tool that can do static, deterministic and stochastic analysis and that support ETCS calculations. ProRail also uses OpenTrack, another commercial microscopic simulation tool. In addition, FRISO is used by both NS and ProRail, which is a homemade microscopic simulator that can do deterministic and stochastic analysis for ATO simulation and to integrate traffic control. SNCF is currently developing its own in-house open-source microscopic simulation tool OSRD.

VTI provides a train driver simulator to evaluate how driver behaviour affects the capacity. This simulator includes ETCS signalling. Furthermore, they have developed a tool EPA for analysing real driver braking data for an ETCS railway line. CAF has a capacity simulation tool designed for testing of TMS from capacity perspective, where they plan to develop functionality for analysis of headway and energy consumption. Indra will develop their own capacity analysis tool (TMS-CAP) for evaluation of different infrastructure layouts to get feedback to the TMS and timetable.

Regarding macro models, the picture is more fragmented, as no tool seems to be used widely around partners. NS mentions using INTROOS (in-house development) while TRV uses PROTON. CEIT uses OPTICON, another macro simulation tool for timetable energy analysis and running time calculation, as well as energy-efficient speed profile generation.

The capacity simulators have different capabilities. An overview per tool can be seen in **Fehler! Verweisquelle konnte nicht gefunden werden.** for the microscopic simulator tools and in **Fehler! Verweisquelle konnte nicht gefunden werden.** for the macroscopic simulator tools. “X” in the table means that the simulator has the required capability, “D” indicates that development is needed before functionality exists and “(X)” shows that only limited development is needed. If a cell is blank, the simulator tool has not the described functionality at all or the tool is not suitable for that type of functionality.

Table 3. Capabilities and development needs of different micro simulation tools

Model name	Simulation of transport plan/ feedback loops	ETCS L2 (braking behaviour)	ETCS HL3	C-DAS	ATO	Driver behaviour/ strategy (coasting)	Train path envelope	TMS - dispatching functional-ities	Low adhesion effects
CAF Tool					X	X		D	
EPA	X	X				X			
ERTMS Traffic Simulation Laboratory		X	D		D			D	
FRISO	X	X	D	(X)	(X)	(X)	D	(X)	D
HESE	X	X	D		(X)	D			D
Open Track	X	(X)	X	D	D	D	D	D	(X)

OSRD (Open Source Railway Designer)	D	D	D	D	D	D	D	D	
RailSys	X	X	X	D	D	(X)	D	D	(X)
TMS_CAP		D		D		D			
VTI Train-Driver Simulator	X	X	D	D	D	X			D

For the microscopic simulation tools, we can see that simulation of transport plan, feedback loops and ETCS L2 are relatively well-developed with several tools capable of handling it. ETCS HL3, ATO and Driver behaviour are areas where a few tools already can handle it but where a lot of development is needed. The last group is C-DAS, Train path envelope, TMS dispatching functionalities and Low adhesion effects where none of the described tools have satisfactory functionality yet. In **Fehler! Verweisquelle konnte nicht gefunden werden.**, the capabilities for the macroscopic simulation tools are shown.

Table 4. Capabilities and development needs of different macro simulation tools

Model name	Simulation of transport plan/ feedback loops	ETCS L2 (braking behaviour)	ETCS HL3	C-DAS	ATO	Driver behaviour/ strategy (coasting)	Train path envelope	TMS - dispatching functional-ities	Low adhesion effects
AnyCrew	X								
INTROOS	X	X		X	X				X
OPTICON	D			D					
OSRD (Open Source Railway Designer)	D	D	D	D	D	D	D	D	
Planif	X								
PROTON	X				(X)				
TMS_CAP		D		D		D			

From Table 4, we can conclude that both functionality and planned development are at a lower level. INTROOS is the only tool having other capabilities than simulation of transport plan, and only limited development is planned for the other tools.

The tools' development needs are specified in different use cases, that in more detail describe the development planned in MOTIONAL WP8/WP9. The use cases are presented in Table 6-5. For the whole use case description, please see deliverable D3.1. In some cases, the use cases purposes are to develop the tools themselves for better analyses while other use cases aim to use the tools existing functionality to investigate new methods, conditions or new technique. The simulation tool in each use case is chosen to best give answers to the specific research question for that use case.

Table 5. Use cases in Motional WP8/WP9 and related simulation tool intended to be used

Nr	Formal use case	Tool	Partner
1	UC-FP1-WP3-30 – Improved railway traffic simulation models for capacity evaluation of ETCS	RailSys	SNCF
2	UC-FP1-WP3-31 – Feedback loop from simulation to planning for large scale networks	Proton	TRV A.E. KTH
3	UC-FP1-WP3-32 – Historical data analysis to improve traffic simulations and traffic planning	Proton	TRV A.E. Lund SNCF
	Demonstrate effect of ETCS level 2 roll-out strategy in terms of drivability, capacity and safety for:	VTI driver simulator	TRV A.E. VTI
4	UC-FP1-WP3-33 – Co-existence		
5	UC-FP1-WP3-34 – Normal ERTMS implementation strategy		
6	UC-FP1-WP3-35 – Special cases		
7	UC-FP1-WP3-36 – Generating plans through different inputs	CAF tool	CAF
8	UC-FP1-WP3-37 – Validation of planning		
9	UC-FP1-WP3-38 – Planification simulation and acceptance		
10	UC-FP1-WP3-39 – Planning changes based on data analytics		
11	UC-FP1-WP3-40 – System effects of different grades of automation	RailSys	TRV A.E. KTH
12	UC-FP1-WP3-41 – System effects of DATO concepts	RailSys	Prorail
13	UC-FP1-WP3-42 – Feedback loops between crew plan and operation	AnyCrew	NSR
14	UC-FP1-WP3-43 – Assess the feasibility of a change in the topology network.	TMS_Cap	Indra
15	UC-FP1-WP3-44 – Effects of C-DAS in capacity	TMS_Cap	Indra
16	UC-FP1-WP3-45 – Effects of introducing ETCS Hybrid Level 3 on lines with dense traffic	RailSys HESE	TRV A.E. KTH
17	UC-FP1-WP3-46 – Effects of C-DAS on energy consumption and capacity	Opticon	CEIT
18	UC-FP1-WP3-47 – Effects from varying adhesion conditions and introducing new generation braking system	RailSys HESE	TRV A.E. KTH

6.14 Conclusions of current simulation usage and development needs

There are several areas where simulation models are in need of improvements. The following development needs have been identified in this section:

- Integrate crew scheduling in simulations, and move towards doing integral simulation (including crew, rolling stock, infra, etc.)
- Simulations of large networks

- Improve the feedback loop between planning and operations, for example using historical data to improve timetables
- Developing a better and easier way to carry out stochastic simulation studies
- Implement different driving strategies depending on relevant factors such as current traffic situation, punctuality and energy consumption, for instance energy-efficient train driving by coasting
- Connecting energy optimisation to capacity analysis
- Possibility to evaluate new digital developments such as next generation brakes, ETCS hybrid level 3 and ATO/C-DAS over ETCS. This may require an improved TMS functionality where the driving strategy depends on the current traffic situation and bottlenecks. Also, an improved method for simulation of disrupted situations in ETCS may be required.
- Evaluation of ETCS level 2 roll-out strategy in terms of drivability, capacity and safety
- Improved evaluation of infrastructure layout and connect it to the timetable design
- Improved dispatcher functionality, for instance to maintain the order of trains according to the timetable and let trains run before schedule

From the mapping of the current simulation tools, it can be concluded that there are several tools available with developed functionality for simulation of a transport plan, including feedback loops, and for simulation of ETCS L2. However, to be able to simulate ETCS HL3, low adhesion effects (next generation brakes), C-DAS/ATO, driver behaviour, and TMS functionality including train path envelope developments are needed.

Developments are also needed for improved feedback loops, including crew scheduling and large networks. In cases where a requested function can be developed in several different tools, it is recommended to prioritise the development of the tools that are used by several partners to get synergies.

7. Research on the verification, calibration, and validation of railway simulation models

Simulation models require data to generate results. With this data it is possible to check if the model works as described by the specification, which is the process of verification. Furthermore, data is used to estimate the parameter settings and possible delay distributions for the simulation model, which is the process of calibration. After running different simulations, it is important to evaluate the results and to compare them with realization data or another simulation model or tool which is part of the validation process. This section considers research on verification, calibration and validation of railway simulation models in practice combined with literature research on those topics. The research aims to determine general guidelines for the verification, calibration and validation of simulation tooling.

The structure of this section is as follows. First, in Section 7.1 a general methodology is described for the verification, calibration and validation process based on (scientific) research. Afterwards, examples of the verification, calibration, and validation processes are shown in Section 7.2 for different European railway companies and research institutes (e.g., Lund University, KTH, and VTI, ProRail, NS, SNCF, and CEDEX). Finally, the section is closed with the conclusions in Section 7.3.

7.1 General methodology

This section describes a general methodology for the verification, calibration, and validation of simulation models. The focus is on simulation models for rail operations that evaluate the infrastructure capacity, timetable stability and/or timetable robustness. The method can be applied for macroscopic, mesoscopic and microscopic simulation models, which consider a different level of detail for infrastructure and interlocking modelling. It starts with the verification process, followed by the calibration and validation process. Based on several studies (e.g., Markewicz (2013), Watson and Medeossi (2014), Medeossi and De Fabris (2018), and Johansson, et al. (2022)) and discussions with experts it can that scientific research regarding verification, calibration, and validation for railway simulation models is limited. Furthermore, if these topics are considered in literature, the focus is mostly on one of them. Exceptions are Markewicz (2013) and Watson and Medeossi (2014), who consider the process of calibration and validation of railway simulation models. In addition, Roungas et al. (2018) developed a framework for the verification and validation of simulation models in general.

A general description for the processes of verification, validation, and calibration of simulation models can be found in Law (2015). Note that all models are simplifications of reality, so there will always be differences found during the calibration and validation processes. Therefore, it is important to have a clear goal for the application of a simulation model and then to determine the minimum quality of the simulation model output.

7.1.1 Verification

The first step considered in simulation modelling is verification, which is defined as whether the specification of the simulation tool is correctly translated into the software by programming (e.g., debugging) (Law, 2015). Most of the simulation models used in practice are developed by external companies where this process has taken place. Verification is then only applied in situations where new specifications defined by the end-user are implemented in the simulation model. During verification the simulation model is first applied applied to simple case studies (e.g., simple theoretical examples) to see if the results are in line with the expectations and/or specifications. If the simulation model behaves as expected, the simulation model can be applied to more complex case studies (e.g., real-life regional or national network) to verify if the model still behaves according to the expectations and/or specifications.

7.1.2 Calibration

After verification, the calibration of the simulation model can be performed. Calibration is defined as determining the parameter settings and delay distributions for a simulation model and checking if the output of the simulation model agrees with the actual, observed outcomes (e.g., realization data) (Law, 2015). Calibration should be applied when there is a significant change in the infrastructure, timetable and/or software (Markewicz, 2013).

Calibration of parameters can be focused on the train motional model that is used to compute the running times in a simulation model. The train motional model includes factors such as the train mass, traction and braking effort, coefficient of the train resistance equation, gradients, curves and tunnels. In addition, the loss of performance due to wear and tear could be calibrated (e.g., a train with a breakdown, old or with advanced engine or brake wear will not reflect the same theoretical values). A preliminary sensitivity analysis can be applied to investigate which parameters have the biggest influence on the model output, before starting the calibration process (Bešinović, et al., 2013). Cunnilera et al. (2023) provide a literature review on the calibration of train motional model using realization data. Calibration of the train motional model is challenging due to a limited availability of accurate infrastructure data describing track geometry, physical sources of train motional model parameter variations (e.g., wind, weather and temperature, available power at the catenary, adhesion, wear of the train and tracks), and driver behaviour. The calibration can be based on online (on-board) or offline (historical) realization data. Offline techniques are more suitable for railway simulation tools, because more data can be used to determine the train motional parameters. For example, NS is experimenting with an accelerometer to determine simulation parameters for ATO.

Delay distributions are only necessary for stochastic simulation (e.g., to evaluate timetable robustness), where random components are introduced for a better representation compared to deterministic simulation (Medeossi & De Fabris, 2018). Realization data is needed to calibrate the simulation model, where primary delays should be filtered from the data, especially for entry delay distributions. This might be a challenge in practice, because it depends on the quality of the realization data. In addition, it is not always clear in the realization data if a delay is caused by the train itself or by knock-on delays of other trains (especially the entry delays and running time delays). Advanced mathematical models can be used to analyse the interlocking data, if available. Otherwise, a rough filtering processes can be applied where first the delays bigger than 5 minutes are filtered and by using a mesoscopic model, the conflicts between arriving and departure trains can be identified at relevant stations (Medeossi & De Fabris, 2018). Another way to filter the primary and secondary delays is described by Palmqvist et al. (2023). They first filter big delays of more than 60 minutes from the realization data. Second, they scale down the empirical delay distributions to a fraction (F) that is considered primary, where this fraction (F) can vary for instance across entry, run and dwell delays, train types, distance, and time windows.

Calibration of dwell time and running time delays can be complex. For the calibration of dwell time delays it is important to consider that departure imprecision and dwell time variability influence the stop time variability (Medeossi & De Fabris, 2018). Departure time imprecision is caused by the fact that even trains that arrived on time do not necessarily depart on time, even if passengers boarded on-time. Dwell time variability is caused by boarding and alighting of passengers and the lay-out of the train (e.g., position and number of doors, distribution of passengers over the platform), but also by driver behaviour, dispatching, and occasional rolling stock failures (e.g., doors failing to close). Calibration of the running times is particularly challenging due to driver behaviour during the different driving regimes (e.g., acceleration, cruising, coasting and braking), especially for freight trains with pneumatic brakes. Additionally, many simulation tools only consider a limited number of performance parameters to be influenced.

For stochastic simulation, dispatching behaviour (e.g., order of trains during delays) consideration should also be given to timetable robustness and resilience. Dispatching measurements of the simulation must be calibrated based on the traffic control rules applied by the dispatcher, which are often more complicated and holistic than can be captured in a simple rule. This becomes even more difficult when major disruptions are considered, since automatic train cancellation and short-turning policies and the relationship with reallocation of rolling stock and crew need to be considered (Medeossi & De Fabris, 2018).

During the process of calibration for simulation models, the application of sensitivity analysis is important to see the correlation between different parameters as well as the effect of the distribution function on the simulation output for stochastic simulation. Applying a statistical experimental design can help determine the parameters that have the most significant impact on the results (Law, 2015), where also the control of randomness is considered.

The calibration process is ongoing until the simulation results approach the input data. However, a question remains about how close to get before the simulation model is considered calibrated, i.e., when are the results “good enough” (Markewicz, 2013; Watson & Medeossi, 2014). This depends on the aim of the study and requested accuracy of the model. The quality of the results can be evaluated using for instance the (mean) delay and (model) punctuality (Lindfeldt, 2015; Medeossi & De Fabris, 2018). For a concise overview of the literature on the calibration of railway simulation models, refer to Johansson et al. (2022).

7.1.3 Validation

After calibration it is important to validate the simulation model. Validation is defined as the process to check if the simulation models the system accurate enough given the objectives of the simulation study (Law, 2015). Validation should be applied when there is a significant change in the input of the simulation model, parameters and/or a change in the software. An example for a railway simulation model is a new infrastructure or new rolling stock that has a big influence on the outcome of the simulation results. Unfortunately, there is no complete, definitive general approach available for the validation of the simulation tool. However, there are multiple methods to validate the tool, e.g., comparison with realization data, comparison with another model and/or tool, expert judgement by analysing timetable graphs and diagrams (e.g., network (delay) graphs, time-distance, speed-distance, acceleration-distance, energy consumption, bar charts), and analysing the animation of trains running through the network (e.g., play the simulation run and see trains running graphically). This section focuses on two common methods for validation: comparison with realization data and comparison with another model. Furthermore, the simulation model should also be credible, which means that the management and/or the project team accepts the tool as correct (Law, 2015).

Validation of railway simulation models can be applied for instance on minimum running times, headway times, conflicts, (model) punctuality and mean delay (and standard deviation). Validation of the minimum running times is commonly applied, because it is relatively easy to get these times from the simulation and realization data. Headway times and conflicts are more difficult to determine, because they depend on the level of detail for infrastructure modelling in the simulation model and on the input data. Furthermore, it is difficult to determine headway times and conflicts in practice, because they are time and distance dependent. Validation of the (model) punctuality and mean delay (and standard deviation) can be applied to validate the general results of stochastic simulation model (Medeossi & De Fabris, 2018).

If realization data is available, validation should be applied based on these data. Since model calibration also requires realisation data, separate data for both processes are recommended to have a general representation of the simulation model instead of only on the input data (Law, 2015). However, Tiong et al. (2023) highlighted that little research considers separate datasets for calibration and validation due to overlooking its significance or misunderstanding the concept of validation with realization data. If no realization data is available, different simulation models and/or tools should be compared to see if the modelling of both tools leads to similar results. However, it is important to have in mind that even if both models/tools produce the same results it is not determined whether either model/tool is valid, because they could both have the same error.

A main question regarding validation is when to stop. While there is no general guideline for this, statistical test (e.g., goodness of fit) can be applied to define criteria when the results are “good enough” (Law, 2015; Tiong et al., 2023) or the extra effort for improvement can be investigated to determine the cost-benefit ratio (Lindfeldt, 2015). Practical criteria to conclude the validation process include insufficient project time/budget and expert judgment (e.g., main differences can be explained and are accepted given the purpose of the simulation model). Roungas et al. (2018) developed a framework for simulation model validation that helps experts to decide the validation method depending on the phase of a simulation study as well as the statistical techniques to apply. Tiong et al. (2023) developed a general framework for data-driven train delay prediction (simulation) model development, where a part is focussed on model evaluation including validation. They gave some statistical techniques to analyse the model accuracy and representational power.

7.2 Practical examples

This section discusses different practical examples of verification, calibration and validation of railway simulation models. The following examples are shown:

- Sweden:
  - Lund University gives an overview of operational simulation tools and calibration methods
  - KTH provides information about the validation of different tools for ETCS regarding running times, headway times and braking curves
  - VTI provides information about the validation of their train driver simulator
- The Netherlands:
  - ProRail describes the application of verification, calibration and validation on their simulation tools
  - NS describes their calibration method for stochastic microscopic simulation models and their validation method for minimum running times and conflicts using realization data
- France:
  - SNCF gives insight in their simulation model calibration and validation by comparing different simulation tools
- Spain:
  - CEDEX provides insights in their validation process of the ETCS simulation laboratory

7.2.1 Practical examples from TRV A.E. Lund

Brief overview of how railway operation simulation tools work

Roughly put, railway operation simulation tools such as RailSys and Proton are event-based simulations. They use a Monte Carlo-based approach to convert a set of inputs to a set of outputs. The key inputs are a network description, rolling stock properties, a timetable, and primary delay distributions. The main outputs are simulated departure and arrival times. From these, delays and punctuality can be derived. These, in turn, are analysed to draw some conclusions about the operations. The scope can be different infrastructure designs, timetable designs, or punctuality forecasts, to name a few, with the number of use cases growing over time.

In the simulation, trains enter the system according to the specified timetable, plus a random draw from an *entry delay distribution*, as in a typical Monte Carlo approach. It then runs along the network, using the scheduled run time for that section (which can be and often is divided into technical minimum run time plus margins in the shape of run time supplements), plus a random draw from a run time delay distribution. The handling of the margins varies from model to model, but essentially, if the draw from the *run time delay distribution* exceeds the allocated margin for a particular stretch, the train is delayed. If the train enters a section with a delay, it is typically allowed to use a portion of its margins to reduce the delay. When the train arrives at a station, the precise handling of dwell times also varies slightly from model to model, but in essence, it stays there for the scheduled dwell time, plus a random draw from a specified *dwell time delay distribution*. In principle it is possible to make a distinction between technical minimum dwell time and dwell time margins, but unlike the case for run times, this distinction does not typically exist in timetables. Some assumptions must then be made regarding the minimum dwell time, and how much can be used to reduce any pre-existing delays. Trains are also typically not allowed to depart a station with a scheduled stop before the scheduled departing time. This is often a good approximation for passenger trains, less so for freight trains, and the handling of the latter varies across models and simulation cases.

Many of the inputs into these models are in principle straightforward to estimate, verify and validate, even if it takes some work to do so. This is not the case for the input delay distributions. The entry, run, and dwell delay distributions specified as inputs should reflect the distributions of primary delays in a scenario. For realistic use cases, they should then reflect the empirical distributions of various primary delays. While large amounts of empirical delay data are often available, there is no clear way to distinguish in the empirical data which delays were primary or secondary. Secondary delays mean knock-on delays to trains as a consequence of an initial primary delay. Some delay attribution systems may have such fields, and attempts are made to link delays back to the primary cause, but these only apply to delay increases above a certain threshold. The threshold is typically three minutes, and some 80% of all delay time is caused by events below this threshold (sometimes called sub-threshold delays). For this reason, it is not known with any certainty what the real distributions of primary delays are. Instead, assumptions must be made and then the model can be calibrated so that the simulation outputs look like the real operational outputs.

Calibration of railway operations simulation models

In the context of railway operation simulation, calibration thus primarily refers to setting appropriate distributions for the primary entry, dwell, and run delays as inputs into the model. These distributions should be calibrated, often in an iterative way, such that the model outputs are in line with the actual, observed outcomes. Most often punctuality is used as the key metric, but some alternatives are possible. In principle, one could use arrival delay distributions, or even the output run and dwell time delay distributions, as the output against which to calibrate. Punctuality has the benefit of being a well-known key metric, and of being highly aggregated and interpretable, as well as often being what one fundamentally cares about. The delay distributions as such have the benefit of being more detailed and granular but are in the end less interesting for decision-makers, carry an increased risk of overfitting, and require much more work to calibrate satisfactorily. A model that is calibrated according to all these outputs simultaneously would be considered very well calibrated and validated to a particular case, but its generalisability for cases or years into the future may still be questionable. Arguably, the practice of targeting a punctuality distribution is thus defensible. It is less precise, but also easier to do, and has less risk of overfitting to the particulars of a case that may not be relevant for the scenario that is to be simulated.

Regardless of the choice of output variable to calibrate against, a fundamental challenge with the calibration of models is that both the models and the real processes are highly stochastic. There are no days without any disturbances, so even if no primary delays are entered into the model, there are no relevant cases with which to compare the outputs. It is also impossible to exactly recreate a given operational day in a simulation cycle, partly because there is no explicit way to tell for sure the extent to which an empirically observed delay is primary or secondary. The delay attribution systems that are used in industry typically only apply once a delay increase reaches a threshold of three or more minutes, smaller deviations are not coded. Often on the order of 80% of delay time comes from events which are below this threshold, and even above the threshold, studies have found an accuracy in reporting of about 80%. This data is thus highly imperfect and insufficient. Instead, Lund University has to work on distributions, ideally over longer periods of time and a large number of simulation cycles.

Current state of the art of calibrating models
Lund University believes that the current best practice for the calibration of railway operation simulation is described in Palmqvist, Johansson & Sipilä (2023). A similar procedure was used in Minbashi et al (2023) and Johansson et al (2022). This included a simultaneous calibration of entry, run and dwell delay distributions, with separate distributions for different train types as inputs and punctuality as the output variable.

The procedure can be briefly outlined as follows. First, empirical delay distributions are generated from operational data. These contain all deviations for the initial departure (entry), run, and dwell times, including positive, negative or zero-values. In the above experiments, Lund University has created distributions across the whole day (without time windows, although this is in principle possible), across the whole network (without distinctions between small, medium or large stations), with separate distributions for different train types (e.g., local, regional, high-speed, freight trains). To avoid extreme outliers, Lund University has omitted run and dwell time deviations of plus minus 60 minutes. Continuing with the empirical delay distributions, one then assumes that a certain fraction (F) of these delays is primary, and the rest secondary. This results in a scaled-down distribution of delays, where the probability of a delay of each size (P) is scaled down to F*P. For instance, if the assumption is made that 50% of delays are primary, and 50% secondary, and there is a 10% probability of a run time delay of 1 minute in the empirical distribution, then there is a 50% * 10% = 5% probability of a 1-minute run time delay in the scaled down distribution, used as the input primary delay distribution. This assumes that the fraction is constant across the entire distribution, and that the remainder of the probability mass is moved to the no-delay case of 0 minutes (rather than to a smaller or negative delay). This procedure is then repeated in steps, assuming different fractions (F) of primary delays, perhaps in intervals of 5%-points. They then end up with perhaps 20 (or more or less, as desired) levels of primary delay distributions.
The calibration proceeds by simulating the operations for a number of cycles. How many cycles are needed depends on the setup itself (network, timetable) and to a large extent on the included (delay) distributions and their characteristics, and which outcome measures are in focus and their degree of aggregation. The output data of the simulations are then compiled to calculate the overall punctuality across the trains and cycles in each given scenario. Lund University then ends up with a number of simulated punctuality figures. The one that is closest to the empirically observed punctuality for that case can then be identified, and that level used as a proxy for the fraction of delays that is primary. In this way, for different geographies and timetables, Lund University has so far arrived at estimates of between 25% and 35% of delays being primary, and 75% to 65% being secondary. Of course, the precise result will vary from case to case.

Next steps for calibrating railway operation simulation models
The next steps improving the calibration of railway operation simulation models are outlined below.

Regarding inputs, it is shown that there is a need to calibrate entry, run, and dwell time delays separately but simultaneously, so that they can be scaled to different levels. This is more challenging, because it greatly increases (cubed) the number of cases that need to be generated and simulated. Instead of simulating 20 cases, one would have to try 4,000, for instance. In practical terms, a design-of-experiments approach is necessary. This involves systematically sampling combinations of the three inputs, rather than attempting exhaustive trials. Also, these distributions should be allowed to vary somewhat across the network, across time windows, and across train types, which further increases the number of combinations to try. The next step is likely to allow the levels to vary across train types (i.e., high-speed, long-distance, regional, commuter, freight), and to postpone detailed calibration across parts of the network and time windows (i.e., night, early morning, morning peak, daytime between peaks, afternoon peak, evening after peak, late evening). To handle these later steps will require more sophisticated experimental designs, and a better understanding of how to avoid over-fitting.

Regarding the output variable against which the calibration is performed, the theoretical ideal may be to use the simulated run, dwell, and arrival time distributions. In the short term, this carries some risks of over-fitting, which needs to be better understood, and is quite a long way off from the current state of the art. A more feasible set of improvements would be to first (as now) target the overall punctuality figure, with a next step of targeting punctuality distributions across simulation cycles and train types. A step beyond that would be to target the arrival time distributions as such, rather than the more aggregated punctuality measures and distributions. Once that is achieved, it may be feasible to also begin calibrating the output run and dwell time distributions, alongside the arrival times. Such a process is likely to require hundreds or thousands of simulated cases, each with at least a few hundred simulation cycles which leads to the need for advanced computational resources. That in turn requires a high degree of automation of the process, and very extensive computation power, so achieving this would require a concerted effort, likely over several years.

7.2.2 Practical examples from TRV A.E. KTH

HESE is a tool intended for calculating technical headways in ETCS L2 setups, with the aim that it should be easy to vary and compare different marker board location setups and to change track parameters such as gradients and line speed profiles. The HESE tool is validated in three steps. Train minimum running time calculations, i.e., calculations of the train speed profiles are validated against RailSys. ETCS braking, intervention and warning curves are validated against ERA Braking Curves Simulation Tool (ERA, 2018). Technical headway calculations are verified against RailSys.

Figure 13 shows example of a train run on a 60 km long track with some gradients, in this case a relatively heavy Swedish diesel-powered freight train. This train does not have enough traction force with respect to its weight to increase or maintain its speed in some situations with positive gradients.

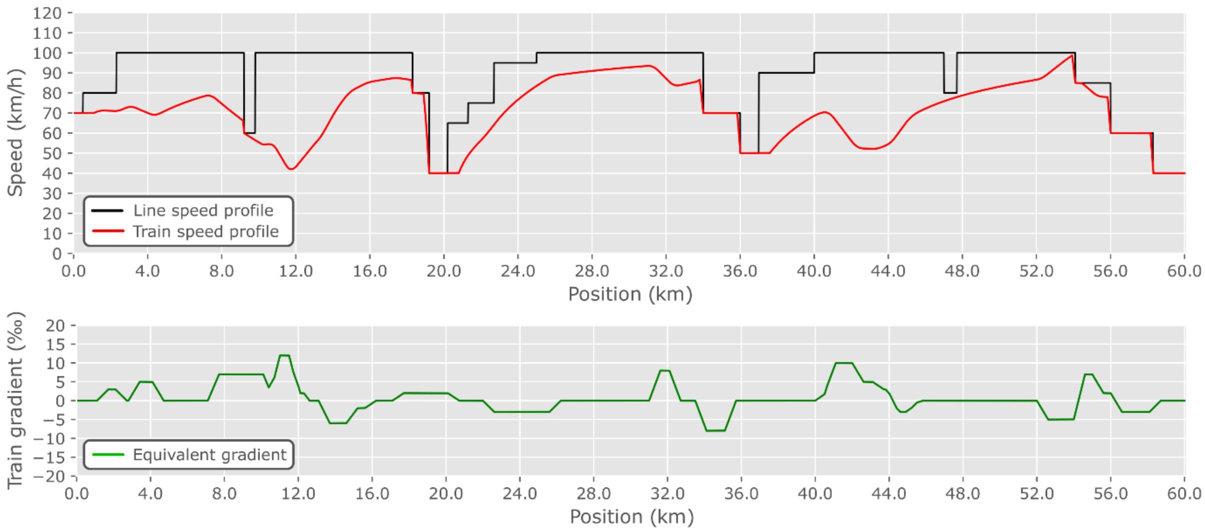


Figure 13: Example of train speed profile for a Swedish diesel powered 600 m long freight train with wagon mass 2000 ton and maximum train speed 100 km/h.

Figure 14 shows a similar example on the same track, but for the Swedish X2000 passenger train. Although this train has a maximum speed of 200 km/h, the maximum line speed sets a lower limit. Unlike the previous example, this train has sufficient traction force with respect to the gradients although the resulting accelerations naturally are slower than on level track.

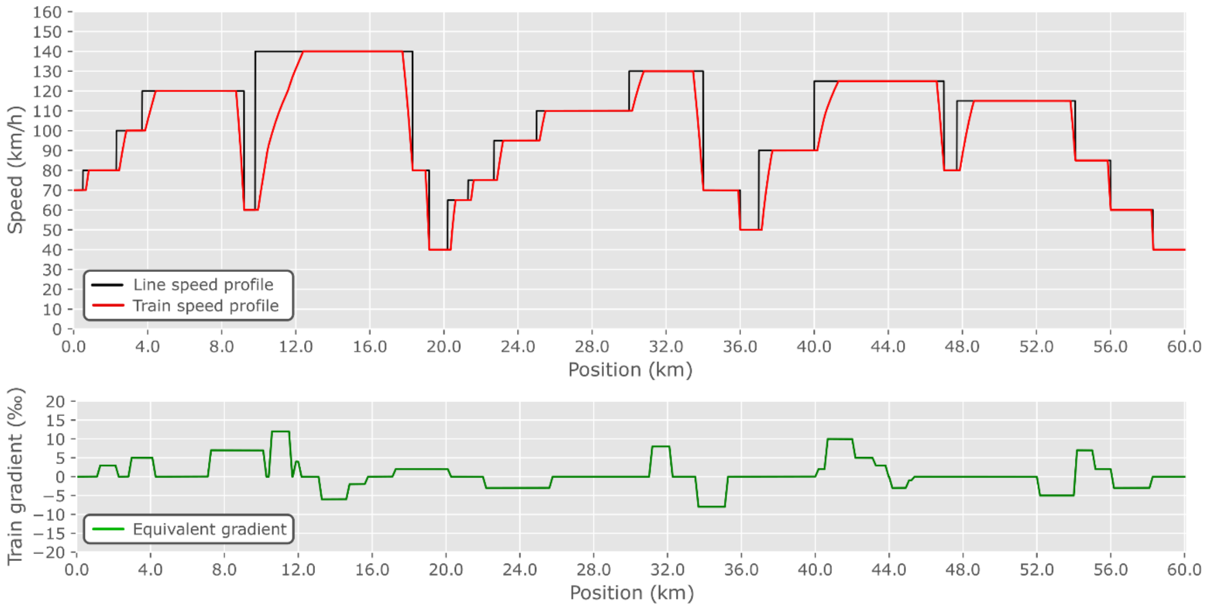


Figure 14: Example of train speed profile for the Swedish X2000 passenger train, 165 m, total mass 365 ton and with maximum train speed of 200 km/h but in this case track speed limits the maximum speed.

The same track section is used both in the train speed module in HESE tool and in RailSys for comparing the running times for a selection of different train types. The running time comparisons were first done using braking values for the Swedish ATC and not with ETCS L2, since the general train speed profile functionality was implemented first before the ETCS L2 braking. Examples of comparisons between the train running module in HESE and RailSys are given in **Fehler!**

Verweisquelle konnte nicht gefunden werden.. The marginal differences between the running times shown in table demonstrates that the train running module is modelling the train runs accurately.

Table 6. Examples of running time comparisons between HESE tool and RailSys using Swedish ATC

Train type ID	Description Loco/unit, wagon load, max train speed	HESE running time (s)	RailSys running time (s)
GR401610	Rc4 + 1600 tons, 100 km/h, Freight	2828	2832
GT422010	T44 + T44 + 2000 tons, 100 km/h, Freight	3220	3220
GEG01510	EG + 1500 tons, 100 km/h, Freight	2714	2718
GB202010	BR185/241 + 2000 tons, 100 km/h, Freight	2847	2850
PX600216	X60 + X60, 160 km/h, Passenger	2380	2383
PX2-2000	X2000, 200 km/h, Passenger	2397	2400
PR600909	Rc6 + 900 tons, 90 km/h, Passenger	2827	2829

Figure 15 visualizes the different curves in ETCS L2. Simplified, the curves are calculated from the EBD via EBI to SBI2, and from SBD to SBI1. Thereafter the resulting SBI is given by taking the most limiting of SBI1 and SBI2. Originating from the SBI, the W, P and I curves are established. The shape of the curves and the relationship between them is affected by train parameters (e.g., braking values), ETCS national values, and other parameters as well as track parameters. There is also a possibility to activate the use of guidance curves (GUI), this is included in the HESE tool but since RailSys 11 does not model these and the train data to HESE is imported from RailSys, there are some braking parameters missing which are required if modelling the GUI curves.

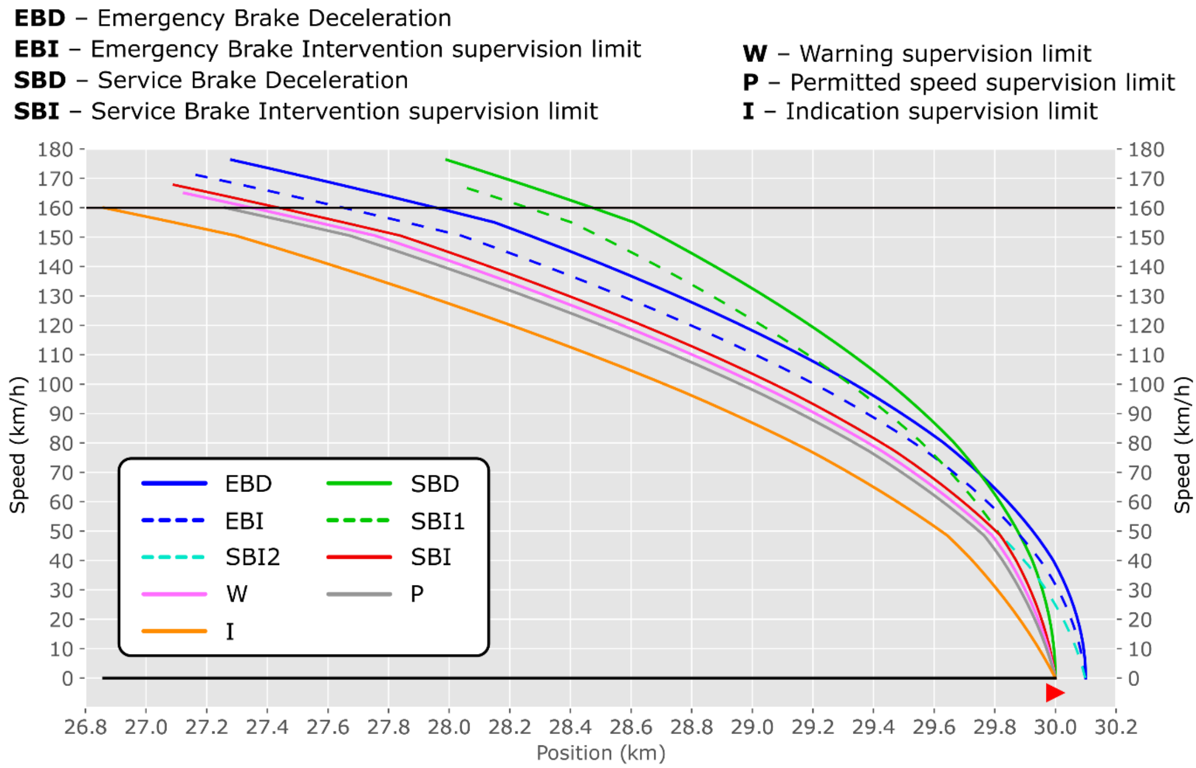


Figure 15: Braking, intervention, and warning curves in ETCS L2 with explanations. Example for an EOA target with supervised location 100 m beyond the target. SBI is defined as the most limiting of SBI1 and SBI2.

Figure 16 shows the principle of how technical headway is computed in the HESE tool with an example. Marker boards (signals) at red tringles, position 32, 34, 36, 38 and so on. Consider the situation where train 1 is occupying block section 38–40 and there is a train 2 following. When train 1 has cleared block section 38–40 with its full train length, including possible distance required for supervised location, a system time (release/setting time and transmission time) is required before train 2 can at earliest reach the point where it would have encountered a restriction if the movement authority (MA) was not extended from 38 to 40. Calculating a technical headway for position 38 (or section 38–40), can be described by:

- Calculating the running time for train 1 from 38 to 40 including the train length and possible supervised location distance
- System time: route release and route setting time, transmission time.
- Calculating the running time for train 2 from the position where it at latest must have an extended MA from 38 to 40 to avoid restriction and up to position 38.

Adding these times together gives the technical headway at location 38. In HESE tool, route release/setting times and transmission times are parameters that can be set in input data. Basically, if the focus is on comparing technical headway between different block section setups (I.e., the difference in technical headway), the sum of these time components can in principle be set to zero.

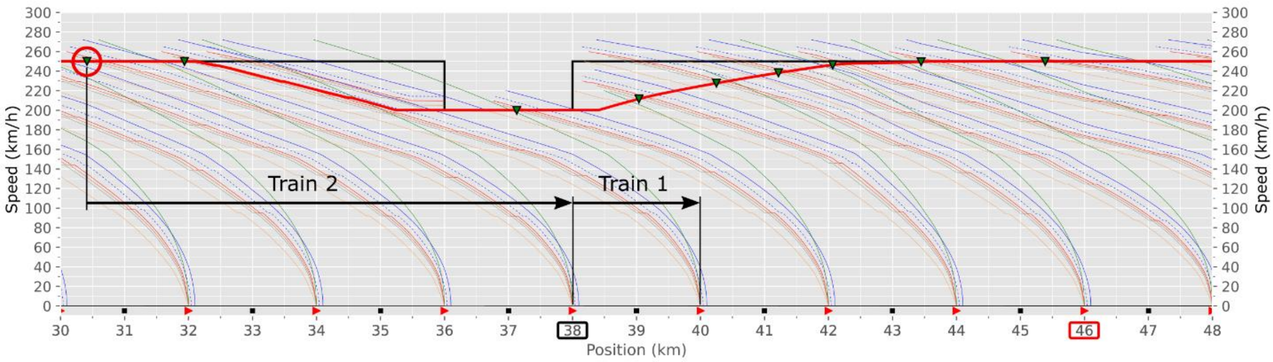


Figure 16. Example of technical headway. After train 1 has cleared section 38–40 and possible supervised location (overlap), train 2 can at earliest get an extension of its movement authority (MA) up to marker board 38. In this example marker boards are red triangles and location balises black squares

The modelling of the different ETCS curves (braking, intervention, warning) is verified by creating a set of cases in HESE and in ERA Braking Curve Simulation Tool, and comparing the curve distances from their respective target locations. Table 7-2 shows ERA and HESE tool curve distances for an End Of Authority (EOA) case, i.e., a case where the target speed is 0. In all four cases a supervised location (SvL) is set to 100 m behind the target. Distances from ERA and HESE tool are in principle identical for all cases and curves.

Table 7. Curve distances (m) with respect to target for ERA and HESE tool in an EOA case with target speed = 0 and SvL = 100 m.

	Gamma train from 200 km/h to EOA (SvL 100 m) with position inaccuracy Gradient = 0		Gamma train from 200 km/h to EOA (SvL 100 m) with speed and position inaccuracy Gradient = 0		Gamma train from 200 km/h to EOA (SvL 100 m) with speed inaccuracy Varying gradients		Gamma train from 200 km/h to EOA (SvL 100 m) with speed and position inaccuracy Varying gradients	
	ERA	**HESE**	**ERA**	**HESE**	**ERA**	**HESE**	**ERA**	**HESE**
EBD	3738	3738	3738	3738	4426	4426	4426	4426
SBD	2735	2735	2735	2735	3060	3060	3060	3060
EBI	5102	5100	5350	5348	4877	4878	6054	6053
SBI	5369	5366	5617	5615	5144	5145	6321	6319
W	5480	5477	5728	5726	5255	5256	6432	6430
P	5591	5589	5839	5837	5366	5367	6543	6542
I	6091	6089	6339	6337	5866	5867	7043	7042

Figure 15 shows distances for cases with a decrease in the line speed profile (most restrictive speed profile, MRSP). The SBD and SBI1 curves are not defined for these cases. Also, in these cases the distances are in principle identical. In addition to the examples shown in Table 7 and Table 8 more comparisons have been made for other train types and cases, also the calculations of GUI curves were checked. Distances, for relevant curves, are also compared with RailSys (but not shown here), these show good agreement with ERA.

Table 8. Curve distances (m) with respect to target for ERA and HESE tool in a case with target speed ≠ 0 (MRSP). The SBD curve is not defined for this type of target

	Gamma train from 270 to 240 km/h (MRSP) with speed and position inaccuracy Varying gradients		Gamma train from 240 to 100 km/h (MRSP) with speed and position inaccuracy Varying gradients		Lambda train from 200 to 160 km/h (MRSP) with speed and position inaccuracy Varying gradients		Lambda train from 200 to 70 km/h (MRSP) with speed and position inaccuracy Varying gradients	
	ERA	**HESE**	**ERA**	**HESE**	**ERA**	**HESE**	**ERA**	**HESE**
EBD	1407	1407	5665	5665	529	529	1397	1397
SBD	-	-	-	-	-	-	-	-
EBI	2219	2220	6284	6289	1040	1041	1902	1903
SBI	2579	2580	6604	6609	1487	1488	2349	2350
W	2729	2730	6738	6743	1598	1599	2460	2461
P	2879	2880	6871	6876	1709	1710	2571	2572
I	3554	3555	7471	7476	2289	2290	3151	3152

Calculation of technical headway is verified by designing a line track section and implementing this in both RailSys and HESE tool. Marker board locations and gradients are not varied but the line speed profile is different depending on train type. First a running time comparison is made for three different train types which follow different line speed profiles, in a similar way as in Table 6, but this time with ETCS L2 activated instead of the Swedish ATC (see Table 9).

Table 9. Examples of running time comparisons between HESE tool and RailSys using ETCS L2

Train type ID	**Description Loco/unit, wagon load, max train speed**	**HESE running time (s)**	**RailSys running time (s)**
ICE3	ICE3, 300 km/h, High speed passenger Gamma braking model	1014	1014
PX550220	X55 + X55, 200 km/h, Passenger Lambda braking model	1308	1307
GR401610	Rc4 + 1600 tons, 100 km/h, Freight Lambda braking model	2771	2771

The agreement in running times is practically identical for these train types. This indicates that the braking curves against targets where the line speed decreases have the same lengths with respect to the targets in both tools. After this check, comparisons of technical headway values are performed for the same three trains. This can be seen as two trains of the same type running one behind the other. If there are any major differences in how the two tools calculate braking curves with respect to ETCS L2 targets and/or how the technical headway is handled, it should appear at this stage.

Figure 18 shows values for technical minimum headway on this test line. The train types used are the same as in Table 9. The agreement is in general good for all three train types with only minor differences. Since RailSys version 11 does not model distance inaccuracy, this feature was naturally inactivated in the HESE tool as well. Tests were performed for additional cases in which some train and ETCS parameters were varied, these also showed good agreement in headway.

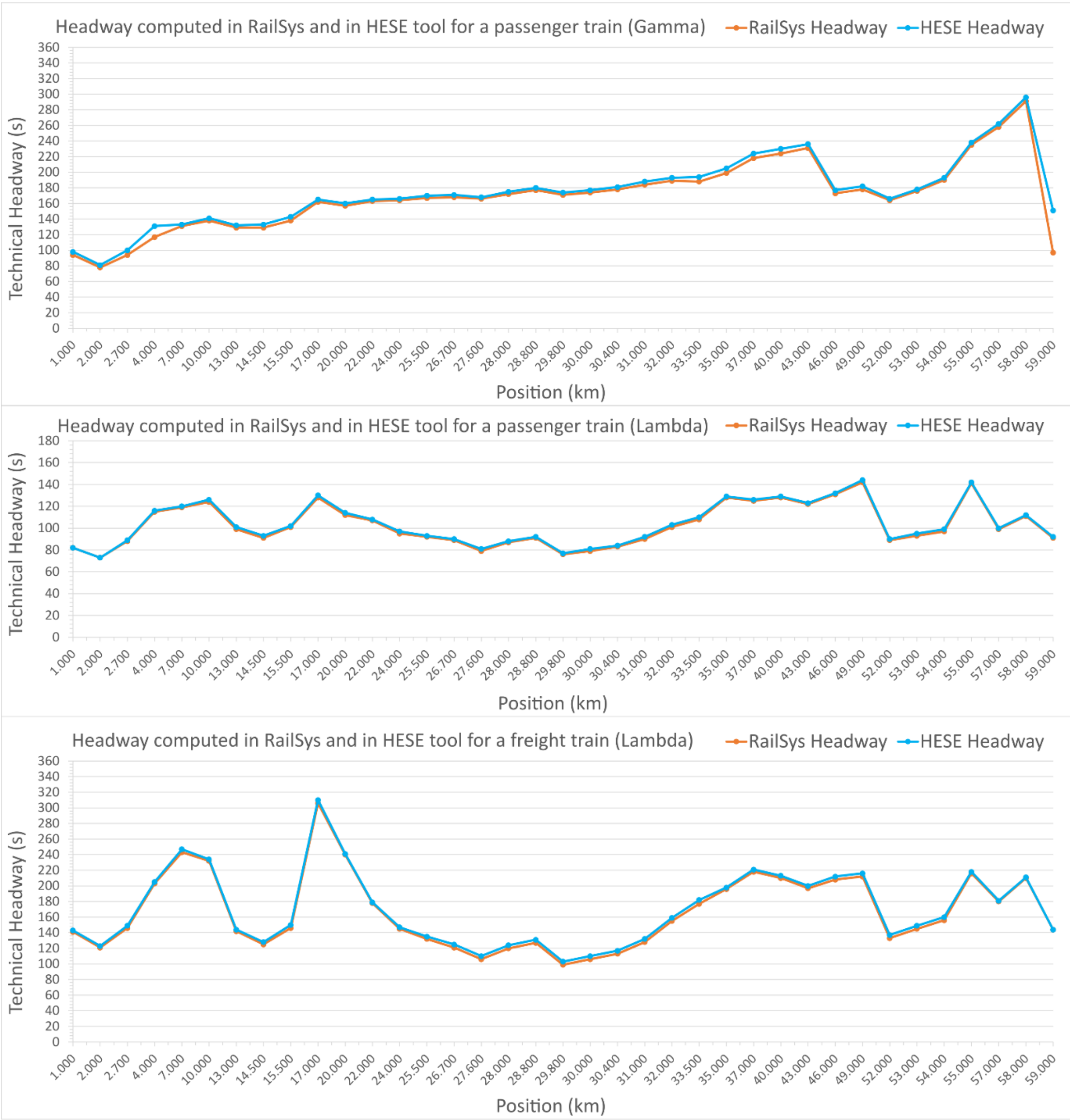


Figure 17: Comparisons of technical headway between RailSys and HESE tool for three different trains.

7.2.3 Practical examples from TRV A.E. VTI

The VTI train driving simulator is in consider in physical terms as low-fidelity (i.e., no moving base, a single screen and incomplete cab, see Figure 18), but with high functional fidelity. The levers used come from real trains, and the retardation and acceleration values inserted in the software correspond to real train driving, as do the values entered in the train protection system. The simulator includes two vehicle models, both of which are among Sweden's most common: the Bombardier locomotive TRAXX (freight) and Bombardier Regina EMU X50 (Passenger train). The tracks used in the simulator were built as real tracks based on infrastructure data collected from

BIS (the Transport Administration database), including coordinates, curvatures, signal positions, level-crossings, gradients and balise positions. To mimic the track surroundings, video films from the routes were used.

Figure 18: One hardware version of the VTI train driver simulator, used within the validation study.

Validation of the train driving simulator on the user level

A driving simulator offers a safe and useful environment for cost-effective, controlled and standardized tests of human behaviour. Due to the difficulty of controlling real traffic situations, there is a need to conduct research on train driver behaviour in a safe, simulated environment where the results are transferable to real train driving. Results that can be generalized to the real-world during an extended and uncontrolled period of time are referred to as ecologically valid.

To test the ecological validity of the VTI simulator, thirty-four Swedish train driver students from two different classes (2019-2021) in the final part of their basic education were assessed in a 45-minute simulator test using the number of driving errors as the performance indicator. The results were compared with the same students' performance at eleven weeks of internship as measured by supervisors grading according to a standard procedure. A significant correlation was found between the number of driving errors and internship grades as by analysing the data using Pearson product-moment correlation, $r = -.449$, $p < .05$ (see Figure 19), indicating a moderately strong correlation with the significance level set to 5%. As the 45 minutes of simulated driving cannot replicate the reality's level of detail nor measure all human behaviour important for actual train-driving (i.e., long-term attentiveness, carelessness or stress-tolerance), the moderately strong correlation should be considered a good result. This concludes that this type of simulator is well suited for measuring real train-driving performance. However, an interesting follow-up to this study would be to compare simulator and actual running-time which is possible as the simulator scenarios replicate actual railway routes in Sweden, including their speed limits, topography, and signal positions.

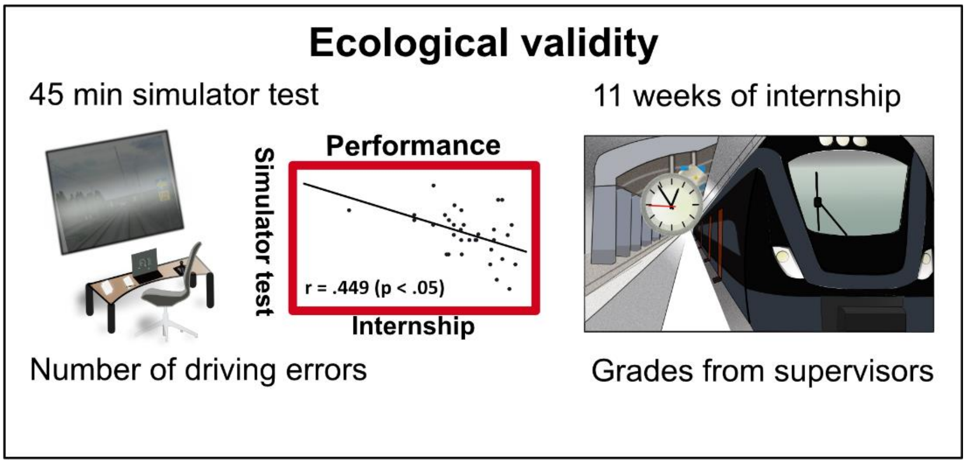


Figure 19: Graphical abstract from the validation study.

7.2.4 Practical examples from ProRail

In the past, several steps have been taken regarding verification, calibration, and validation of the timetable planning and simulation software. To start with the last two: there has been and still is more effort on the calibration of planning software than on simulation software. A few reasons that planning software compared to simulation software is checked more, are:

- The final timetable of the planning software is directly and on a daily basis linked to realisation data in analytical software "Sherlock", which makes analyses much easier compared to manual comparison of ad-hoc simulation studies.
- Planning software is used by many more users (factor ± 20).
- Planning software is used for a wider variety of train types, routes, timetables etc.

The lessons learned from the checks and adjustments in planning software are usually applied in the same way in simulation software. For the planning software this is done by an expert group of the largest train operators, senior planners, staff from the analytics bureau and software developers. If a major impact is expected, the changes are applied first in simulation software to analyse the effect on the yearly or daily timetable. Examples include the changes in the running time module regarding acceleration and deceleration and changes in the running time supplements.

Next, the verification, calibration and validation are considered. Verification is done during the software development. However, most simulation software is not developed in-house, so this is usually done by the external software developer. Later on, or after upgrades with new features/calculation methods, it is done on a smaller scale by end users of ProRail.
Calibration and validation are done based on realisation data or by cross checking. Since the realised timetable and operational data of all train traffic are stored in one database and a special team is available, this can be done on a structural basis but also focussed on changes. One method that has been developed is to create a "running time book" of planning data, simulation data and the 10^{th} percentile fastest running times from realisation, given the most logical route. By doing this for most part of the country, there has been a check on both planning and simulation. Another

way of validation was the cross check between different simulation software for situations that sometimes did not have realisation data yet, such as ERTMS on lines where that has not been implemented yet. By tweaking both simulation software in the comparison (i.e., RailSys, OpenTrack, and FRISO) a check has been done if the behaviour in both is of the same size and direction or that there were differences, that either could be clarified or not.

7.2.5 Practical examples from NS

Calibration

NS applies calibration for the project stochastic microscopic simulation (SMS). Calibration is defined as setting appropriate parameters and delay distributions using realisation data. The aims of this project are (1) to gain insights in the relationship between norm times and the operational performance, (2) to gain experience within NS with stochastic microscopic simulation, and (3) to compare (verification) different simulation tools (RailSys and FRISO). For FRISO NS considers the complete Dutch railway network, whereas RailSys is only applied on a partial network (due to too much manual work). The main simulation settings (e.g., number of simulations and driving behaviour) are determined during meetings with realization data and simulation experts.

The delay distributions are calibrated by filtering the realisation data (i.e., only delays smaller than 5 minutes are included to avoid influence of traffic control). It is also important to filter primary delays (not caused by other trains) from this realization data. ProRail and NS developed a methodology called STIPT to automatically derive the cause of the delay in realization data using multiple data sources. One these sources is the signal data that is used to determine secondary (knock-on) delays. The system checks if the signal in front of a train was hindering this train, which was caused by another train that occupies the block section on the infrastructure. NS used the method STIPT to filter the secondary delays out of the realization data, so they could use the primary delays to derive the delay distributions.

NS considers the following delay distributions in their models: entrance (i.e., entering the model), dwell time (i.e., process of boarding and alighting of passengers), departure time (i.e., process after departure signal turned on and before driving the train), and running time (i.e., process of running the train). Entrance delays are determined at the borders of their railway network using realisation data. For the dwell time, NS clusters the data based on different station categories (in total 6 station types) by sorting the stations according to intervals of the median of the dwell time. In addition, the data was clustered on different train classes (i.e., Intercity and Sprinter), so in total there are 11 different classes. Departure time delays are only considered at bigger stations. Finally, running time delays are only applied for FRISO, where a delay distribution is generated for the arrival time at the next station by an energy-efficient train driving strategy based on coasting. After running multiple stochastic simulation runs, the simulation results are compared with realization data (validation). Multiple iterations are performed to adjust the delay distributions and parameter settings of the simulation.

Validation

NS develops a methodology for the validation of timetabling and simulation tools. They define validation as comparing the output of a tool with realization data. Validation is necessary due to the more accurate way of timetabling and simulation at NS (i.e., blocking times and planning in

1/10 minute) and due to the limited available capacity on the Dutch railway network (high-capacity utilization). Validation needs to be applied when there is a change in input (e.g., rolling stock or infrastructure data) or the tool itself (e.g., parameter settings), and if realization data is available for the validation.

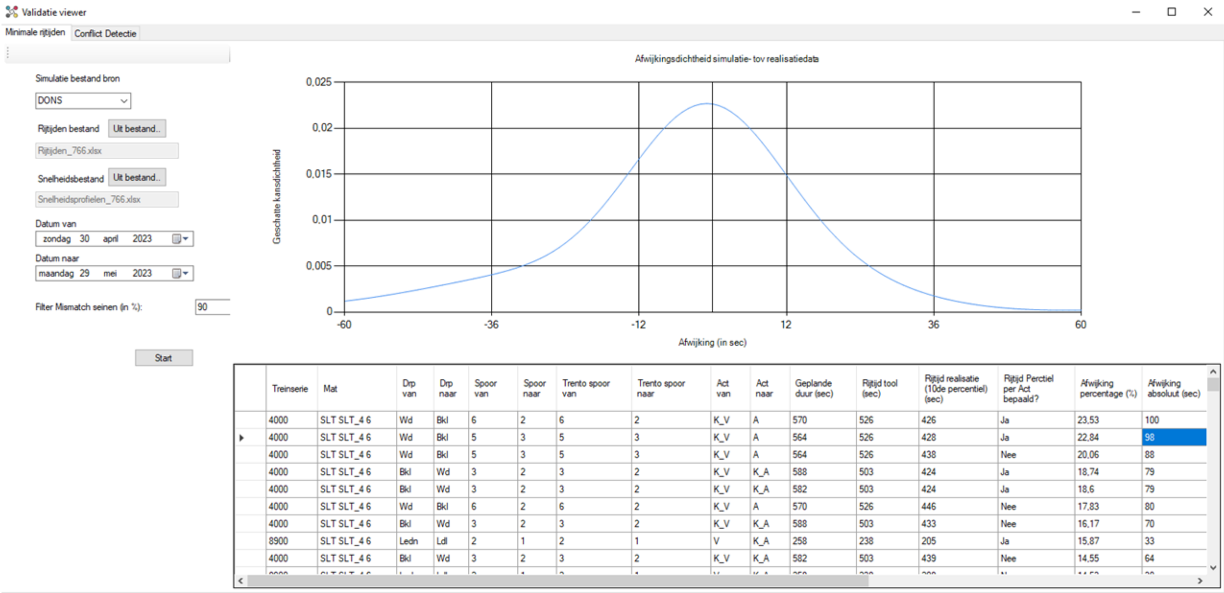
NS develops and tests a methodology for validation on some case studies for both timetabling tool DONS (developed for ProRail and NS) and simulation tools OpenTrack and FRISO, which are all based on microscopic conflict detection (blocking times). The research is conducted in collaboration with ProRail, because they have realization data and tooling (Sherlock) to analyse the data, and to compare it with the results of the tools.

Research regarding validation started in 2016 at ProRail and NS, where a first step for a methodology was developed in 2017. The focus was on validation of minimum running times and headway times. However, during that period the accuracy of the timetabling tools changed from macroscopic (headway norms) to microscopic (blocking times). Therefore, in 2018 NS initiated a new project to develop a basic method for validation of minimum running times, headway times and conflicts. First, minimum running time is defined as the fastest possible running time between two stops by maximum acceleration, cruising at the speed limit and maximum braking. Second, headway times are defined as the time interval between two successive trains (measured at the head of the trains). Third, conflicts are defined as two trains that would like to claim the same infrastructure at the same moment (both physical occupation and reservation). NS selected these three because running times and conflict detection are the core part of timetabling and simulation tools. This led to a first version of the methodology in 2019. The main result was that the method for validation of the minimum running times was available, but the workflow should be more automized and the method should be improved to deal for instance with long distances between stops. Second, a first step in the development of validation of conflicts was delivered, but it should be improved to better match the block sections of conflicts in realization data and to automatize the process. Third, it can be concluded that the validation of headway times is not of added value, since NS already validates the conflicts (which are related to headway times).

Therefore, in 2022 a follow-up project was started to further develop and automatize the method for the minimum running times and conflicts. This project is still ongoing in 2023. One of the questions NS also would like to answer is when a tool is validated (when is good 'good enough'). Some ideas NS has based on the project SMS: P20 (20^{th} percentile), P50 and P80 of the tool output are in line with realization data or the biggest differences can be explained and are accepted by the realisation data and simulation experts. NS considers the following steps for validation of the minimum running times:

- Compare the minimum running time of the tool with 10^{th} percentile of realization data (i.e,. the 10 percent fastest running trains). NS filtered the realization data so that only unhindered trains in their data (only primary delays) are considered, and they filtered the data based on the track usage and rolling stock configuration. Then NS analysed histograms with the difference between the tool and realization. An example is shown in Figure 7-8.
- Face validation by using GPS plots to understand driver behaviour (acceleration, cruising, coasting, braking), and checking the maximum line speed and stopping locations of trains, where the minimum running time of the tool is compared with realization data. Unfortunately, most part of this analysis needs to be done manually. Furthermore, not all trains do provide GPS data and the data is not filtered. Therefore, face validation is only applied if NS would analyse the behaviour of

a specific train line based on the analysis of the percentiles of running times. An example can be found in Figure 22.



Treinserie	Mat	Drp van	Drp naar	Spoor van	Spoor naar	Trento spoor van	Trento spoor naar	Act van	Act naar	Geplande duur (sec)	Rijtijd tool (sec)	Rijtijd realisatie (10de percentiel) (sec)	Rijtijd Perctiel per Act bepaald?	Afwijking percentage (%)	Afwijking absoluut (sec)
4000	SLT SLT_4 6	Wd	Bkl	6	2	6	2	K_V	A	570	526	426	Ja	23,53	100
4000	SLT SLT_4 6	Wd	Bkl	5	3	5	3	K_V	A	564	526	428	Ja	22,84	98
4000	SLT SLT_4 6	Wd	Bkl	5	3	5	3	K_V	A	564	526	438	Nee	20,06	88
4000	SLT SLT_4 6	Bkl	Wd	3	2	3	2	K_V	K_A	588	503	424	Ja	18,74	79
4000	SLT SLT_4 6	Bkl	Wd	3	2	3	2	K_V	K_A	582	503	424	Ja	18,6	79
4000	SLT SLT_4 6	Wd	Bkl	6	2	6	2	K_V	A	570	526	446	Nee	17,83	80
4000	SLT SLT_4 6	Bkl	Wd	3	2	3	2	K_V	K_A	588	503	433	Nee	16,17	70
8900	SLT SLT_4 6	Ledn	Ldl	2	1	2	1	V	K_A	258	238	205	Ja	15,87	33
4000	SLT SLT_4 6	Bkl	Wd	3	2	3	2	K_V	K_A	582	503	439	Nee	14,55	64

Figure 20: Screenshot from the tool Sherlock of a histogram to analyse difference in the minimum running time between the tool and the realization data.

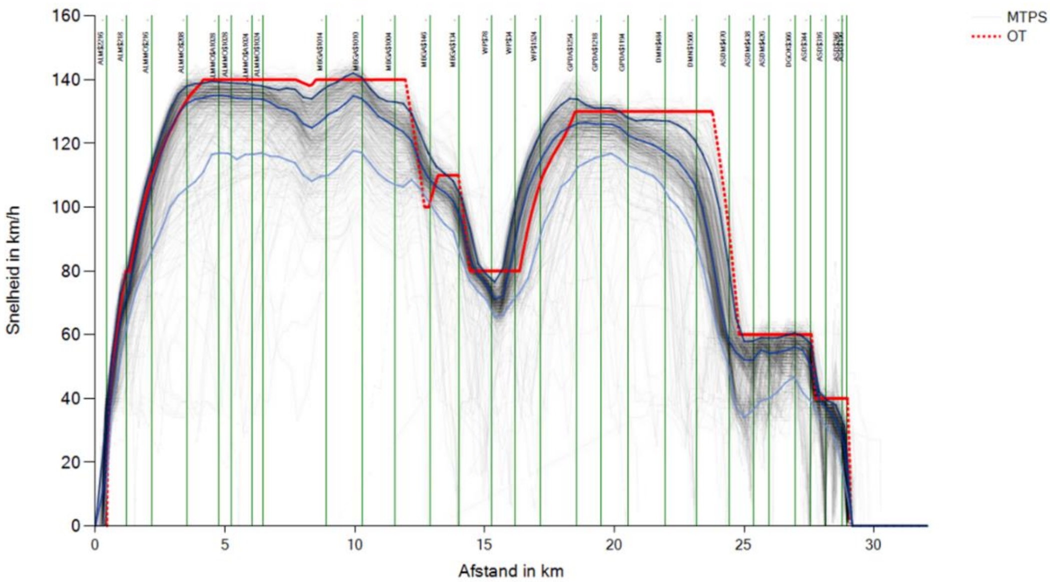


Figure 21: Example of a speed profile using GPS data to validate the minimum running times of OpenTrack (OT) with realization data (the dotted lines and thick lines the 10th, 50th and 90th percentile). The thick red line is the scheduled speed profile according to the timetabling tool Donna

The method for validation of conflicts is still under development. Basically, NS considers a microscopic level of detail for conflicts (blocking times) and for realization data they check the hindrance at a signal level by analysing the headway times. For the tool NS considers conflicts as block overlap (timetabling) or unscheduled red signal approaches (simulation). Figure 7-10 gives an example from the tool Sherlock, where the tables are used to determine if a conflict in the tool was also present in the realization data, and the graph indicates for the conflict the headway time in realization data with the frequency.

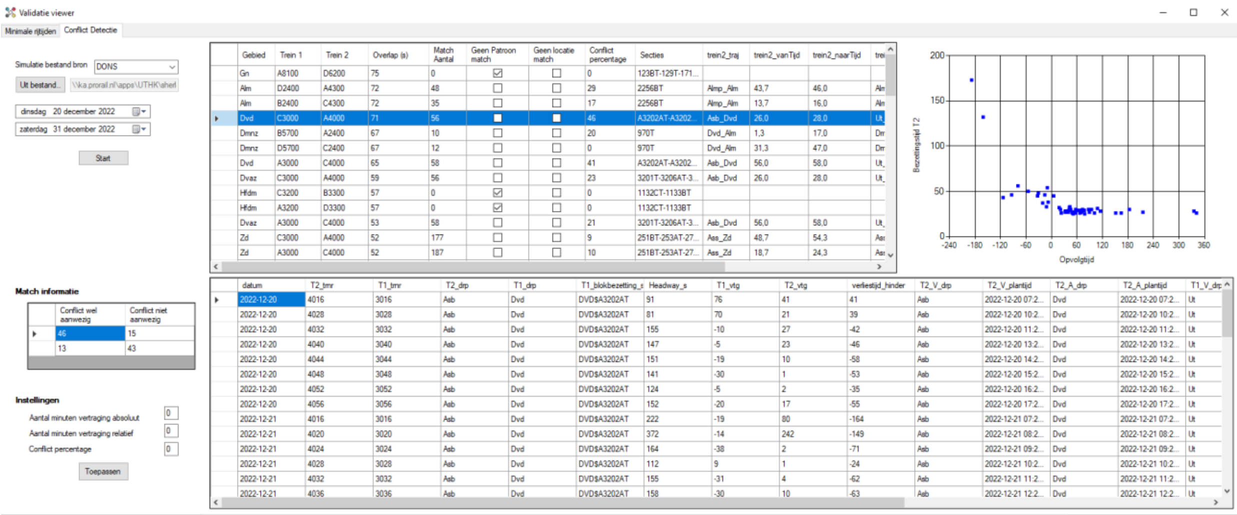


Figure 22: Screenshot from the tool Sherlock for the validation of conflicts. The table gives background information about the conflict match and the graph indicates the frequency of the headway time of the conflict in realization data.

7.2.6 Practical examples from SNCF

Calibration of the simulation models at SNCF Réseau is carried out in two different ways. First, in the case of static models, the data is calibrated using static operating reports available for the entire French rail network. These reports, known as "norms of operation", list the different timetable planning norms used by the timetable operators to develop the timetables according to the lines, configurations and rolling stock (see Figure 23). These timetable planning rules are set empirically over the years. Furthermore, SIPH (official tool for timetabling, TPS developed by HaCon) is used to calibrate conflicts. The conflict detection of SIPH allows a more reliable and precise calibration of the simulation data, compared to the old norms. The conflict detection is set according to the signalling and interlocking settings (microscopic conflict detection). The norms are defined for the most disadvantaged routes. Unfortunately, this process requires manual verification of conflicts by timetable operators.

- **Espacement des sillons de même sens**

Vitesse sillon	Parcours	Voie 1	Voie 2
≥ 120	Plaisir-Grignon – Dreux (BAL)	3'	3'
100		4'	4'
90 - 80		5'	5'

- **Intervalles minimaux entre tracés incompatibles**

	1er train / particularités	Valeur minimale	2ème train
Plaisir - Grignon	BV VC banalisée (250m)	garage 4' devant, sortie 2' 30" après	train voie direct
	Bif de Plaisir	itinéraires divergents, convergents, sécants : 4'	train voie direct
Dreux	BV Voie 4 banalisée (310m)	garage 4' devant, sortie 2' 30" après	train voie direct
	Gge impair 8 voies		

Figure 23: Example of norms of operation for French Rail Network.

Second, for dynamic models, the calibration of the data is essentially based on the actual delays (realization data) ("Bréhat" or "ORE" data). In this case, the Denfert (RailSys developed by RMCon) parameterisation is adjusted in order to reproduce, on the basis of the selected origin incidents, the same regularity data on the simulated timetable. Denfert parameterisation includes adding delays, routing rules, priority between trains, simulation parameters, allowance in train path, and allowance for lateness reduction.
For stochastics simulations, the delay distributions are calibrated by filtering the realisation data. SNCF includes the following data into the model:

- Only the delays smaller than 7 minutes, or.
- The delays that represent less than 80% of the realization data.

These values may vary depending on the study.

SNCF considers the following delay distributions in their models:

- Entry: i.e., delay of the train before entering the model.
- Dwell time: i.e., delay during boarding and alighting of passengers.
- Departure time: i.e., delay before the departure of the train at the first station - delay in preparing the train.

SNCF does not consider running time delay distribution, because they find this distribution difficult to implement and manage with the current simulation models.

SNCF also run unit tests (deterministic simulation): add a single delay of more than 10 minutes into the timetable on a selected train and analyse the consequences for other trains (knock-on delays). During this test SNCF discusses with the operators and with local correspondents where to introduce the delay (critical point of the infrastructure) and what regulation rules to apply.

The validation of the data used in SNCF Réseau's simulation models is limited by the fact that the

Denfert simulator (RailSys) is mainly used for upstream studies, i.e., design of long-term projects 5 to 15 years before being put into service. Therefore, there is little feedback on whether the simulation parameters match the reality of operations.

The main validation of models used in simulation tools such as "Denfert", consists of comparing the minimum running time calculated in the simulation tool with the minimum running time calculated in SIPH. It is only possible to compare the running time for the current infrastructure. The model is validated if the minimum run time calculated by Denfert is the same than the one in SIPH, or if SNCF can explain the differences.

For validation, SNCF also checks if there is any incoherence on the train run (speed limit, signals aspect, gradient, etc.). This is an empirical validation (see Figure 24).

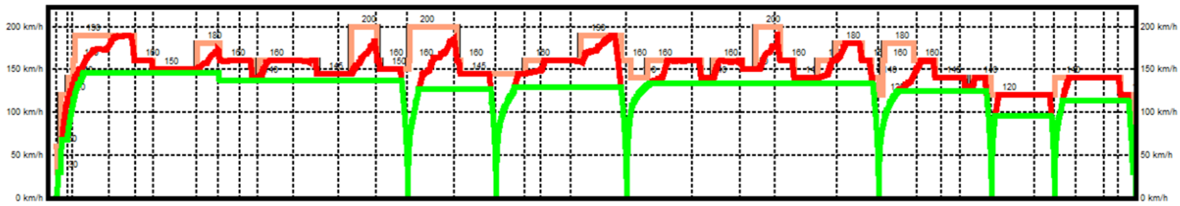


Figure 24: Example of the visualisation of the train run in Denfert. With this visualisation, SNCF can check if there is any incoherence on the train run.

The timetables plotted in Denfert are not directly used in operations: the SIPH tool produces the real timetables, and these are used in operations. However, the timetables simulated in Denfert are generally compatible with those plotted in SIPH, in particular because the fixed margins added allow for flexibility in driving behaviour.

7.2.7 Practical examples from ADIF A.E. CEDEX

As already mentioned in Section 6.9, the ERTMS Traffic Simulation Laboratory at CEDEX is a powerful tool where industrial ETCS components can be connected to the test environment and the participation of drivers and dispatchers in the simulation is perfectly possible. In this way, the CEDEX test environment can be used to compare the outcome of pure simulation environments for capacity analysis and assess the validity of the new algorithms to be implemented in the frame of FP1 WP8 and WP9. It could serve as an intermediate validation step prior to the comparison with data obtained from real train operations.

At this point, it is very important to stress that the accurate acquisition of data from real train operations can be, at least, complex. However, the ERTMS Traffic Simulation Laboratory at CEDEX, offers a lot of advantages in this regard, including the possibility to increase the data logging in order to cope with the validation needs.

Regarding the validation of the results obtained in the ERTMS Traffic Simulation Laboratory, the methodology employed have been the comparison with real operations. Besides, the validation exercise at CEDEX laboratory was focused on the ETCS performance. This validation target influences the next steps in the validation process design. In the end, the following requirements were defined:

1. Ensure that the ETCS equipment on track and in the laboratory are the same: the same software version and the same configuration/data preparation.
2. Design suitable and significant scenarios for the comparison. The definition must include any action affecting the dynamic scenario: driver actions, train dynamics, and dispatcher action. The scenarios must be shared with the infrastructure manager and the train operator to ensure its viability from a formal/safety point of view.
3. Ensure the capability to log the data for the scenario's comparison, especially on the track/train. On the train, the use of the train recorder (with the harmonized ETCS juridical data) is very helpful. On track, it is important to recover interlockings and Radio Block Center's logs, with the added difficulty of the proprietary format of this data, if possible.

As can be observed, the validation exercise is not an easy one. Actually, it is recommended to perform in the first stage the scenarios on the track, since there can be last minute modifications to the scenario's design, due to any incidence on the track, or the train. These modifications can be afterwards incorporated into the laboratory scenarios execution.

The experience at CEDEX with these exercises is that, for ERTMS Level 1 lines, the resemblance of the ETCS performance in the laboratory and on the track is very good. Figure 7-13 proofs this for a case study between Valdemoro and Villarrubias on a high-speed line in Spain where the permitted speed calculated by the ETCS OBU in track and in laboratory is shown. A description of the case study can be found in Figure 27.

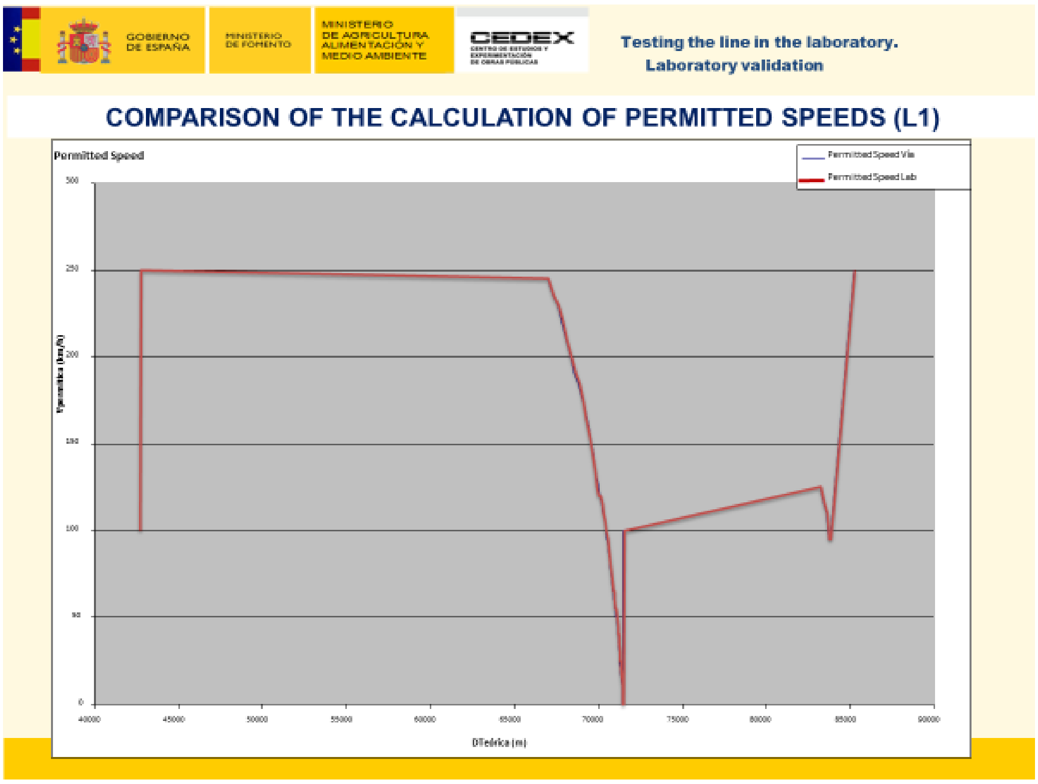


Figure 25: Comparison of Permitted speeds on the section Valdemoro-Villarrubias of the HSL Madrid-Valencia equipped with ERTMS Level 1.

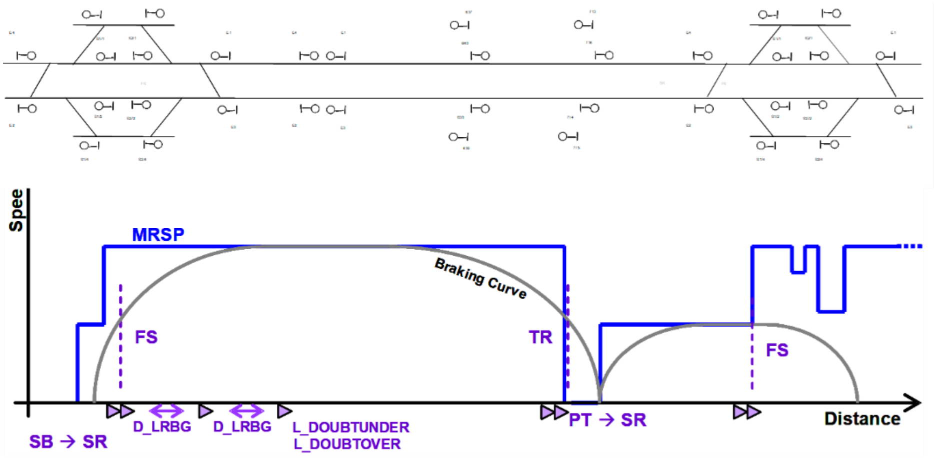


Figure 26: Description of the scenario on the section Valdemoro-Villarrubias of the HSL Madrid-Valencia equipped with ERTMS Level 1.

However, in ERTMS Level 2, the experience is more complex. In this case, the comparison exercise was done in the Madrid Commuter lines, around the Atocha station, see Figure 7-14. Although the comparison of permitted speed calculated onboard was very good, other aspects of the simulation presented differences. For instance, the ETCS messages exchange between the track and the train showed differences in time and distance. The main reasons were due to the real GSM-R network performance in the Atocha station and the impossibility at that time to register properly the signalman actions.

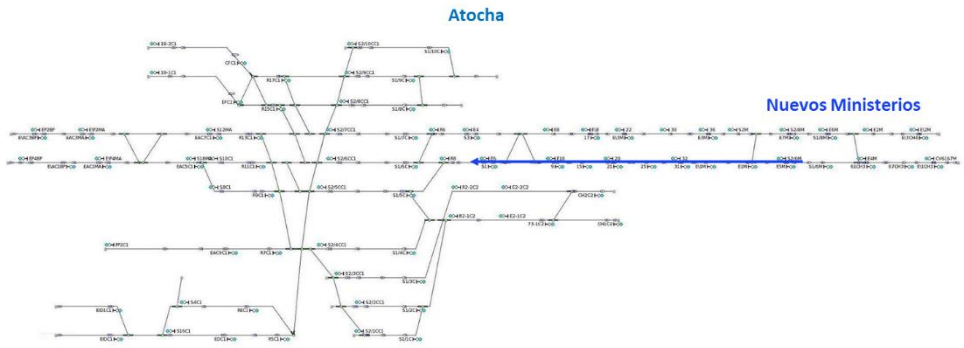


Figure 27: Description of the scenario on the section Nuevos Ministerios-Atocha in the Madrid commuter line equipped with ERTMS Level 2.

As a summary, the learned lessons from the validation exercises are:

1. Clearly define the scope/target of the validation.
2. Define a suitable process considering the scope of the validation, the different alternatives and its viability.
3. Deep analysis of the results, considering that they might affect previous steps of the simulation design (development, verification and calibration).

7.3 Conclusions of verification, calibration, and validation of railway simulation models

In this section the processes of verification, calibration and validation of simulation models applied for infrastructure capacity analyses and timetable evaluation (i.e., stability and robustness) were discussed. Verification is defined as the process where the output of the simulation model is compared with the specifications to build or improve the model. This mostly concerns debugging of the code of the software. During the process of calibration, the parameter settings and delay distributions for a simulation model are determined by comparing the simulation results with realization data. Validation is defined as the process to check if the simulation output is accurate enough given the objectives of the simulation study. If realization data is available, validation should be applied with them. Otherwise, different simulation models can be compared. The section also gave different examples from European railway companies and research institutes regarding verification, calibration and validation of simulation models.

Although the processes of verification, calibration and validation can be time consuming, costly and complex, they are very important for the quality of the results of the simulation model, because the quality is determined by the input data and the parameter settings. Limited scientific research is found that combines the processes of verification, calibration and validation of railway simulation models (Markewicz (2013), Watson and Medeossi (2014), Medeossi and De Fabris (2018), and Johansson, et al. (2022)). Current focus is mostly on one of the topics. Therefore, future research should integrate verification, calibration and validation with general guidelines, including examples of application.

8. Research on feedback loops between operation and timetable planning for improved capacity and punctuality

In order to improve the quality in railway traffic it is important to use feedback loops to exchange knowledge from one area to another. Feedback loops can be made in several levels and between several areas, as is shown in Figure 8-1. For example, planners do not have complete knowledge of how the traffic is operated and how the trains will perform in reality when making the timetable. However, with the use of real-world performance data, it is possible to analyse the data and connect the outcome to decisions made by the planners and make improvements. Instead of analysing real-world performance data it is also possible to simulate the traffic in a simulation tool and analyse the simulated outcome. It is especially useful if we want to test several hypothetical timetables, compare their qualities and give input to the timetable planning.

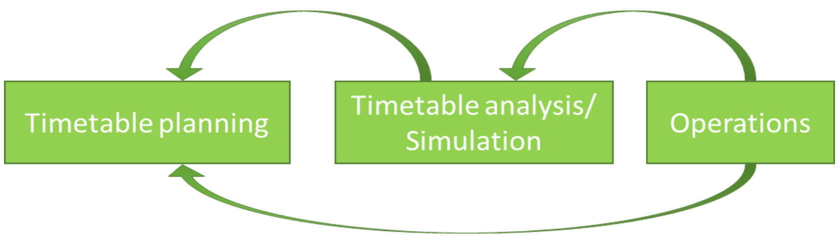


Figure 28: Feedback loops between planning, operations and timetable analysis/simulation.

Examples of different existing feedback loops can be given by ProRail in the Netherlands, some more formal and more structural, others more informal:

- The most formal one is the Performance Analysis Bureau (PAB) within ProRail. They daily get data of all train detection systems, train measuring systems, signalling systems, traffic control, and planning systems. Based on these large data sets they give feedback to the respective other departments, such as Asset Management on the setting of timers, timetabling on running times and Traffic Control on route setting strategies. They provide the information during design workshops or otherwise directly to the people that are responsible to implement the changes.
- Another large group providing feedback (but in a less structural way) are the train drivers and traffic controllers who can send their feedback from operations directly or indirectly via their managers. This feedback can go different ways to get to the responsible department. A route via PAB might make it faster, otherwise usually a check with the PAB is carried out, before implementing or declining the suggestions.
- Another smaller group providing less structural feedback loop are the specialists from other different departments carrying out inspections (e.g., visibility of signs, realized stopping positions) measurements (e.g., timers, running times) during train operations and implement their feedback sometimes directly themselves in their systems.
- Like train operators within ProRail also calculations are made on passenger flows based on passenger smartcard and ticket data of the TOCs. This information is fed back into the long-term timetable planning and capacity analyses for future infrastructure.
- Feedback can also go before the practice outside. Simulations are run within the timetabling department based on realisation data with new timetables to judge the feasibility and robustness. The feedback is given back to the planners.

Within this project, we will give and overview of what has been done in Europe regarding feedback loop methodologies, both in research and in practice. The focus is on the feedback loop between operations and timetable planning, but we also assess the important feedback loops between traffic simulation and timetable planning and between operations and simulation. Timetable analysis and simulation can give useful outputs, as a complement to operational performance, to improve the planning and feedback from operations to simulation is needed for improving methods and models.

In the following three sections (sections 8.1, 8.2 and 8.3) each of the three feedback loops are assessed. They are followed by a section with development needs within feedback loops (section 8.4). Finally, a summary section is included with the main conclusions extracted from the feedback loop topic (section 8.5).

8.1 Feedback loops between operations and timetable planning

One major purpose of the feedback loops from operations to timetable planning is to improve timetable quality. At TRV, several performance analyses are made yearly of railway lines, services and specific locations with the aim to decrease delays. For example, Trafikverket (2016) and Trafikverket (2018) describe two analyses where operational performance is analysed to develop new timetable planning rules. Solinen and Palmqvist (2023) describe the process of implementing new timetable planning rules and how operational performance is used to evaluate the timetables. In Sweden, Lund University has worked extensively around creating, investigating and improving various types of feedback loops between operations and planning. Three areas we can highlight are punctuality, timetabling, and dwell times.

Palmqvist (2019) summarised and synthesised several analyses regarding train delays and timetabling. There, the difficulty that timetable planners face in realising the feedback loops from their planning is highlighted. The making of the plan often takes place more than a year in advance, and it often covers thousands of train paths per planner, with several variants for each path over the year. It is clearly a challenge to learn from these stochastic outputs, and take these lessons into how to precisely fine-tune the size and location of dwell times and run time supplements. Figure 8-2 summarises this work, and illustrates both "single-loop" and "double-loop" learning in timetable planning. "Single-loop" learning is essentially about ensuring that we "learn to do things right", sticking within the existing frameworks and methods, but gradually learning to do so better. A timetable planner, learning over the years, where to place time supplements, etc. "Double-loop" learning is about "learning to do the right things", or in a similar context, learning that leads to updating of planning rules and tools.

Two studies that led up to this include Palmqvist, Olsson & Hiselius (2018), which covered interviews with planners and gave their perspectives on the process, and Palmqvist, Olsson & Hiselius (2017), which identified a number of different strategies in timetabling, and quantitatively evaluated their performance in terms of punctuality.

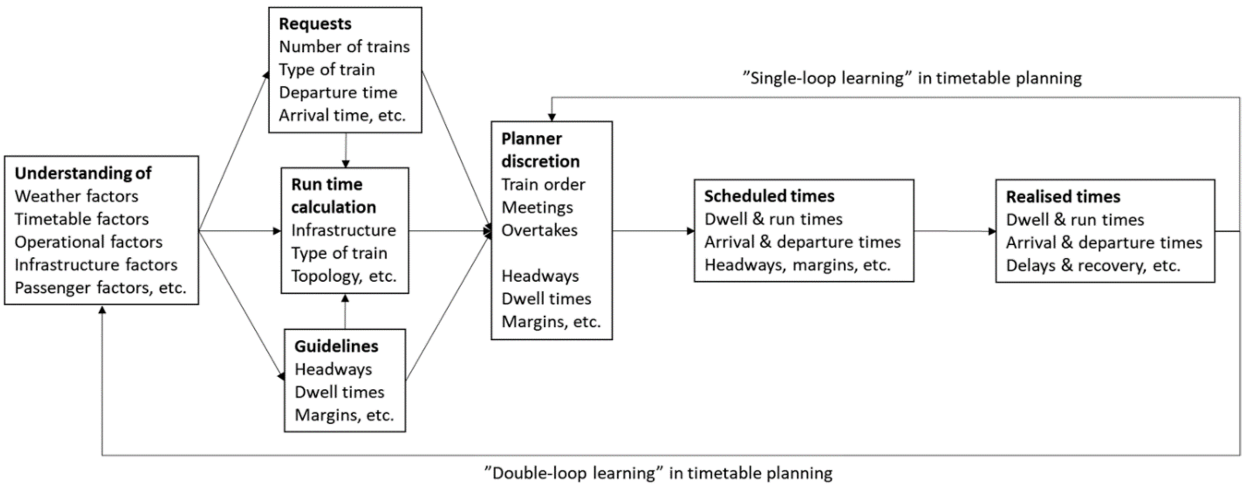


Figure 29: Single- and double-loop learning in timetable planning (Palmqvist, 2019).

A particular challenge that has received more recent attention is regarding scheduling of dwell times. Palmqvist, Tomii & Ochiai (2020) looked into these in more detail, using data from both Japan and Sweden, to help enforce the feedback loop between operations and planning of dwell times. Kuipers & Palmqvist (2022) also looked at learning from how dwell times vary over time, with the intent of using this to help schedule dwell times in the future, with less delays as a consequence. As a final example, Palmqvist, Lind & Ahlqvist (2022) investigated the operations of freight trains, along with why these deviations occur, aiming at a longer feedback loop than is usually practiced, with takeaways that range from practical timetabling choices to the design of the capacity allocation process itself.

At NS the timetable process receives feedback from operations in the following ways:

Basic operational data

Data on arrival punctuality is used to flag timetable issues. Weekly reports are made during the year, where the punctuality of train series is monitored, especially when the timetable has changed. This involves connections that are important to passengers, or operational processes such as extra shunting.

Advanced operational data

The automatic fare collection system 'OV-chipkaart' gives NS detailed data about passenger trips. There, the planned arrival time for each trip is determined and it is possible to see if the passenger arrived ('tapped out') at the planned time. A data analysis team reports issues back to the responsible operational and planning departments. Issues can for example be patterns or interchanges that are responsible for many late passengers.

Train telemetry

NS trains transmit position and speed data from GPS to their data warehouse. In collaboration with ProRail, speed distance plots are made for a train series, that show the speed of the trains. This is used to spot situations where drivers encounter a non-optimal signal aspect and have to reduce speed. Also, deviations from the permitted or advised speed are monitored.

Automated delay explanation

Data from NS and ProRail is used in a mutual system called 'Sherlock' that assigns delays to the most likely cause. This ranges from interlocking, rolling stock to infrastructure issues. This is done automatically for all delays registered on the network.

Customer feedback

NS has in its app a feedback option to ask customers about their experience. This can be about punctuality, seat availability, cleanliness and more. Also, complaints filed via customer care can be used in timetabling decisions.

Employee feedback

Employees give feedback, via an app or their managers, on timetabling issues and train capacity. This can result in small improvements of the timetable.

At SNCF, upstream operating studies are regularly based on exchanges with the timetable production chain:

- Timetable experts for the planning of train paths and capacity allocation.
- Local timetable offices for the design of track occupation charts and technical movements in stations.
- Regularity experts for the robustness and analysis of regularity data on the routes.

In the Spanish ADIF network, there is currently no automatic feedback loop between operations and timetable planning for improved capacity and punctuality. The process that is followed is shown in Figure 8-3 and described in more detail below.

Figure 30: The current planning process in the ADIF network.

As a first step, the Railways Undertakings (RUs) make the train path requests to the Infrastructure Manager (IM). These requests are studied by the area of the IM dedicated to the Capacity and Planning and the train running times are calculated as well as possible conflicts with other requests. Once it is adequate, the IM confirms the paths to the RUs. The Transport Plan is created and prepared as part of this step.

After this, the TMS receives the Transport Plan with the set of trains that will run through its control area. Since there is no continuous and bidirectional communication (see above) between the Capacity and Planning area and the TMS, the Transport Plan (enlarged with new train paths) is sent to the TMS through a conventional communication channel. Once received on the TMS side, it can be imported for tracking and audit.

The TMS is responsible for managing the train movements in its control domain. Using the information coming from the trackside signalling system (Interlocking (IXL) and RBC systems), the TMS determines if the train movements comply with their planning, or the trains have delays or are early.

The monitoring information from the TMS, i.e., deviations with respect to the planning (delays, etc), is collected in a repository which is used for the future replanning.

The concept of feedback loops between operation and timetable included in the WP8/WP9 is already implemented in the Spanish ADIF network, but this does not follow an automatic process. This feedback is done by reports from an analysis and traffic monitoring department provided to the Capacity Planning staff. The information of planning deviations from the repository, previously mentioned, is used in the feedback loop between operation and planning to achieve the improvement of the planning (optimisation). This means the use of the historical data (identification of recurring and repetitive events in the operation).

The current status of the TMS at Indra implies that there is not any direct channel to obtain data from operation to be implemented in planning. The current system follows the schema shown in Figure 8-4.

Figure 31: Feedback loop between operation and planning.

As can be seen in the figure, every input related to operation for timetable management and simulation is currently provided by the Post-Analysis and Reporting module. This means that there is not a direct relation between operation and timetable planning.

8.2 Feedback loops between operations and simulation

There is two-way feedback between simulation and operations. Operational data is used in simulation studies, mostly in the form of stochastic distributions derived from these data. These types of data are utilised for simulation studies at NS:

- Arrival delay distributions
- Running time disturbances
- Dwell time disturbances
- Entry time disturbances
- Turn-around time disturbances
- Failure rates of rolling stock
- Failure rates of infrastructure
- Probability of absent crew

Feedback from simulation to operations is more spontaneous, when needed. An example is when we observed in a simulation model that the anticipation trigger of the interlocking was too late for a train series, causing it to brake and have a delay on arrival. This was communicated to traffic control, who could change the settings of the interlocking to avoid the delay.

At SNCF Réseau, feedback loops between operational data and simulation models are relatively few in the context of upstream operational studies. The assumptions used in the simulation models (stopping times, turnaround times, margin recovery in the event of disruption, etc.) are nevertheless consolidated on the basis of feedback from the operational professions.

From a driver feedback point of view, two different approaches have traditionally been used - onboard data/measurements and simulator studies. However, these methods are associated with drawbacks. Onboard measurements are often time consuming and a logistic challenge, because simulator studies demand a well-prepared software environment. In countries with close relationships between infrastructure manager and operator(s), it is generally easier to get access to train onboard data than in countries with a deregulated market.

The EPA (ERTMS Protocol Analyser) tool (inhouse VTI development, see Section 6.4) is enabling a new method to remotely get access to driver data aimed for countries with a deregulated market. With the EPA method, basic train dynamics data for all types of trains running on ERTMS (with at least Level 2) can be collected over time in a smooth way. The general idea to use data in the ETCS protocol, process it and provide a sufficient accuracy in speed and acceleration predictions. From a radio communication perspective, ETCS data transfer is logged between the Radio Block Center and train, forming the input to EPA. The algorithm is time saving when it comes to train driver behaviour studies where several trains and drivers can be followed over time, and act as input to studies concerning acceleration and braking behaviour connected to signalling points, ETCS braking curves, and ATO development, see Figure 8-5.

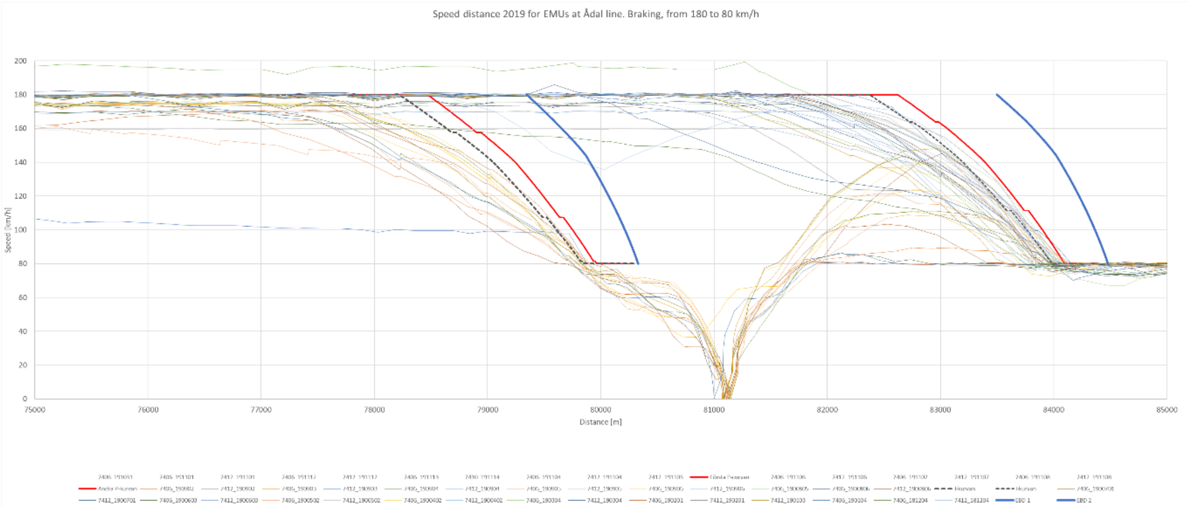


Figure 32: Braking pattern for EMUs at Ådal line evaluated by EPA.

The EPA tool is validated with EMUs (Alstom Coradia and Bombardier Regina), both configured with modern odometer systems. For a freight train, slip and adhesion is probably higher, implying a possibly higher position uncertainty. Therefore, there is a need for validating the method with freight trains.

The simulation environment provided by Indra, allows to modify the behaviour of the trains while running the simulation. This simulation considers the characteristics of the trains, the driving mode, the catenary and the topological characteristics of the tracks. Based on those characteristics, the simulator applies an algorithm that simulates the dynamic movement of the train in a realistic way. The simulator allows editing the following characteristics:

- Train movement: the system simulates the dynamic behaviour of the trains. This simulation shall consider the characteristics of the train composition (train physics), the driving mode, the catenary and the topological characteristics of the track where they run on.
- Driver management: The system simulates the behaviour of a train driver that operates under the parameters established in his data sheet.
- Interactive train driving: When simulation is running, the system provides a view to check and modify the train´s behaviour on the fly.

As it operates nowadays, the simulator performs its activity based on post-analysis and reporting data that can run after the operation. The aim is to improve the timetable by including in the simulation environment timetable modifications based on C-DAS data.

8.3 Feedback loops between simulation and timetable planning

Timetable analyses and traffic simulations are made to evaluate timetables and give feedback to timetable planning. The aim is to detect conflicts and improve timetable quality.

Often, the tool used by timetable planners is at a macroscopic level. It is unnecessary to include too many details, that might slow down the system when creating timetables. However, using a macroscopic model, means that the timetable might not be microscopically feasible. There is no way that the timetable planners can find out if the timetable is operationally feasible and conflict free or not. Therefore, it is useful to make microscopic timetable analyses and simulations, to give feedback to the planners. For example, until 2023, simulation experts of TRV needed to analyse the conflicts in the annual timetable using simulation tools. In 2023, a microscopic timetabling tool was introduced and the need for a separate conflict detection decreased. However, future traffic solutions are still evaluated via timetable analyses and simulation to give input to the timetable planning. Both new traffic patterns and infrastructure changes can be evaluated.

One of the main advantages of simulation is its ability to evaluate timetables before they are established and put in operation. By simulating a timetable (or a set of alternative timetables), the effect of stochastic primary delays can be estimated. By analysing the simulation results, timetables of poor quality can be avoided, and potential robustness issues can be detected and resolved. As noted in the previous section, timetable planners work manually with large-scale problems covering thousands of trains. A major challenge of using simulation as input to the planning is therefore that the resulting amount of data can be overwhelming. Furthermore, understanding the relationship between adjustments and expected outcomes can be difficult to manage in practice, which may require repeated trial-and-error to obtain desired quality.

To overcome these issues, Högdahl (2022) has developed methods that combine simulation and optimization to automatically improve a given timetable. The developed methodology is based on first generating a draft timetable, simulating the draft to evaluate its robustness, use the resulting delay data to generate an adjusted timetable by solving an optimization problem, and finally, validating the optimized timetable using simulation. In Högdahl et al. (2019), an optimization model for minimizing the weighted sum of scheduled travel time and predicted mean delay was proposed for the scenario with a fixed train order and fixed departure times. To overcome the limitations of only considering the given order of trains and departure times as fixed, Högdahl and Bohlin (2022) generalized the method by proposing an improved model for predicting delays. In Högdahl and Bohlin (2023), a model for explicitly maximizing punctuality was proposed.

In Högdahl and Bohlin (2023), a comparison between different strategies for improving timetable robustness was carried out. The comparison included their method for maximizing punctuality; their previous method from Högdahl and Bohlin (2022); light robustness (Fischetti et al., 2009);

robustness in critical points (Andersson et al., 2015); and two naïve strategies for allocating time supplements. The methods were evaluated in two scenarios using three performance measures. The evaluation indicated substantial differences in punctuality despite that the evaluated methods generated similar timetables. The two combined simulation-optimization methods (Högdahl and Bohlin, 2022, 2023) generated timetables with better overall punctuality than the other methods, whereas relatively small differences between the evaluated methods were observed in terms of end-station punctuality. None of the evaluated methods were able to outperform the others, which indicates that planners should not only rely on a single method for improving timetables.

8.4 Development needs within feedback loops between operations and timetable planning

Feedback loops between operations, simulations and timetable planning are essential to improving capacity of the line. There are several developments needed to advance the area.

One development need for the feedback loops from simulation to timetable planning is that we need models to simulate large networks. Today we are not able to evaluate traffic in large networks, which might result in missing information about widespread chain effects. We then risk not getting the complete knowledge of how the planning influences the traffic system as a whole. Another need is related to both operation, simulation and timetable planning and that is comprehensive analysis of historical data. Historical data of how train perform with a certain timetable can be used to give direct feedback to the planners. It is also important to use historical data to create accurate disturbance distributions that can be used in stochastic traffic simulations as well as to validate simulation models (see previous section on validation).

After many years of using simulation for assessing timetable robustness, a similar need arises at crew planning. As there is no way yet of evaluating a crew plan on robustness, simulation is seen as a promising method. By introducing similar or identical disturbances in a simulation model of a crew plan, different plans can be compared on robustness indicators. These simulation models and indicators have to be developed.

As explained more in detail in Section 8.1, in the Spanish ADIF network there is currently no automatic feedback loop between operation and time table planning for improved capacity and punctuality. Currently, this feedback is manually done by reports from an analysis and traffic monitoring department provided to the capacity planning staff. To be able to achieve an automatic process (this means an automatic feedback loops between operation and timetable planning) for the Spanish ADIF network, it was identified as required to define a standard communication channel/protocol (this means the way to transmit the information) and the format of the information to be exchanged.

It is important to always include the RUs in the information flow since they are the ones must request/accept the changes proposed by the IM. The RUs as users or consumers of data should participate in the communication channel definition.

To improve the planning process, the capacity and punctuality of lines with energy efficiency, it is essential to take advantage of the operational feedback from the historical data (e.g., driver behaviour data or usual delays in specific areas), and from the use pf the innovative technologies such as C-DAS, ATO over ETCS, ERTMS Hybrid Level 3 and ERTMS moving block. Continuing to investigate and develop these aspects is relevant for process improvement.

Thanks to the refinement of the train running that could be achieved with the use of historical data on the train operation and the introduction mainly of ATO over ETCS on the lines, the planning could be improved and adjusted.

The inclusion of innovative technologies such as ERTMS HL 3 or ETCS L3 with moving blocks in the process is expected to have a significant effect in the increase of the line/network capacity.

Indra has identified a development need in performing feedback loops between C-DAS and TMS. The following aspects need to be taken into account:

- Define the protocol exchange channel between C-DAS and TMS.
- Define how it is simulated
- Define how it affects the current plan (how data from operation can affect the short-term planning by considering real time data).

An identified need from CAF is improved data analytics to the continuous optimisation of planning, based on real data. When the timetable planning is made, an ideal result is obtained in which there is no disturbance or inconvenience, but we have to be able to adapt these schedules to the habits of passengers. In other words, we need to be able to design patterns that allow us to explain the disturbances that occur in order to adapt the planning. Initially, this is intended to be done manually, as we must understand how these patterns work and generate good planning recommendations. In the future, we would consider generating these patterns in an automated way.

In order to be able to do all of the above mentioned, we will have to work with Big Data fed by for instance passenger information, time at stops, switches, accelerations, brakes. This will be done according to the diagram shown in Figure 8-6.

Figure 33: The data process phases needed to feed the Big Data.

The following phases as shown Figure 33 are considered:

1. Data obtention: this is the way in which we are going to obtain all the data and put it into the database. The main sources are all those that will allow us to obtain data that will help us to generate the KPIs to improve the capacity and punctuality of the service. We would have as the main source the data provided by the signalling systems as ATO, ATP, TMS systems and other relevant subsystems. Also, other sources of data will be provided by third parties, such as the driving information, historical data, etc.

2. Data collection: this is the collection of raw data. It is the input that will feed our Data Analytics process with different information such as: data received by different operators, train characteristics, line topology, and speed limits. It needs to be cleaned to remove any errors, inconsistencies, or missing values. This step involves checking for outliers, duplicates, and formatting issues.
3. Data preparation: this is the process where all the data is going to be normalized. We will develop an algorithm that will allow us to group all the data following a model generated on the basis of the data received that allow us to properly analyse it. There may be different databases interrelated among them, prepared as a cluster.
4. Data analysis: when we have the database normalised (meeting the requirements that we have defined), through different algorithms we start to process the data to obtain KPIs (for example, we have energy consumption data and with this algorithm we obtain a KPI that will allow us to improve this consumption over time).
5. Data output: they are the KPIs that we obtain given the data analysis. They will give us reliable information, reliability of the results to be able to apply it.
6. Storage: it is the physical storage of this information so that it is always available.
7. Visualization: generation of data graphics that allow us to visualise these KPIs and detect possible incidences.

Once we have completed the process, data interpretation will take place. The results need to be interpreted to extract conclusions and make decisions. This involves using the insight gained from the analysis to be able to make informed decisions and take the optimal action.
The place where all this information accumulates is already structured within CAF environment, but it is not yet defined how to make use of the data. The communication to transmit information from one site to another would correspond to the arrows in Figure 34.

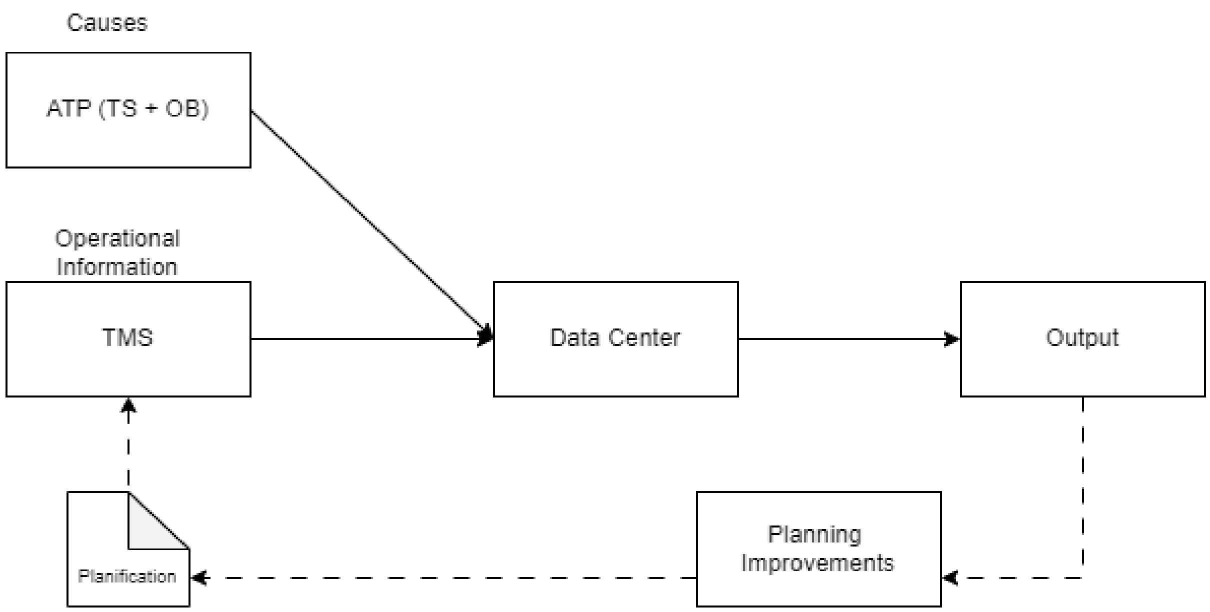


Figure 34: The communication flow.

Through the communication in Figure 34, we can reach a number of different conclusions or benefits:

- We will mainly use it to be able to correct any errors that may arise in planning in order to adapt to the new realities that may appear.
- In addition, we can achieve other benefits such as detecting anomalies at the operational or technical performance level that may be subject to conditions or patterns that require not a change in planning but a change in the setup of the system.

We also need to be able to represent different KPIs through this data analytics to detect situations that change the transport model (i.e., in the case of a station closed for construction work, the habits of users may have changed, and they may opt for other means of transport, making it necessary to rethink the original planning).

8.5 Conclusions of feedback loops

Feedback loops between operations and planning are essential to improve railway planning. When we evaluate operational or simulated performance, we can measure the quality of the plan and give input to the planning. Within this area, the following development needs are found:

- Improved simulation models for large networks
- Better analyses of historical data
- Improved evaluation of crew plan robustness
- Automatised feedback from operation to planning
- Feedback loops between C-DAS and TMS
- Improved data analytics and better communication between data and planning modules

9. Research on capacity effects of TMS-C-DAS/ATO/ETCS

There is a need to develop more knowledge about the capacity effects of TMS-C-DAS/ATO implementation because these systems aren't (fully) implemented in simulation systems and the effects in comparison with the current practise aren't fully known. Furthermore, research on capacity effects of ETCS developments will be included, because of the development of Hybrid Level 3 and because capacity effects of different ways of implementation aren't fully clear or fully satisfying. This section describes the current status, development and evolution in different countries.

The content in this section is written per partner, because then we only have to introduce the methodologies once. Most partners have included information about research on ETCS as well as TMS/C-DAS/ATO. The order of the paragraphs is from northern partners to southern partners and grouped by country.

A question that can be answered at the end of this section is whether there are synergies possible in research topics and methods.

Factors affecting the capacity and how they are measured

The capacity effects of ETCS can be measured by differences in running time, headway, and capacity utilisation according to UIC-leaflet 406 and punctuality. These factors are influenced by changes in the braking curves, speeds, distances between the trains and disruptions/delays (see Figure 35 for detailed information).

ATO does not change the block sections, but might change the drive behaviour of trains, i.e., how trains react on the provided information.

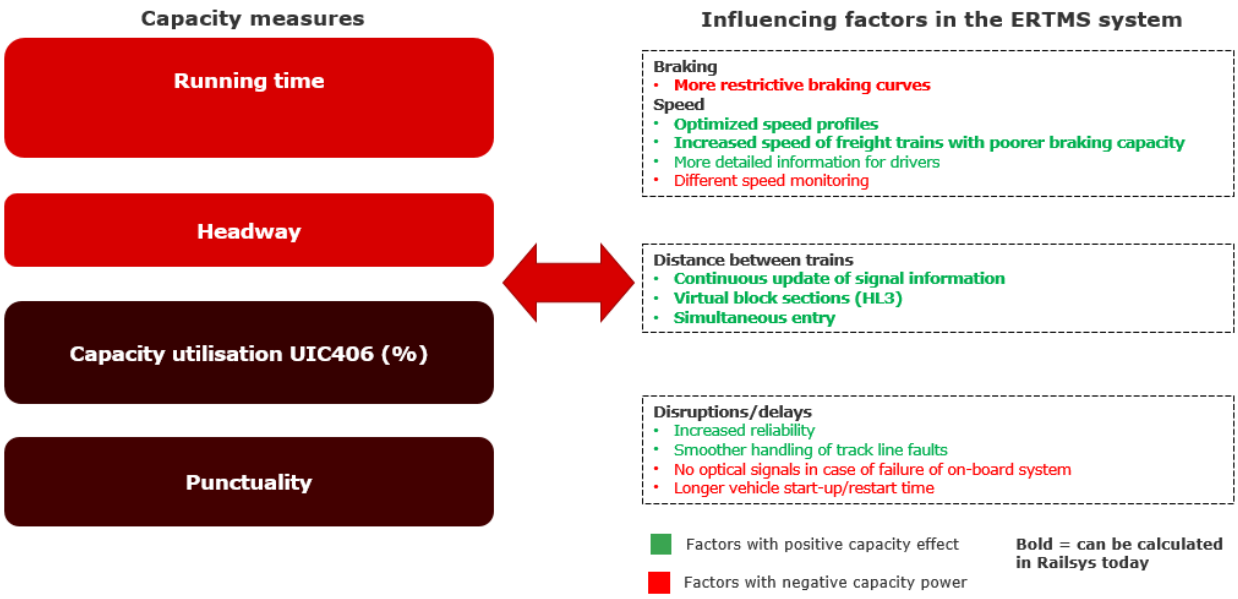


Figure 35: Capacity measures and influencing factors in the ERTMS system.

9.1 Research on capacity effects of TMS-C-DAS/ATO/ETCS at TRV

ETCS capacity research at TRV

ETCS capacity studies are performed at TRV in early stages of the ERTMS design phase. The aim is to find possibilities to improve the signalling layout and ensure that the capacity will be at least the same as with the current signalling system. These studies are in general performed in RailSys, where the ETCS signalling system can be modelled in detail. For this purpose, TRV has a national infrastructure model in RailSys and certain guidelines for RailSys ERTMS studies (Trafikverket, 2022).

There are still very few lines installed with ERTMS and according to the current implementation plan it will take a long time to have most of the lines equipped with ERTMS.

Capacity study of ETCS L2 and HL3

Capacity effects of implementing ERTMS on the Southern mainline Stockholm – Malmö in Sweden have been evaluated in 2021. Comparisons have been made between the current signalling system ATC, ETCS L2 1:1 (same signal positions as in ATC) and ETCS Hybrid L3. The running times and headway have been calculated in RailSys, where the ETCS signal positions are modelled in the same locations as in ATC. Any differences in driver behavior between ATC – ERTMS due to different driver interfaces have not been included in the study. In the study, it has been assumed that the driver brakes according to the ETCS permitted curve.

Running times

In **Fehler! Verweisquelle konnte nicht gefunden werden.** an overview of different signalling s ystems and running times is presented. The running times with ERTMS were extended compared to the current signalling system ATC due to more restrictive braking curves monitoring the driver. By optimizing the speed profile during the ETCS implementation, positioning the danger point in front of the EOA and to equip the onboard units with service brake feedback, the extended running times may be reduced and even shortened.

Table 10. Running time differences ERTMS-ATC for different train types

Running time difference ERTMS - ATC, Flemingsberg – Lund	High speed	High speed	High speed	Regional	Regional	Regional	Freight	Freight
	X2	X2	X50	X50	X50	X50	BR185, 1400 ton	BR185, 1400 ton
	0 stops	9 stops	0 stops	9 stops	20 stops	36 stops	0 stops	6 stops
Without speed optimization, Without SB-feedback	1.2%	1.3%	0.9%	0.9%	1.5%	2.0%	0.6%	1.3%
Without speed optimization, With SB-feedback	0.7%	0.6%	0.7%	0.5%	0.6%	0.8%	0.1%	0.6%
With speed optimization, Without SB-feedback	-0.6%	0.4%	-3.5%	-2.9%	-1.7%	-0.4%	-0.1%	0.8%
With speed optimization, With SB-feedback	-1.1%	-0.4%	-3.8%	-3.4%	-2.5%	-1.7%	-0.4%	0.1%

ERTMS provides the possibility of a more detailed speed profile without additional signal objects in the track and also individual speed profile for different train categories. By the service brake feedback function the braking curves will be updated depending on the brake pressure, and thereby it will be possible to reduce the restrictiveness of the braking.

Headways

The principle for HL3 is shown in Figure 36, where virtual blocks in combination with trains with integrity control can improve the headway and capacity. When physical blocks are subdivided in multiple virtual blocks, trains with train integrity can be followed quicker by the next train. So, without increasing the number of objects in the tracks, the track capacity can increase when short following is necessary.

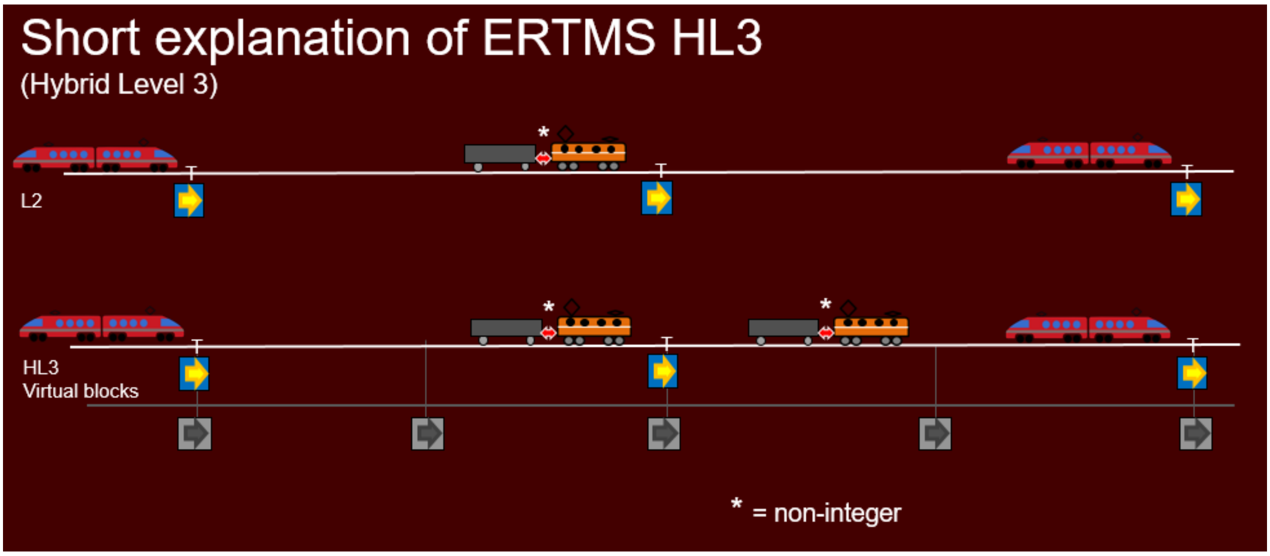


Figure 36: Factors affecting the capacity for a double track line.

Headway calculations for the Southern mainline indicates that for two following passenger trains without stop the headway may be improved by roughly 30 seconds by ETCS L2 and 1 min for ETCS HL3 compared to the current signalling system ATC, see Figure 9-3.

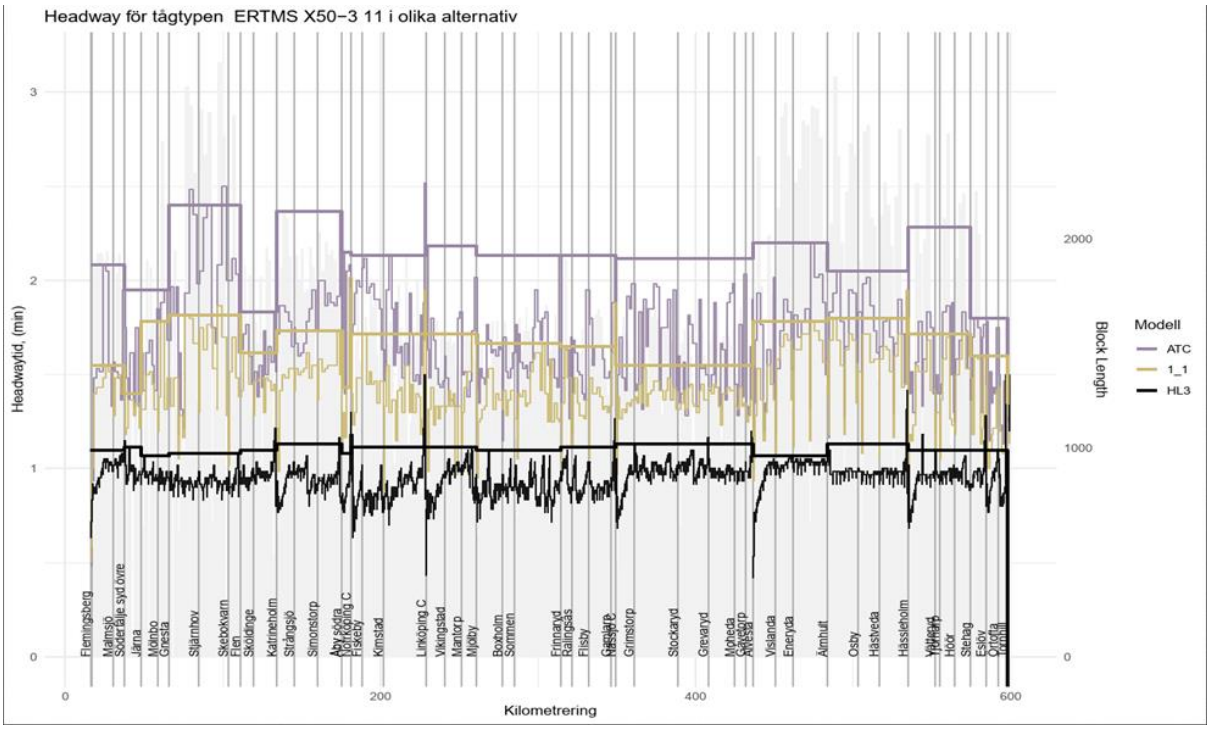


Figure 37: Variation in headway for the Southern mainline between ATC, ETCS L2 and ETCS HL3.

Capacity utilisation

For double track lines two main factors affecting the capacity are speed differences between trains and headway. The speed differences between different train types are in general large in Sweden and can prevent the possibility to achieve higher capacity even if the headway would be reduced.

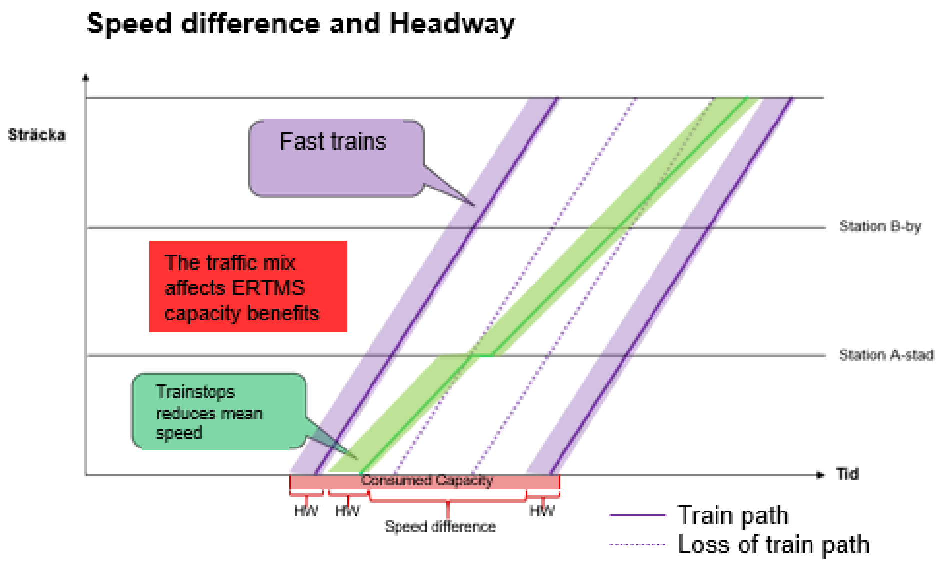


Figure 38: Graph to explain how headways are influenced by speed differences between trains.

Hence, the capacity benefit of ETCS varied between different sections of the network depending on the mix of traffic, see Figure 39. On parts where the traffic was homogenous with mostly integer passenger trains and small speed differences the capacity effect was larger, while on sections with a high level of mixed traffic and large proportion of freight trains the capacity benefit was smaller.

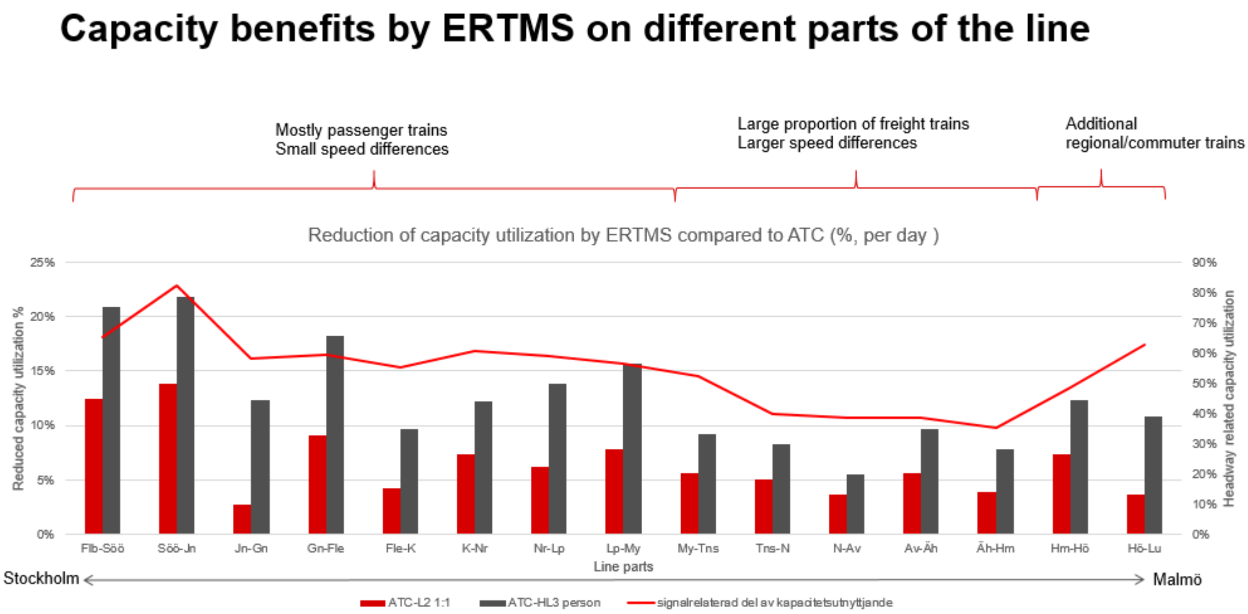


Figure 39: Capacity benefits by ERTMS on different parts of the line.

Conclusions

From the study it can be concluded that the biggest effect of ETCS HL3 was on lines with uniform traffic and small speed differences. The capacity gain corresponds to approximately + 1 train paths/maximum hour and direction or decreased capacity utilization per day by an average of 11% compared to the current signalling system ATC. It was also found that if freight trains are equipped with train integrity, capacity utilization will instead be reduced by 14 %.

Furthermore, the study recommended to consider the following questions in the further work with ETCS implementation:

- Ensure operators have the right functionality in the locomotive's on-board system, where it is important to enable optimized braking curves by implementing service brake feedback.
- By optimizing the speed profile when implementing ERTMS the negative effects of the more restrictive braking curves can be reduced or even give shorter travel times.
- In strategic locations it may be motivated to optimise the signalling layout for L2 until HL3 is implemented.
- Prioritize development of virtual functionalities such as train integrity, HL3, and virtual simultaneous entry (i.e. without extra signals).

Experience of ATO in Sweden
In Sweden there are no installations of ATO on mainlines and there are no plans to implement ATO at the moment. An initiative to at least explore what the railway industry, mainly the RUs, think about ATO started in spring 2023, since the capacity benefits are interesting.

A study made by Lindbergh (2022) on the capacity gains on a commuter train tunnel through the centre of Stockholm with GoA 2 showed a reduction of the occupation of the lines by around 4% by reducing the driver variation. In a GoA 3 or 4 system there could be other benefits. In a study by Jansson (2020) the effects of drivers on delays during 2019 were analysed in Sweden based on four categories: Driver (general category), Driver missing, Crew change, and Break. For passenger trains the driver categories amounted for 1.1% of the total delay time and 6.1% for freight trains. The same study also studied the effects by train drivers on requested stops by the train operators. For each requested stop the train operator should state the purpose of the stop and there are two actions closely related to the driver: Crew change and Break. During one year (2019) the total amount of dwell time due to stops that only included Crew change or Break amounted to 180 hours for passenger train and to 7,000 hours for freight trains. The results show that freight trains need to stop because of the driver much more than passenger trains. These stops would not be needed in a GoA 3 or 4 system.

GoA 3 and 4 systems would also entail new challenges as the train driver will no longer be in the cabin. One of these challenges is handling trackside sensor alarms. Trackside sensors are placed along the railway network in order to monitor the passing vehicle for damages. The purpose of the trackside sensors is to prevent serious accidents such as derailments. When an alarm occurs the train driver needs to stop and manually control the alarms, today resulting in a delay of 20-25 minutes in Sweden. In a GoA 3 system this task could most likely be performed by the train attendant that is still onboard the train. In a GoA 4 system, where there is no personnel onboard the train, a person needs to be sent to the train and control it. This would entail additional delay in the form of the response time besides the 20-25 minutes required for the actual control.

Up to GoA 3 the system effects on the capacity are expected to be positive in terms of automated driving with no driver variation. In addition, the delay caused by the driver is expected to be reduced. However, in a GoA 4 system the need of external personnel to handle unplanned events on a train would entail delays in terms of response time. If GoA 4 is to be implemented on the mainline the number of unplanned events needs to be reduced in order to reduce the frequency of sending out external personnel.

Experience of C-DAS in Sweden

TRV and the railway undertaker LKAB performed a test on Malmbanan 2011-2016, where C-DAS was installed on all iron ore trains, which was approximately 70 % of all traffic on the line. The C-DAS system consisted of both an on-board system and a Traffic Control Centre system.

The resulting effects of the test mainly consisted of improved communications between the traffic controller and the driver, a better mutual understanding of the traffic flow between traffic controller and driver and lower energy consumption of the train. However, no direct results on capacity or punctuality were noted. Reasons for this could be that not all trains were equipped with C-DAS and a resistance from some drivers to use the system due to early technical issues a lack of confidence of the system.

However, there might be some indirect capacity effects, as fewer phone calls and better mutual understanding between driver and traffic controller lead to a lighter workload for the traffic controller. This allows the traffic controllerto focus more on better planning and reduces the number of mistakes resulting from a high workload.

The next C-DAS test phase with focus to investigate capacity effects of C-DAS is postponed since digital graphs are not available at the moment due to the implementation of a new timetable and TMS system. When the digital graphs are available again new parts of the railway net will be available for train operators to use for C-DAS with our existing solution, which is not adapted to SFERA standard. There is parallel work ongoing of implementation towards the SEFRA Standard and the plan is to have a pilot running before the end of 2023.

9.2 Research on capacity effects of TMS-C-DAS/ATO/ETCS at TRV A.E. VTI

Driver simulator

The VTI train driver simulator includes an ETCS version that has been used for some years within train driver training and for research purposes. In Sweden, three main lines are equipped with ETCS level 2 and the simulator replicates one of those lines (Ådalsbanan). While the Train operating companies (TOCs) use the simulator for training purposes, VTI have used it for human behaviour research, for example in the following studies:

- A study using an experimental design was conducted with 16 experienced ERTMS drivers of which half used the simulator for practice and the other half practiced in reality, according to the TOCs standard procedure (the theory was the same for both groups, Olsson, N., Lidestam, B., & Thorslund, B. 2022). The participants then conducted a simulator test to measure the learning outcome through number of driving errors and instructor evaluation. Also, a questionnaire was filled in to capture how different factors, as number of repetitions, age and experience correlated with performance. The simulator test results showed a significantly better performance by the simulator group who committed 38 percent fewer driving errors and received a 34 percent higher score from the instructor compared to the control group. An important reason for the better performance was number of repetitions of special cases. A conclusion from this is that simulator training, to be effective and thus increase railway punctuality, should focus on training of special cases and especially those that reality does not offer sufficient training.

- Effects of baseline 3.6 on train driver behaviour. In this study approximately 50 driving students were driving a test track with two different versions of the ETCS onboard version, 3.4 and 3.6. The effects on capacity, energy usage and driver workload were measured.

ATO and C-DAS

In order to fulfil the ATO GoA2 research objectives, there is a need for simulator development. Generally, support for TSI 2022 and the ATO over ETCS functions need to be integrated in the simulator baseline. In TSI 2022 proposal a new packet switched data channel to the train is introduced, and the ATO driving functions are defined in Subset 126, see Figure 9-6.

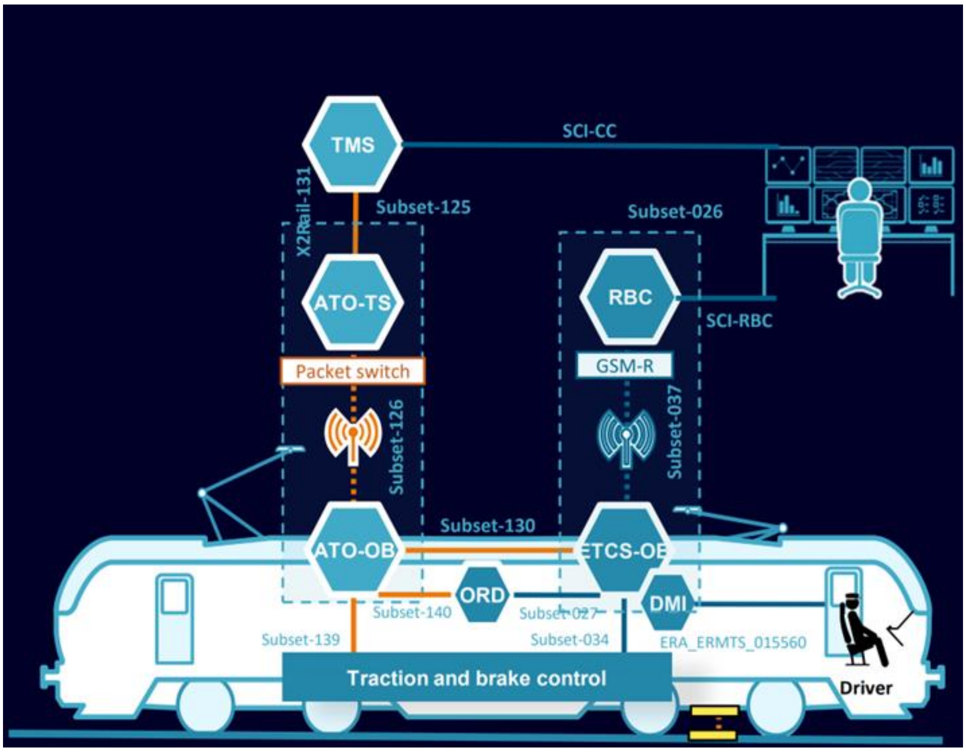


Figure 40: The two communication channels with its corresponding specification.

From a development horizon the new ATO functions must interact with the simulator baseline software according to Figure 41.

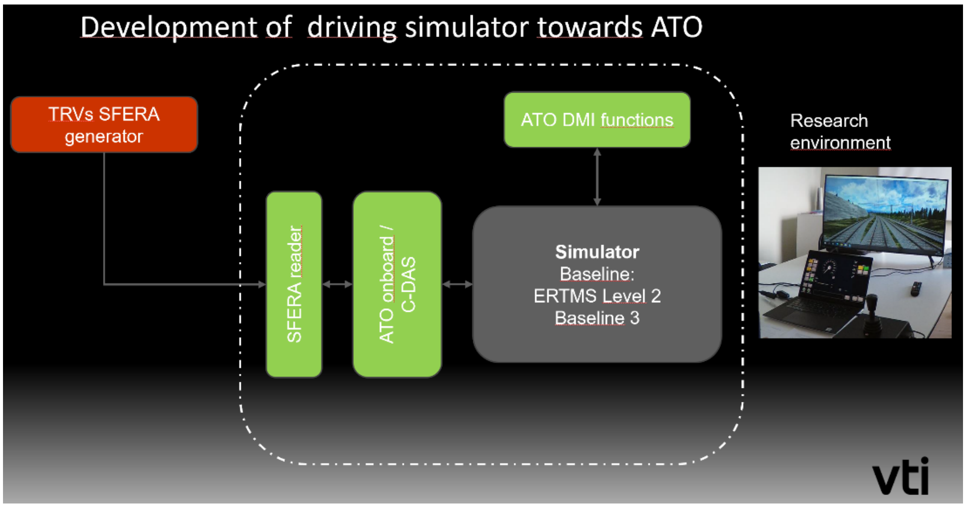


Figure 41: System overview of development of VTI train simulator towards ATO GoA2 and C-DAS. Add-on modules in green, Trafikverkets SFERA-generator supporting Subset 126 and ATO TMS interface.

In the software block *ATO DMI, functions* are added to support the new DMI which is defined in IRS 90940. The ATO onboard block has the responsibilities for the automatic driving according to the Time Table Speed Management (TTSM) or Supervised Speed Envelope Management (SSEM) speed computation strategies. The SFERA reader supports the SFERA standard of reading journey profiles and segment profiles from an TMS environment.

EPA

The strength of the EPA tool is its ability to contribute extensive statistics from operations under normal conditions, resolving driver behaviour and giving a fair view of the capacity gains going towards automated operations. More information about the tool can be found in Section 8.2.

HL3

VTI have conducted simulation of Hybrid Level 3. Key performance indicators, headway on the microscopic level; and capacity utilization, punctuality and robustness for the network level are identified and used to evaluate the effects of implementing Hybrid Level 3. A case study has been conducted to look at how the performance indicator headway is affected by different lengths of virtual blocks. Using a microscopic model in the simulation tool RailSys, the interaction between two trains has been studied. The results show that the headway is similar for different virtual block lengths. Furthermore, two case studies of a real line in Sweden, with the addition of virtual blocks, study how the performance indicators punctuality and capacity utilization are affected by the share of Level 2 and Level 3 trains in a HL3 system. Simulations have been conducted in RailSys, using a real-world infrastructure and a complete timetable. The result shows that the share of Level 2 and Level 3 trains has minor effect on punctuality and capacity utilization. However, some limiting factors have been identified that acts as bottlenecks for the potential capacity gains. For example, stations (platform capacity) and switches on the line, as well as restrictions of the simulation tool (shorter blocks reduces the effectiveness of detecting conflicts in the simulation).

9.3 Research on capacity effects of TMS-C-DAS/ATO/ETCS at Prorail

Research on the Dutch railway network has proven that the traffic within the coming 10-20 years cannot be handled given the current signalling and train protection system as well as the current infrastructure. ETCS is obligatory according to European rules. However, knowledge is limited and the capacity benefits are unclear. Although ATO has been implemented a lot, the implementation on a classical railway network is something new. In the next subsections the current status of research and further needs of ProRail are described.

ERTMS

Within ProRail numerous calculations on the capacity effects of ERTMS L2 have been conducted, usually to check if a future infrastructure project will provide the same capacity as currently or the necessary capacity increase for future timetables. The investigated corridors range from regional traffic only, mixed traffic of Intercities and regional trains and the complete mixture with freight trains added. In the past these simulations were done in an in-company developed tool called FRISO and are still performed with OpenTrack. On the densest corridor between Schiphol Airport, Amsterdam, Almere and Lelystad (the SAAL-corridor) it has been proven that only L2 is not enough for the capacity requirements of the future. Therefore, studies with ETCS HL3 have been conducted, both without and with ATO combined, to investigate if that would lead to the necessary capacity. The results of these studies, performed by using RailSys, can be found in Jansen (2019) and Vergroesen (2020). Besides demonstrating good capacity effects for HL3, the studies also indicate that perturbations dissolve more quickly compared to L2. This leads to a reduction of secondary delays (see Figure 42).

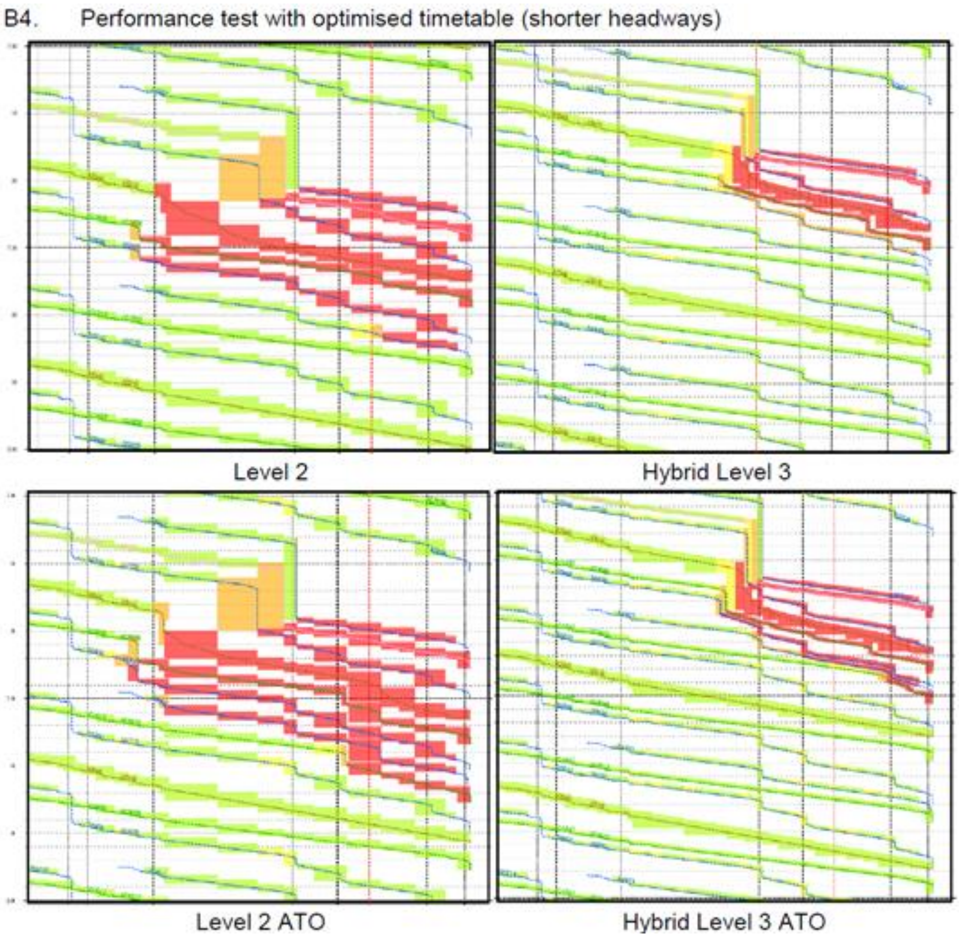


Figure 42: The same perturbation dissolves quicker with short blocks with HL3 compared to L2, although even more trains are run in the HL3 scenarios compared to the L2 scenarios (Vergroesen, 2020).

For future simulation studies development is needed for the braking behaviour, both to have a better understanding how train drivers react on the braking curves and to implement this knowledge into the tools. A better implementation will also lead to more stable results compared to automatic driven trains. Another necessary development is to have an easier implementation of HL3 in order to make quicker analysis of the difference between L2 and HL3.

ATO

In the Netherlands ATO GoA2 is researched as a possible future addition to the rail network. First estimates for planning norms with ATO were investigated. This has led to the following assumptions for timetable planning:

- The maximum average braking rate (when not limited by the ETCS braking curve) is increased to 0.8 m/s^2 instead of 0.5 m/s^2 for passenger trains.
- The running time slack can be reduced from 7% to 4%.
- The headway buffer time can be decreased by 15 s. These parameters are used for future timetable design with ATO.

In general, these assumptions seem to be on the conservative side since higher values are found around Europe for ATO GoA2 projects.

The dynamic effects of ATO and TMS were investigated with FRISO on the Betuweroute freight corridor. The robustness of the timetable has been improved due to a higher punctuality level with ATO and TMS and less unplanned stops were measured.

For a case study on the SAAL corridor it was concluded that the track occupation rate will decrease on average 18% if according to expectations the buffer times and running times will be shorter (Vergroesen, 2020). The combination of HL3 and ATO is complementary, so combined they can result in up to 43.8% on average lower occupation rates according to the same study.

TMS, C-DAS, NG-Brake
Hornung (2023) concluded that a non-centralized TMS can mitigate delays by an additional 5% and the proposed algorithm can decrease the difference between the maximum and minimum delay by 90% compared to the current manual work at traffic control for the Brabantroute. Further research is needed on how to incorporate the TMS in current simulation tools to study the network effects. This means that the behaviour of a single train will not only depend on its own signal aspects anymore, but also on other trains in the network.
Within ProRail no railway capacity studies on C-DAS and NG Brake systems are conducted yet.

Use cases
For studies on the different system developments mentioned above ProRail proposes the SAAL-corridor (dense passenger traffic, with hardly any freight trains) and Brabantroute (mixed traffic with many freight trains).

9.4 Research on capacity effects of TMS-C-DAS/ATO/ETCS at NS

NS started their ATO research program in 2019. Broadly speaking, there were two main reasons for doing so:

- **Internal developments and future passenger growth** - more capacity is expected to be needed in the future to meet increasing passenger demand from passengers and freight operators (Meijeren & Van der Lande, 2021). In addition, we have goals in the area of cost-efficient operation and sustainability.

- **International developments** - worldwide, ATO concepts are being widely experimented with and the first tenders or implementations are realized. At the administrative level, too, Europe is committed to automation and digitization in the rail sector (Chéron, 2020).
  a. Research on the current status of ERTMS/C-DAS/ATO capacity effects and the need of future development. - the speed of technological change is growing and new technology is becoming more accessible. This brings the realization of autonomous vehicles ever closer. This may have a negative impact on the competitive position of rail transport and lead to an undesirable change of transport mode. There is a sense of urgency in the rail sector to innovate more quickly.
  b. Current studies for ATO capacity Hanzelijn (CAF, 2019). This was the first step in learning how to run with ATO and to gain insight in its sub-technologies. During the experiments we showed that ATO is able to drive very accurately according to the timetable. This is also confirmed by other European carriers and rail parties. However, given the small number of

trips driven and the circumstances in which the experiments took place, it was not yet possible to draw conclusive conclusions about ATO's performance in the Dutch context.

Human factors research during the experiments taught us that, in order to achieve optimal performance with ATO, the interaction between driver and system must be explicitly designed. The tested ATO over ETCS solution did not yet sufficiently take into account the human need from the driver.

In 2020 NS performed their first experiments with ATO over STM (ATB-EG). This testing project was completed in 2021. The reason NS wants to experiment with ATO on a class B system is because of the long-planned lead time of ERTMS rollout on the infrastructure in the Netherlands. We investigate ATO over STM as a potential migration step towards ATO over ETCS. The first tests confirmed the technical feasibility of the concept and showed sufficient potential for NS to continue developing and testing the solution.

Follow up endurance tests (200 test runs) with both ATO GoA2 over ETCS and STM are planned in 2023 with a retrofitted SNG train in cooperation with CAF. This will include a more mature and further integrated ATO solution compared to our earlier testing solutions. The same train is also equipped with technology to conduct our first experiments in a shunting context using remote train operation (RTO) with GoA1, RTO with GoA2 and GoA4. These experiments are also planned in 2023.

9.5 Research on capacity effects of TMS-C-DAS/ATO/ETCS at SNCF

At SNCF several operational studies using the tool DENFERT (based on RailSys) have been carried out to determine the capacity effect of ERTMS L2, including stochastic simulations. The main difficulties were determining the specific system times to be taken into account according to the type of signal box and determining the ERTMS parameters for each item of rolling stock. SNCF also had difficulties to implement the transition between ERTMS and national systems (BAL + KVB) since RailSys isn't able to correctly implement of the transition yet. Furthermore, trains always brake according to the ERTMS curves, even if they are stopping at a station. This leads to inaccurate run time, indeed, in France, trains mostly stop at stations without any restrictive signalisation.

These studies mostly conclude that ERTMS L2 has a positive effect on the capacity. The level of the effect depends on the current performance of the signalling system, and the type of traffic (varying speeds, presence of freight trains, etc.).

SNCF has not carried out studies on ERTMS HL3, so they have to go deeper into the HL3 modelling in RailSys. However, they have carried out a study that combinedERTMS L1, C-DAS and ATO. The aim of the study was to assess the capacity contribution of the introduction of ATO on a French network line. This line would have been operated by trains in ATO, and others in C-DAS only. The ATO would have been set up on ETCS L1. SNCF has only worked on microscopic timetable construction and has not carried out dynamic simulations.

To model ATO and C-DAS, SNCF has taken certain assumptions into account:

- Reduction of buffer times
- Better braking performance for ATO trains

These hypotheses have been defined in conjunction with SNCF Réseau's railway robustness experts.

SNCF has not gone into detail in the modelling of the ATO in Denfert, because they soon realised that it would be difficult to implement the system on this line and improve the capacity. This is due to the characteristics of the infrastructure on the line studied (need for additional heavy infrastructure such as new points, mixed traffic, ERTMS L1 with punctual data transmission making the ATO less efficient). To date, there are no plans to deploy ATOs on the French rail network. For this reason, the need for ATO simulation in France appears limited.

Simulations have been made for "Nexteo" system which is a CBTC. This system is about to be implemented on the line E of the Paris RER. It is an autonomous driving system but it is not based on ETCS, it is a specific system for Paris RER developed from CBTC solution. The simulation of this CBTC system is outside the scope of this work package and will therefore not be dealt with here.

9.6 Research on capacity effects of TMS-C-DAS/ATO/ETCS at ADIF/RENFE/CEDEX/INECO

In Spain there is no mainline equipped with fully operational and functional ATO. The closest one is the C5 commuter line in Madrid, but it is an LZB installation that is more than 25 years in service and that is planned to be replaced by ERTMS. There is also no ATO/ERTMS. The closest thing is the ERTMS L2 mounted on some lines, but it has no commercial stop data, among others. In on-board systems, Siemens ICE 3 tracks the traction curve, but is not an ATO as such.

In terms of capacity allocation, the implementation of TMS has a positive impact on punctuality by obtaining more accurate and regular schedules, which improves the reliability of rail service as the system can provide more accurate information on the location and speed of trains. The TMS can enable better capacity management in high-demand situations, as the system can automatically adjust the speed and time interval between trains to optimize line capacity. In addition, it can also improve strategic planning and decision making at the railway level, which can contribute to better capacity allocation and larger efficiency in rail operation. There is an improvement in safety, helping to reduce the risk of rail accidents by providing real-time information on train location and speed.

The focus should be on the ATO over the ETCS. The reasons for this are:

- It is already foreseen to include the ATO GoA 2 in the CCS TSI 2022/2023
- In FA2, the different partners are already working on the next level, the specification for GoA3 and 4 of the ATO over the ETCS

The implementation of ATO has a positive impact on the scheduling and capacity management of

the rail line, as the system can provide more accurate information on travel time and line availability. In addition, it can enable better capacity management in high-demand situations, as the system automatically adjusts the speed and time interval between trains to optimise line capacity. We have increased safety of the rail system by reducing the risk of accidents due to human error, this which could allow for greater capacity allocation by reducing safety risks. Finally, we will have a better energy-efficiency rate by optimising train speed and minimising sudden braking and acceleration.
With the use of the ATO over ETCS it is possible to refine the trains running and with this to improve the planning. Thanks to the refinement of the train's running, the planning can be adjusted, reducing the safety margins (used to represent human factors…), and therefore managing to contribute to an increase in capacity.

ATO is able to achieve a more precise acceleration and following of braking curves. Although the ATO over the ETCS helps to increase the capacity, the innovative technology that is expected to impact beneficially and directly on the capacity is the implementation of the ERTMS level 3 (Hybrid level 3 or L3 moving blocks). ATO is a system that allows trains to operate without human intervention, or with minimal human intervention, by using a combination of sensors, algorithms, and control systems. The ETCS onboard computer calculates braking distances in real-time by predicting the decrease of the train speed in the future, from a mathematical model of the train braking dynamics and of the track characteristics ahead. The European Union Agency for Railways has published a document on the harmonized braking curves for the ETCS (ERTMS UNIT (europa.eu)).

The implementation of ETCS has a positive impact on the scheduling and capacity management of the rail line, as the system can provide more accurate information on train location and speed. In addition, it can enable better capacity management in high-demand situations, as the system can automatically adjust the speed and time interval between trains to optimize line capacity. In addition, it improves interoperability between different rail systems across Europe, which facilitates capacity allocation on international lines. In short, we have improved efficiency and safety, reduced travel times and achieved larger regularity and punctuality.

CEDEX is committed to implement the ATO interfaces in the ERTMS Traffic Simulation Test Environment, as well as to introduce laboratory modules associated to absolute safe positioning, communications (FRMCS oriented) and automation. This transversal strategy is also participated with other FP, mainly 2 and 6. As mentioned in 6.9.1, the main target is to prepare the laboratory for the evaluation/validation of lines implementing the new technologies for ETCS (mainly, ATO + Level 3 + FRMCS). CEDEX is also planning to increase the usability of the test environment for operational issues, not strictly for network capacity analysis, but for the evaluation of specific operational concepts (train journeys, timetables, degraded situations, etc).

9.7 Research on capacity effects of TMS-C-DAS/ATO/ETCS at Indra

The objective of Indra is to evaluate the capacity of the infrastructure with new elements such as C-DAS analysing if it can be used to improve the timetable. To meet the goal, the starting point is the hypothesis that current plans are performed based on the most restrictive speed among all the affecting speed restrictions.

During the operation, a scheduled plan is not successfully accomplished as trains do not usually run at maximum speed, they usually run at lower speed depending on external conditions initially not taken into account. One of the systems that affects the real speed of a train during operation is the C-DAS (that can be simulated), which indicates to the driver the speed of the train.

To achieve the objective previously mentioned, the procedure will be based on comparing the initial plan of the trains to the real run followed by trains with C-DAS. Trains can be simulated and KPI can be defined to measure the results.

9.8 Research on capacity effects of TMS-C-DAS/ATO/ETCS at CAF

The aim of CAF in this section is to be able to compare ATO driving with the driving of different drivers (both juniors and seniors) to get information about how they differ from each other. In order to make this comparison we will need driving data provided by different operators. The larger the amount of data we have, the more dispersion of results there will be and the better comparison of driving and response times (both junior and senior drivers) will be possible. To make the comparison we will also need to make use of a real ATO into which we will input the necessary data to observe what the dispersion is between this situation and that realised by the drivers. All this will be achieved on the basis of diverse and well-labelled data. Diversity will be sought in various segmentations:

- Driver expertise
- Type of line: suburban, regional, medium distance, metro...
- Type of vehicle
- Existing slack in the planning
- Rush hour or off-peak hour

What has been mentioned in the previous paragraphs will allow us to measure, compare and develop result reports based on categories or segments, such as the ones mentioned in the previous paragraph.

All this will be completed by making use of the KPIs obtained in the data analysis explained in the previous section, which will allow us to have more real values of the dispersion that exists between our real ATO and the driving data of the drivers (seniors and juniors).

9.9 Research on capacity effects of TMS-C-DAS/ATO/ETCS at CEIT

CEIT's aim at this section is to develop a new platform using as a basis its existing OPTICON, CET and RANSS software tools to provide a full tool capable of simulating real operation between C-DAS/ATO trackside and C-DAS/ATO onboard. CET tool allows the modelling and accurate simulation of the wireless communication between trackside and onboard systems. Additionally, an accurate navigation modelling software called RANSS will help to simulate navigation performance. The integration of these two tools, together with OPTICON, will provide a helpful tool to analyse the influence of communication and navigation systems on operations, namely energy demand and capacity of the line. The communication and navigation systems deployed might not behave as ideally as it could be considered in traditional simulations, therefore the tool proposed will offer at laboratory more realistic results.

9.10 Conclusions of research on capacity effects of TMS-C-DAS/ATO/ETCS

In this section, the current status of capacity effects research related to system developments has been described, as well as the need for further development of this research.

In all 4 countries, and even within most individual partners, research on all the major themes ETCS, ATO and TMS has been done, but there is always a need for further improvement. The reason is that comparison with real-world data is still limited, so data has to be interpreted and has to be compared with other countries in order to gain insights.

Within 3 out of 4 countries at least a part of the previous and future research is (to be) conducted with RailSys. This is very beneficial to create a common set of specifications for the software developer and makes it easier to adjust simulations and compare results of different case studies. Due to availability of data or methods also sometimes local software tools need to remain in place.

10. Interaction with other work packages and Flagship Projects

There are several other work packages within Europe's Rail that are related to these work packages 8 and on simulation of capacity effects. These related ones will be shortly listed and explained in this section. Several partners of WP8/WP9 also participate in these work packages, so that questions and knowledge can easily be shared and explained by people taking part in both WPs.

FP2, WP17: Next Generation Brake Systems with adhesion management functions – Phase 1: Demonstrator preparation and pre-validation

The capacity effects of brake system developments can be incorporated into the simulation models to study the capacity effects. Based on the planning horizon and implementation chances these can be incorporated in more or less studies of WP8/WP9.

FP2, WP32: DATO Assessment and Potential identification

Capacity effects of DATO will need to be calculated within WP8/WP9 and returned to WP32 so that the business case of the DATO application can be calculated there. These calculations relate to the need described by several partners of WP8 in the previous section to better understand the train driver behaviour for a good comparison with DATO. Another reason is to have a better implementation in simulation tools of braking curves of the train versus the behaviour of the train driver.

FP2, WP37: ETCS HL3 Deployment Strategies

The capacity effects of different strategies of HL3 need to be calculated in WP8/WP9 based on the specifications written in deliverable D37.3. As mentioned in the previous section, some development is needed in the different simulation tools to have an optimal modelling of HL3 that relates to questions from WP37.

11. Conclusions

The aim of this document is to identify the need for future development of methods and models for capacity simulations and feedback loops between planning and operations. This includes capacity simulation of new digital technologies such as ETCS HTD and C-DAS/ATO that will be developed in Europe's rail.

From the mapping of the current simulation tools in Section 6, it can be concluded that there are several tools available with developed functionality for simulation of a transport plan including feedback loops and ETCS L2. However, to be able to simulate ETCS HTD, low adhesion effects (next generation brakes), C-DAS/ATO, driver behaviour and TMS functionality including train path envelope development is needed. Developments are also needed for improved feedback loops including crew scheduling and large networks. In cases where a requested function can be developed in several different tools, it is recommended to prioritise the development of the tools that are used by several partners to get synergies.

Section 7 aimed at giving a general guideline for the processes of verification, calibration and validation of simulation models. During verification the output of the simulation model is compared with the specification of building or improving the model. The calibration process considers parameter settings and defines delay distributions for a simulation model by using realization data. During validation the simulation output is analysed to see if they are accurate enough given the objectives of the simulation study. First, a general methodology for verification, calibration and validation was derived using a literature review. Afterwards, practical examples of verification, calibration and validation were given of different IMs, RUs and research institutes. Current research shows that a combined methodology for the verification, calibration and validation of railway simulation models is missing. Future research should focus to integrate these three topics together including the application of them on practical examples.

Section 8 aimed at giving an overview of feedback loops between operations and planning, and to identify development needs. Feedback loops between operations and planning are essential to improve railway planning. The quality of the plan can be measured and input can be given to the planning when the operational or simulated performance is evaluated. Timetable analysis and simulation can give useful results as a complement to operational performance, to improve the planning. In addition, feedback from operations to simulation is needed for improving methods and models. In order to achieve better simulation models, data improvement by continuous feedback of historical information available for analysis is needed and this will increase the reliability of the railway network simulation. There are also other areas where development is needed, typically related to specific cases where there are missing functionalities today, for example, feedback loop in large network simulations and feedback loops between C-DAS/ATO and TMS.

Section 9 aimed at giving an overview of the current status of research on capacity effects of system developments such as ETCS, ATO and TMS. Since the capacity needs and European obligations are comparable most partners have already performed studies on one or more of the listed developments. Data from real world instances is still hardly till not available, so much research is still dependent of major assumptions. By doing research on the same topics and within the same modelling environments the comparison of projects and results is easier. Furthermore, it will be possible to generate a common list of requirements for the software developer. The major fields for development are better understanding and modelling of the train driver behaviour amongst others for comparison with ATO, the implementation of TMS within software that has only limited support for that yet and a better support for modelling ETCS HTD.

The results from this report will be used in the future work with the developments for improved feedback loops and capacity simulations of ETCS and C-DAS/ATO. The results from this work package will also be given back to several work packages of FP1 and FP2 where the several techniques (HTD, TMS, etc.) are developed or business cases about them are calculated.

12. References

Andersson, E.V., Peterson, A., Törnquist Krasemann, J. *Reduced railway traffic delays using a MILP approach to increase robustness in critical points,* 2015, Journal of Rail Transport Planning & Management, Volume 5, Issue 3, pp. 110–127.

Bešinović, N., Quaglietta, E., Goverde, R.M.P. *A simulation-based optimization approach for the calibration of dynamic train speed profiles,* 2013, Journal of Rail Transport Planning & Management, 3(4), 126-136.

CAF. *CAF Signalling, success in the ATO test on ETCS in NS of Netherlands*, 2019. https://www.cafsignalling.com/en/caf-signalling-success-in-the-ato-test-on-etcs-in-ns-of-netherlands/

Chéron, C. *Rail Strategic Research & Innovation Agenda* (Rapport ERRAC), 2020, https://errac.org/wp-content/uploads/2020/12/RAIL-Strategic-Research-and-Innovation-Agenda-2020-_FINAL_dec2020.pdf

Cunillera, A., Bešinović, N., Lentink, R.M., Oort, N. v., Goverde, R.M.P. *A Literature Review on Train Motion Model Calibration,* 2023, IEEE Transactions on Intelligent Transportation Systems, 24(4), 3660-3677.

De Fabris S., Medeossi G., Montanaro G. *Trenissimo: Improving the microscopic simulation of railway networks*, 2018, WIT Trans. Built Environ., 181, pp. 199–211.

ERA. *System Requirements Specification, Chapter 3, Principles. SUBSET-026-3*, 2016, https://www.era.europa.eu/era-folder/set-specifications-3-etcs-b3-r2-gsm-r-b1

ERA. *ERA Braking curve simulation tool 4.2*, 2018, https://www.era.europa.eu/content/braking-curves

ERA. *Control Command and Signalling TSI*, 2023, https://www.era.europa.eu/domains/technical-specifications-interoperability/control-command-and-signalling-tsi_en

Fischetti, M., Salvagnin, D., Zanette, A. *Fast approaches to improve the robustness of a railway timetable*, 2009, Transportation Science, Volume 43, Issue 3, pp. 321–335.

FR8RAIL II. *Deliverable D3.4 Demonstrator on improved planning - verification of demonstrators*, 2021. FR82-WP03-D-TRV-008-01_-_D3.4_Demonstrator_on_Improved_Planning_Verification_of_Demonstrators.pdf

Goverde, R.M.P., Bešinovic, N., Binder, A., Cacchiani, V., Quaglietta, E., Roberti, R., Toth, P. *A three-level framework for performance-based railway timetabling*, 2016, Transportation Research Part C: Emerging Technologies, 67, pp. 62–83.

Hansen I. A., Pachl J. *Railway timetabling and operations, 2014,* Eurailpress, Hamburg, Germany.

Högdahl, J. *Improving timetable robustness and punctuality of railway traffic: A combined simulation-optimization approach for nonperiodic timetabling on double-track lines,* 2022, PhD Thesis, KTH Royal Institute of Technology. https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-320765

Högdahl, J., Bohlin, M., Fröidh, O. *A combined simulation-optimization approach for minimizing travel time and delays in railway timetables*, 2019, Transportation Research Part B: Methodological, Volume 126, pp. 192-212.

Högdahl, J., Bohlin, M. *A combined simulation-optimization approach for robust timetabling on main railway lines*, 2022, Transportation Science, Volume 57, Issue 1, pp. 52-81.

Högdahl, J., Bohlin, M. *Maximizing railway punctuality: A microsimulation evaluation of robust timetabling methods*, 2023, Presented at the 10th International Conference on Railway Operations Modelling and Analysis (ICROMA), RailBelgrade2023.

Hornung, G.A.M. *Hourly Traffic Density based Decomposition in Non-Centralized Railway Traffic management*, 2023, Delft.

Janecek, D., Weymann, F. *LUKS – Analysis of lines and junctions*, 2010, In: 12th World Conference of Transport Research. Lisbon, Portugal.

Jansen, J. *ERTMS/ETCS Hybrid Level 3: A simulation-based impact assessment for the Dutch railway network*, 2019, TU Delft.

Jansson, E. *Automatic train operation i Sverige - En förstudie [Automatic train operation in Sweden - A prestudy]*, 2020, http://www.diva-portal.org/smash/record.jsf?pid=diva2:1516663

Johansson, I., Palmqvist, C.-W., Sipilä, H., Warg, J. Bohlin, M. *Microscopic and macroscopic simulation of early freight train departures*, 2022, Journal of Rail Transport Planning and Management, Volume 21, Article ID 100295.

Kircher, K., Rosberg, T., Thorslund, B., Ahlström, C., Prytz, E., Bernheim, L., Olsson, N., Moertl, P. *Train driver attention is influenced by the type of railway signalling system,* 2022, Proceedings of the 8th international conference on driver distraction and inattention.

Kuipers, R.A., Palmqvist, C.-W. *Time-varying dwell times for commuter trains*, 2022, Presented at the 9th Transport Research Arena TRA Lisbon 2022, Portugal.

Law, A. M. *Simulation Modeling & Analysis,* 2015, New York, NY, USA: McGraw-Hill Education.

Lindbergh, J. *ATO on Citybanan to increase capacity and punctuality*, 2022, Master thesis at KTH Royal Institute of Technology. https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-315335

Lindfeldt, A. *Validation of a simulation model for capacity evaluation of double-track railway lines*, 2015, In: Proceedings of the 6th International Seminar on Railway Operations Modelling and Analyses (RailTokyo2015), Tokyo, Japan.

Markewicz, B. *A rail model calibration methodology*, 2013, Paper presented at the 36th Australasian Transport Research Forum (ATRF), 2013, Brisbane, Queensland, Australia.
Medeossi, G., De Fabris, S. *Simulation of Rail Operations,* 2018, In R. Borndörfer, T. Klug, L. Lamorgese, C. Mannino, M. Reuther, T. Schlechte (Eds.), Handbook of Optimization in the Railway Industry. Cham, Switzerland: Springer.

Meijeren, J., Van de Lande, P. *ATO, Gamechanger voor de toekomst van het spoor; position paper Automatic Train Operation., 2020, (*Rapport TNO, in Dutch).
https://publications.tno.nl/publication/34637578/LC3k3H/TNO-2020-ato.pdf

Minbashi, N., Sipilä, H., Palmqvist, C.-W., Bohlin, M., Kordnejad, B. *Machine learning-assisted macro simulation for yard arrival prediction,* 2023, Journal of Rail Transport Planning and Management, Volume 25, March 2023, Article 100368.

Nash, A., Huerlimann, D. *Railroad simulation using OpenTrack*, 2004, Computers in Railways IX, Vol. 74, pp. 45–54.

Olsson, N., Lidestam, B., Thorslund, B. *Effect of Train-Driving Simulator Practice in the European Rail Traffic Management System: An Experimental Study*, 2022, Transportation Research Record: Journal of the Transportation Research Board, 036119812211358.

ON-TIME Optimal Networks for Train Integration Management across Europe, *Task 3.1 Assessment of State-of-Art of Train Timetabling*, 2013, ONT-WP03-I-EPF-008-02

Pachl, J. *Railway Operation and Control*, 2009, VTD Rail Publishing, Mountlake Terrace, WA, USA.

Pachl, J. *Timetable design principles*, 2014, in: Hansen, I. A., J. Pachl, eds., Railway Timetabling & Operations, Eurailpress, Hamburg, Germany, pp. 13–46.

Palmqvist, C.-W. *Delays and Timetabling for Passenger Trains,* 2019, Doctoral thesis. Lund: Lund University.

Palmqvist, C.-W., Johansson, I., Sipilä, H. *A Method to Separate Primary and Secondary Train Delays in Past and Future Timetables Using Macroscopic Simulation*, 2023, Transportation Research Interdisciplinary Perspectives. Volume 17, January 2023, article 100747.

Palmqvist, C.-W., Lind, A., Ahlqvist, V. *How and Why Freight Trains Deviate from the Timetable: Evidence from Sweden*, 2022, IEEE Open Journal of Intelligent Transportation Systems. Volume 3, pp. 210-221.

Palmqvist, C.-W., Olsson, N.O.E., Winslott, Hiselius, L. *An Empirical Study of Timetable Strategies* *and Their Effects on Punctuality*, 2017, presented at the 7th International Conference on Railway Operations Modelling and Analysis (ICROMA) – RailLille in Lille, France.

Palmqvist, C.-W., Olsson, N.O.E., Winslott, Hiselius, L. *The Planners' Perspective on Train Timetable Errors in Sweden*, 2018, *Journal of Advanced Transportation*.

Palmqvist, C.-W., Tomii, N., Ochiai, Y. *Explaining dwell time delays with passenger counts for some commuter trains in Stockholm and Tokyo*, 2020, Journal of Rail Transport Planning and Management, Volume 14.

PLASA. *Deliverable D2.1 Summary of state-of-the-art in simulation*, 2017.

PLASA-2. *Deliverable 3.3 Smart planning: Approaches for simulation with incomplete data,* 2021.

Radtke, A., Hauptmann, D. *Automated planning of timetables in large railway networks using a microscopic data basis and railway simulation techniques*, 2004, Computers in Railways IX, Vol. 74, pp. 615–625.

Rosberg, T., Thorslund, B. *Radio communication-based method for analysis of train driving in an ERTMS signaling environment*, 2022, European Transport Research Review, 14, Article 18.

Roungas, B., Meijer, S., Verbraeck, A. *A Framework for Optimizing Simulation Model Validation & Verification*, 2018, International Journal on Advances in Systems and Measurements, 11(1 & 2), 137-152.

Scheepmaker, G.M., Goverde, R.M.P. *Multi-objective railway timetabling including energy-efficient train trajectory optimization*, 2021, European Journal of Transport and Infrastructure Research, 21(4), 1–42.

Scheepmaker, G.M., Goverde, RM.P., Kroon, L.G. *Review of energy-efficient train control and timetabling*, 2017, European Journal of Operational Research, 257 (2), pp. 355–376.

Scheepmaker, G.M., Willeboordse, H.Y., Hoogenraad, J.H., Luijt, R.S., Goverde, R.M.P. *Comparing train driving strategies on multiple key performance indicators*, 2020, Journal of Rail Transport Planning & Management, 13:100163.

Solinen, E., Palmqvist, C.-W. *Development of new railway timetabling rules for increased robustness*, 2023, Transport Policy. Volume 133, pp. 198-208.

Su, S., Tian, Z., Goverde, R.M.P. *Energy-Efficient Train Operation. A System Approach for Railway Networks*, 2023, Springer, Cham, Switzerland.

Tiong, K.Y., Ma, Z., Palmqvist, C.-W. *A review of data-driven approaches to predict train delays*, 2023, Transportation Research Part C: Emerging Technologies, 148, 104027.

Trafikverket. *Omtag tilldelning Värmland. (New guidelines for traffic planning in Värmland, in Swedish)*, 2016, TRV 2017/127, Publication from the Swedish Transport Administration.

Trafikverket. *Nya konstruktionsregler på Bergslagsbanan Borlänge - Frövi. (New planning rules for the Bergslagen line, in Swedish*), 2018, TRV 2018/097, Publication from the Swedish Transport Administration.

Trafikverket. *Handledning för analyser med ERTMS i tågtrafiksimuleringsverktyget RailSys på Trafikverket*, 2022, TRV 2022/83045, Publication from the Swedish Transport Administration.

UIC. *Code 405 OR – Links between Railway Infrastructure Capacity and the Quality of Operation*s, 1996, Paris, France

Vergroesen, R. *ERTMS/ETCS Hybrid Level 3 and ATO: A simulation-based capacity impact study for the Dutch railway network*, 2020, TU Delft.

Watson, R., Medeossi, G. *Simulation*, 2014, In I. A. Hansen & J. Pachl (Eds.), Railway Timetabling & Operations: analysis, modelling, optimisation, simulation, performance evaluation. Hamburg, Germany: Eurailpress.

Zinser, M., Betz, T., Becker, M., Geilke, M., Terschlüsen, C., Kaluza, A., Johansson, I., Warg, J. *PRISM: A macroscopic Monte Carlo railway simulation*, 2019, In: Proceedings of 12th World Congress on Railway Research (WCRR) 2019. Tokyo, Japan.