
# The influence of railway signalling characteristics on resilience

**Rogier Jakob Simons**

**4732839**

***Master thesis - Transport, Infrastructure and Logistics***

***Supervisors:***
**Prof. dr. R.M.P. Goverde (TU Delft, Chair)**
**Dr. ir. E. Quaglietta (TU Delft)**
**Dr. W.W. Veeneman (TU Delft)**
**Ing. R. Koops (Arcadis)**

***To be defended on:***
**25-11-2019**

# Summary

Currently, the railway signalling system ETCS is being implemented Europe wide. This is expected to bring benefits with respect to safety, capacity and interoperability. Resilience is another factor that is likely to be increased with the coming of ETCS. The design of the signalling system affects the resilience, next to the available infrastructure, the timetable and the quality of the contingency plans. This research focuses on the resilience of railway signalling characteristics and its impact on the design and implementation of future ETCS projects.

The aim of this research is multiple. Firstly, it aims to set up a method to test and quantify the resilience of railway signalling characteristics. Subsequently, it aims at comparing the current Dutch signalling system, NS'54/ATB, to the new standard for signalling in the Netherlands, ETCS Level 2. Secondly, it aims at finding the stakeholders in the railway sector that are interested in increased resilience, and those that have the power to increase the resilience of the railway system. And lastly, it aims at creating a method which can support the decision-making around and design of ETCS projects.

To accomplish these aims, several distinctive methods have been applied. Literature review has been used (1) to identify the proper performance indicators for resilience, (2) to find the relevant signalling characteristics with respect to resilience, (3) to see the role of simulation in decision making, and (4) to find common disruptions and disruption management strategies in the Netherlands. Interviews and expert knowledge have been used (i) to find the role of resilience in the Dutch ETCS-decision making and designing process, (ii) to create a power-interest diagram with respect to resilience, and (iii) to discuss the set-up of simulation, the disrupted scenarios and the applied dispatching strategies.

The simulation tool Xandra, developed within Arcadis, has been used to simulate seven disrupted scenarios over three signalling configurations. The corridor Utrecht - Den Bosch and the related traffic have been modelled using microscopic, deterministic simulation. The outcome of the simulation, the delay of trains, has been processed using the computing environment Matlab.

The block section lengths and in-cab signalling vs line-side signalling have shown to be the most relevant signalling characteristics related to resilience. Using the ETCS L2 in-cab signalling has reduced on average 10% of the delay in the seven scenarios, compared to the Dutch NS'54/ATB signalling system. By implementing minimal block sections lengths in combination with ETCS L2, 20% of the delay has been saved. This increase in resilience will be most interesting for operators and passenger organizations as they benefit from it; the power, on the other hand, lies in the hands of ProRail and engineering firms, as they are the ones responsible for the design of the lay-out of the signalling and infrastructure. This creates a tension field. There are three levels in the design phase of ETCS where resilience can be of support: full comparison between alternatives on all factors, including resilience; comparison between alternatives, but only on resilience; to find the relation between alternative designs, based on expert judgement of the results of a resilience study.

# Contents

# Chapter

# 1. Introduction

Various legacy train protection systems across Europa, such as the Dutch NS'54/ATB-EG, the German Indusi/PZB and the Belgian/French Crocodile, have been developed and installed in the mid of the previous century and stem from a time with far less trains. And although some of these systems still have a high capacity standard, they will need to be replaced in the coming decades as their technology is ageing or outdated.

The new European standard for railway signalling is the specification ERTMS (European Rail Traffic Management System). It aims to enhance safety, increase efficiency of train transports and enhance cross-border interoperability of rail transport in Europe (ProRail, 2019a). This is done by replacing national legacy signalling equipment and operational procedures with a single new Europe-wide standard for train control and command systems, specified as ERTMS.

In ERTMS, the trains as well as the infrastructure need to be equipped, the both of which communicate by means of balises in the tracks and/or radio-communication. The train control component of ERTMS is ETCS, which translates lineside information to the driver's cab, where the movement of the train is supervised. There are several so-called 'levels' of ETCS, in which the technique/intelligence predominantly shifts from the track to the trains. Goverde et al. (2013) provides a clear overview of the differences between the levels. A quick overview of the main difference between the working of NS'54/ATB and level 2 and 3 of ETCS can be seen in Figure 1. Further on, these levels will be explained in more detail.

In the Netherlands, the transition from NS'54/ATB to ETCS is a present-day topic of discussion. A number of lines have already been solely equipped with ETCS level 2 in the Netherlands, such as the high speed line between Schiphol and Belgium, the Betuweroute. On the Hanzelijn and the line between Utrecht and Amsterdam, ETCS level 2 has been installed on top of the Dutch legacy protection system. The parliament has recently come to terms that in 2030, several international freight corridors and main lines are to be equipped with ETCS and that in 2050, the whole of the Netherlands should be equipped (ProRail, 2019a). This decision, however, is still flexible and may change due to new developments or insights.

Many aspects have been considered in this decision making process of where and when to replace the current NS'54/ATB with ETCS, such as the cost and available budget, the probable capacity and safety increase, and the ageing of current infrastructure. Capacity increase is thus one part of the whole decision making process. If it can be proven that capacity can be increased significantly at reasonable costs, ETCS may be implemented sooner and/or on more corridors.

ETCS has the potential to reduce headways and hence increase capacity, but in normal operations

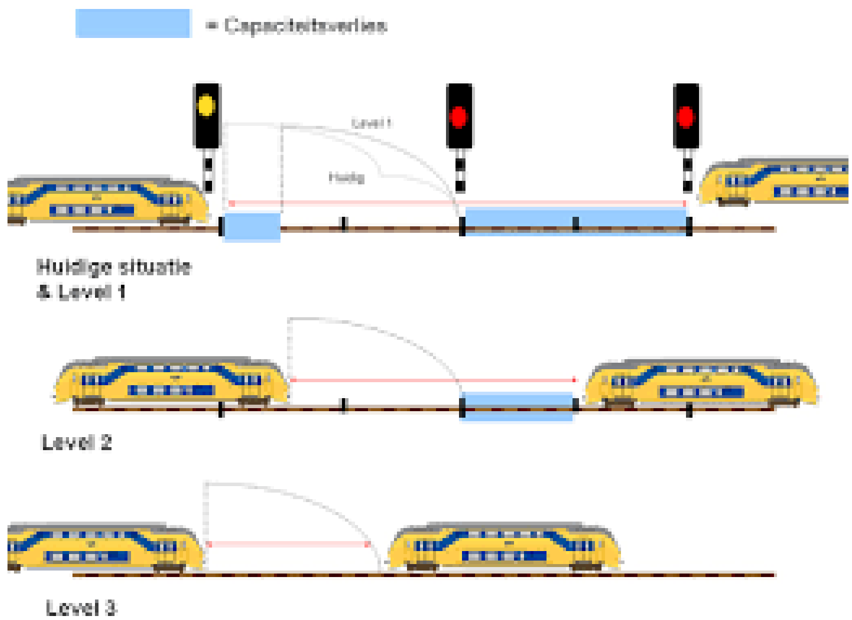



Figure 1: Difference in working between ATB and ERTMS (Kivi.nl, 2019)

only significantly if track circuit arrangements and block boundaries are redesigned optimally, rather than simply being ported over from the previous conventional schemes. Yet, it has shown that the increase in capacity during planned operations with ETCS level 2 is not as much as hoped in the beginning compared to NS'54/ATB with optimized signal distances (Barter, 2008). This can be explained as the capacity is moreover determined by junctions, the homogeneity of speeds and station capabilities, things for which ETCS can only do little, by minimizing or optimizing block lengths around junctions and stations. The lost capacity due to heterogeneous train traffic may not change that much with ETCS, maybe leading to one or two extra train paths per hour. The benefit of improving headways is mostly felt when trains of the same speed and stopping pattern follow each other, so-called branching of train services.

Yet, the potential capacity increase of ETCS may not only manifest itself during normal traffic conditions, but also in disrupted and disturbed conditions. As trains can follow a more optimal speed pattern and braking curve better, ETCS may lead to a faster or better handling of the train traffic in these situations than legacy protection systems such as NS'54/ATB-EG. Furthermore, ETCS-equipped trains have a direct communication link to the radio block centre and thus more precise measures may be taken by the traffic control centre to solve disruptions. It is likely that during disruptions, such as a broken vehicle or a track failure, ETCS will help reducing the delays and that the initial schedule can be resumed more quickly than legacy protection systems.

Goverde et al. (2013) have already evaluated the capacity increase of ETCS during small disturbances and have shown that there is a considerable gain in terms of reduction of infrastructure occupation and increase of punctuality for ETCS compared to NS'54/ATB. Their evaluation referred to the stability of the railway system under different signalling systems and traffic management. However, so far no literature has been found on the quantitative benefits of ETCS over legacy protection systems under disrupted conditions, where disruption management is needed. This then refers to the resilience of the railway system, which is how quickly and smoothly a system can recover from a disruption to the system, with interference of the traffic dispatcher. It is likely that during disruptions, such as a broken vehicle or a collision, ETCS will help reducing the delays and that the initial schedule can be resumed more quickly than legacy protection systems. The question then remains, how much more resilient is ETCS, one of the answers to be sought in this research.

# Chapter

# 2. Methodology

In this chapter, the methodology of this research will be explained. Firstly, by stating the goal and motivation behind the research. Secondly, by formulating the research (sub-)questions to be answered. Consecutively, by clarifying the approach taken. And lastly, by describing how all the data is collected and how the output data is analyzed.

This research aims at identifying the influence of signalling system characteristics on railway resilience and to set up a method of how railway resilience can be tested, focusing mostly on the characteristics of the Dutch legacy system NS'54/ATB-EG and ETCS L2.

## 2.1. Research questions

From the previously mentioned objective flows the following main research question:

***What is the influence of railway signalling system characteristics on resilience?***

This question can be split into several components which all consider a part of the main question. The main research question is therefore subject to the following sub-questions:

- *How does resilience influence the implementation and design of ERTMS in the Netherlands?*
- *Which signalling system characteristics have effect on resilience?*
- *What difference in resilience does ETCS level 2 have compared to NS'54/ATB-EG?*

## 2.2. Methodological approach

To answers the main research question, several steps will be considered as shown in Figure 2, which follow the order of the sub-questions.

Firstly, a literature study will be carried out to see which signalling characteristics could influence resilience, what types of disruptions are frequently occurring, and what traffic management strategies are commonly adapted during these disruptions. Another literature study is performed to find the role of simulation and scientific research in infrastructural decision making processes, and to see how a research like this may fit in that process for ERTMS.

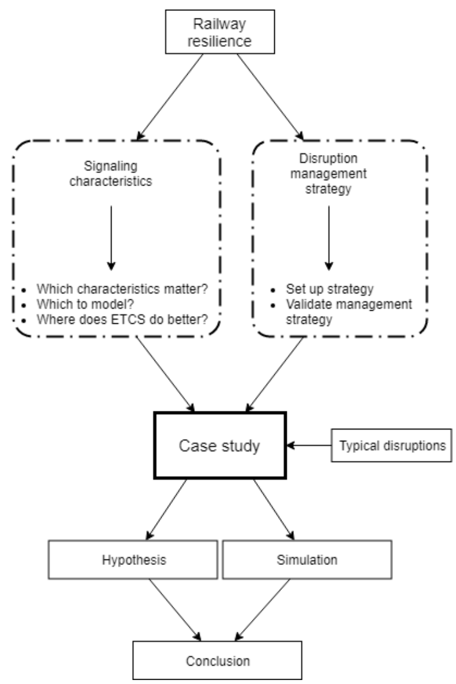



Figure 2: Overview of the methodology

The literature on the signalling characteristics, disruptions and traffic management is then combined into several case studies. The approach is to simulate various disruption scenarios over a case study corridor where multiple signalling system configurations can be implemented and where other factors, such as timetable and rolling stock, are to be controlled for, so that the impact of the signalling system can be found. These cases are simulated to quantify their resilience and to test the hypotheses that ETCS is likely to be more resilient than the Dutch legacy protection system. Simulation is used as this allows assessment of the performances of the rail network repeatedly under different signalling configurations, whilst controlling for all other factors.

To find the true impact of the safety system, different signalling technologies and layouts will be assessed repeatedly over several scenarios, for given characteristics of the timetable, rolling stock, infrastructure and traffic dispatching. The current configuration of NS'54/ATB is tested, as well as two configuration of ETCS: an Level 2 which uses the existing blocks of the current signalling and an Level 2 with reduced blocks, which is approximation of Level 3.

To find the the potential impact of this research on the decision process around ERTMS in the Netherlands, firstly an interview is held with an expert on decision making in the railway sector, and secondly a power-interest diagram is put up to understand the broader environment of this research and to find the stakeholder for whom this research is relevant, following the initial work of Freeman & Reed (1983). This power-interest diagram has been validated using the expert knowledge of two Arcadis employees who have a clear perception of the railway sector as a whole.

## 2.3. Methods of data collection

The corridor considered for simulation experiments is Utrecht - Den Bosch in the Netherlands, as this traffic on this corridor is quite heterogeneous, i.e. intercity as well as sprinters and freight trains, and there are several spatial dispatching possibilities.

Four types of data are needed as input for the simulation: (i) infrastructure data, (ii) interlocking data, (iii) rolling stock characteristics and (iv) scheduled operation data (Goverde, 2018). The level of detail of the network in the simulation needs to be microscopic as this provides accurate modelling of the signalling and the interlocking system, which is the level of detail this research focuses on. Microscopic means that the infrastructure is represented in detail and that the modelling of signalling and interlocking systems is accurate. This opposed to macroscopic models, which are an abstract representation and do not contain details about track characteristics and signalling, and mesoscopic modelling, which is an approximate model of the signalling and infrastructure. Deterministic input data will be used to remove statistical bias from scenario analyses.

Data for the simulation is provided by the consulting company Arcadis. The type of rolling stock that is used on the corridor is open information shared by the NS and can be found on many open platforms. However, more detailed information about the rolling stock, such as the length, mass, and acceleration and braking characteristics, is familiar within Arcadis due to previous projects with this rolling stock.
The microscopic infrastructure and signalling model in Xandra has been built up from scratch and has been set up manually. Information about the infrastructure and interlocking, such as the number of tracks, the placement of switches, signals, speed indication signs and stops, is imported from the track and yard drawings ("OBE-bladen"). The aspects that signals can show and how those signals aspects are enforced by other signals, e.g. a stop aspect requiring a danger aspect on the preceding signal and that signal enforcing a yellow '8' on its preceding signal, is imported from the "OS-bladen" (overview signal aspects drawings).
Detailed information about scheduled operations of the passenger services can be found on the forum *somda.nl*, which gives the stopping and transit times of all services at each station, accurate to the minute. This will be used to precisely reconstruct the current timetable.

The experiments, or more specifically the disruption scenarios, have been designed using statistics on the type of disruptions most frequently occurring on the corridor Den Bosch - Utrecht, and their average length. Seven scenarios distinctive in cause, severeness and length, have been created. For each of these scenarios a specific location and duration has been chosen and the expected differences between signalling system configurations are highlighted. After an initial simulation of a disrupted scenario, an appropriate dispatching strategy is devised and implemented for that scenario. The same dispatching strategy is then the used in all signalling configurations.

In reality, operators and ProRail have developed predetermined contingency plans for several types of disruptions, which are applied when it has become clear what the type and severeness of the disruption is. These contingency plans have not been consulted for the disruptions in the model. For the appropriate dispatching measures in the model, expert knowledge by a former dispatcher has been taken into account. In the model outcomes, it can be seen that trains are standing still for some time when the disruptions happen. This is to imitate the time it takes to find the correct contingency plan and implement it.

## 2.4. Method of analysis

For the simulation, the software program Xandra is used due to the interest of Arcadis in this research. This program is developed internally by Arcadis and approved by ProRail for simulations.

There are several other methods that, in theory, could have been used to answer the research questions, for example a mathematical description of the situation, real-life tests, or a data analysis from real-life disruptions. Describing this problem only mathematically would not suffice, as there is the need to capture patterns and effects which are hardly describable by pure mathematical models. Testing the resilience in reality would be nearly impossible without great disturbance to passengers, would be too costly, and/or would not allow to control all other factors than the signalling characteristics. Gathering data from real-life situations from past disruptions under different signalling systems would require an extensive amount of disruptions to be able to find the effect of the characteristics of the signalling system, since several dozen other (country dependent) factors have influence on the impact of a disruption.

There are a few limitations to Xandra which have been of influence to the scope of this research. Firstly, traffic management is not automated, thus all dispatching measures had to be done on own judgement and be implemented manually. The dispatcher measures may therefore not be as optimal as possible. Secondly, the performance indicators relevant to this research were not calculated by the program, so this had to be calculated separately. And lastly, the behaviour of drivers is not fully accurate.

The Matlab online computing environment has been used to import, analyze and visualize the data which was exported from the simulation program. The data exported from this program consisted of the passing time of each train at each station or timetable point. From this, the relevant performance indicators to measure resilience, as explained in Section 3.3.1, can be calculated. Graphical comparisons are added as these allow for a better recognition of patterns and differences in the data.

The timetable of each train at each timetable point has been calibrated to the second, so that there are no early arrivals. These could count as 'negative delays' and now every small delay can be measured.

# Chapter

# 3. Literature research

First in Section 3.1, the many definitions of resilience will be explored to understand all the aspects of resilience and to determine the right definition to use in this research. Then in Section 3.2, all previous work related to the research question is examined, to find the research gap to be filled. Thirdly in Section 3.3, literature is searched for how one can evaluate resilience in railway. Subsequently in Section 3.4, all railway signalling characteristics related to resilience are analyzed. And lastly, in Section 3.6, the various types of disruptions that can happen on the railway are researched.

## 3.1. Resilience

In literature, many definitions can be found regarding resilience. A few of these will be discussed to get a better overview of what resilience is in general, how its defined in the railway sector and which definition is to be used in this research. Moreover, resilience is closely related to stability and robustness. The differences and similarities between these terms will also be discussed.

Firstly, Baroud et al. (2014) established a paradigm to explain the resilience in networks. This is shown in Figure 3, which shows three distinct states in which a system can operate: (i) its original, as-planned state, (ii) its disrupted state, caused by a disruption to the system and (iii) its recovered state that results from a recovery effort. Resilience is then defined in a general way as the time-dependent ratio of recovery over loss.

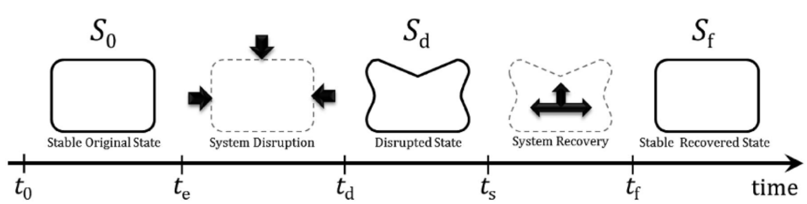



Figure 3: System state transition with time (Baroud et al., 2014)

The *National Academy of Sciences* have defined resilience as "the ability to prepare and plan for absorption of, recovering from, and more successfully adaption to adverse events" (Fisher, 2015). A slightly different definition is provided byAven (2010) who define resilience as the ability of a system to mitigate the consequences of any perturbation on the system. This definition only reflects the ability of the system to mitigate consequences, and does not focus on the effort that the system should made in order to recover an acceptable state. Yet another definition is one by Haimes

(2009). Haimes defines resilience as "the ability of a system to withstand a major disruption within acceptable degradation parameters and to recover within an acceptable time and composite costs and risks". As also concluded by Chen & Miller-Hooks (2012), from all previously mentioned definitions it is clear that resilience is an indicator of the recovery capability of a system, with (predetermined) disruption management.

When looking into resilience in the railway sector, the European ON-TIME project has defined timetable resilience as 'the flexibility of a timetable to prevent or reduce secondary delays using rescheduling' (ON-TIME, 2014). This is underpinned by Corman et al. (2018) as the ability of a timetable to absorb delays if some rescheduling actions is taken. It is assumed that this is not just a valid definition for a timetable but also for a signalling system.

Hence, the definition of resilience to be used in this research is that *resilience* is *the ability of a system or timetable to recover from a disruption given a predetermined control management plan.*

From this it follows that resilience thus consists of two parts: the characteristics of the system itself and the disruption management strategy of the traffic control. The role of both parts on railway resilience is explained in the Sections 3.4 and 3.6.3.

### Differences between resilience, robustness and stability

Goverde & Hansen (2013) provide a differentiation between the terms resilience, robustness and stability. Stability refers to the possibility of a timetable to absorb initial and primary delays so that delayed trains return to their scheduled train paths. Robustness refers to the possibility of a timetable to cope with process time deviations due to design errors, parameter variations and changed operational conditions, excluding any real-time rescheduling actions, thus the timetable order is considered to be fixed. Resilience is defined as the flexibility of a timetable to reduce secondary delays using rescheduling actions such as re-timing, reordering, rerouting, cancelling trains, etc.

## 3.2. Prior work

One of the studies that this research builds upon is the one from Goverde et al. (2013), who made a capacity assessment of a Dutch corridor of different signalling configurations under both normal and disturbed railway conditions. For the analysis of the disturbed conditions, they used the train dispatching system ROMA, so that an assessment can be presented of the potential gain of next generation signalling and train dispatching systems. Their results showed that in delayed operations, there is a considerable gain for ETCS compared to NS'54/ATB, since the braking distances decrease when delayed trains run at lower speeds, having a stabilizing effect on headway times, delay propagation and throughput.

Another study in this direction is the one from Quaglietta (2014), who investigated the influence of the block layout of a metro line on network performances, weighed against the cost. Shortening the block sections reduces the headway of the line and the energy consumption, but increases the number of block sections and therefore the investments costs. Using a "black-box"optimization algorithm, it was found that for a metro line in Naples (Italy) the investment costs could be strongly decreased whilst only slightly increasing energy consumption, the minimum line headway and the average train delay. The presented simulation-based framework is applicable to any kind of railway networks and can be easily applied for the optimal and robust design of different elements related to both infrastructural (track layout, power supply, station areas, etc.) and operational components (e.g. recovery strategies, service timetable). A simplified version of this framework of iterating e.g. different block lengths over several disruptions until the optimal signalling layout is found, is useful in this research for finding the influence of several signalling characteristics and the optimal signalling configurations regarding resilience. However, unlike that study, this research will not weigh the economical aspects of e.g. increasing the number of blocks into detail.

Dicembre & Ricci (2011) developed a methodology which allows to have an analysis of sensitivity by modifying technical features without modifications of delays distributions. Their findings were that higher values of practical capacity are attainable when reducing block sections and that for a fixed maximum speed, the influence of block sections length is higher for homogeneous train traffic and lower for heterogeneous traffic. The influence of block sections length on resilience can thus be tested more clearly in a case study with homogeneous train traffic than a case with heterogeneous traffic. Another finding was that reduced sub-sections near stations have almost no practical effect on capacity for plans with mixed traffic, as it will mostly reduce headway between trains with stops at these stations. The greater the number of trains without stops, the less is the effect of such reduced sub-sections.

## 3.3. Evaluation of railway resilience

### 3.3.1. Key performance indicators for resilience

When looking at the literature, one can find a large range of articles that describe the evaluation of a railway timetable, from passenger perspective as well as from operator or infrastructure manager perspective. However, none describe which performance indicators are appropriate to measure the intrinsic performance of a railway signalling system.

As the evaluation of a timetable and the signalling system is likely to be closely related, we will look into what timetable performance indicators are proposed in literature for either robustness, stability or resilience.

Goverde & Hansen (2013) have suggested five main indicators of timetable performance:

- Infrastructure occupation: the share of time required to operate trains on a given railway infrastructure according to a given timetable pattern. It can be computed using the timetable compression method, UIC 406;
- Timetable feasibility: the ability of all trains to adhere to their scheduled train paths. A performance measure can be the amount of scheduled train path conflicts.
- Timetable stability: the ability of a timetable to absorb initial and primary delays so that delayed trains return to their scheduled train paths. A measure can be the settling time until

which a delay has been absorbed.

- Timetable robustness: the ability of a timetable to withstand design errors, parameter variations, and changing operational conditions. A measure of timetable robustness could be the percentile of process times that should be realized within the scheduled process time.
- Timetable resilience: the flexibility of a timetable to prevent or reduce secondary delays using dispatching (re-timing, re-ordering, and re-routing). Performance measures for timetable resilience are statistics such as punctuality, average delay, or maximum secondary delay, as well as the actual average track occupation with rescheduling.

Hofman et al. (2006) mention two indicators of performance during disturbances, both from the perspective of the passenger: (i) regularity, the percentage of late departures over the total number of departures, measuring the departure delays, and (ii) reliability, the percentage of actual departures over the scheduled number of departures, measuring the number of cancelled services.

Corman et al. (2014) have deviated between objectives from infrastructure managers and dispatchers, and passenger related objectives. The objectives for the IM and dispatchers are average total delay, average consecutive delay and total travel time spent. For the passengers, the objective was to minimize the waiting time on the first train, the in-vehicle time, and the waiting time for connections.

Vromans et al. (2006) have used the average arrival delays and the observed punctuality as measures for evaluating reliability. Other possible measures they mention are the percentage of realized passengers transfer and the average delays of passengers. Binder et al. (2015) have also looked into the performance of a given timetable, with key indicators that are passenger related. they have used indicators based on travel time, transfer, departure time shift and the number of disrupted passengers.

Nicholson et al. (2015) have used three measures of resilience which allow for an assessment of performance of timetables, control methods and railway operations, which are based on the deviation of the system measure. These measures are (i) the maximum deviation during a time period, which can be seen as the maximum of the sum of the delays at a moment in time, (ii) the time to recover to an acceptable threshold and (iii) the deviation area, or the total sum of delays. The meaning of the three measures is shown in Figure 4.

For the evaluation of train schedules, Goverde et al. (2013) have used the following four indicators: (i) running time, (ii) minimum headway time, (iii) minimum cycle time, and (iv) infrastructure occupation. To evaluate the behaviour of various signalling systems variant they have computed delay statistics and infrastructure occupation using the following indicators: (i) average delay, (ii) maximum consecutive delay, (iii) average consecutive delay, (iv) punctuality, (v) dynamic minimum cycle time, (vi) dynamic infrastructure occupation, and (vii) travel time spent.

As shown, there is enough literature to be found on how to evaluate railway timetables and schedules, and also a few on the evaluation of resilience in railway operations. However, only Goverde et al. (2013) have described how to compare the performance of the railway signalling system itself. The difference is that said research was aimed at investigating the combined effect of optimal train rescheduling for various signalling systems, whilst this research has the aim to use the

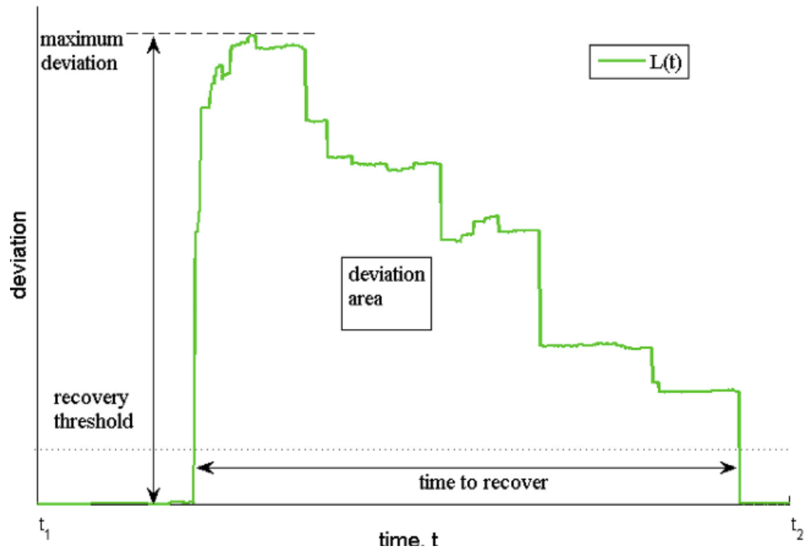



Figure 4: Three key measures for resilience proposed by Nicholson et al. (2015)

same rescheduling rules for the various signalling systems.

On top of the found performance indicators, this research introduces a new indicator for resilience, the Resilience Index , which is the ratio between the total train delay of a given signalling system over all scenarios, and the delay of a benchmark signalling system, multiplied by the number of scenarios.

$$Resilience\ Index = \frac{N}{\sum_{n=1}^{N} \left[ \frac{\sum_{t=1}^{T} Final\ Delay_{t,SignSys}}{\sum_{t=1}^{T} Final\ Delay_{t,RefSys}} \right]_n}$$

where $N$ is the number of scenarios, $T$ is the number of trains, *Final delay* is the delay of a train at its final destination, and *SignSys* is the given signalling system which is compared to the reference signalling system, *RefSys*.

The resilience index of the reference configuration is set to 1 by multiplying the ratio of the delays by the number of scenarios. An index larger than 1 then indicates a more resilient signalling configuration, and an index smaller than 1 a less resilient configuration.

Taking the reciprocal of the resilience index gives the average percentage of delay a configuration can save compared to the base configuration. This delay saving can be used to translate the benefit of a configuration for an economic evaluation and identify the corresponding monetary savings.

### 3.3.2. Evaluation method

The key performance indicators chosen to measure the signalling resilience in this research are a combination of indicators found in Nicholson et al. (2015), namely time until recovery, maximum delay and the sum of the delays, captured together in the deviation area, as well as indicators found in Goverde & Hansen (2013), such as punctuality measured at 3 and 5 minutes, the precision standard of ProRail (ProRail, 2019b).

A design is made in Table 1 of how to evaluate the different protection systems. Per single disruption this table will be made in which the different configurations of the train protections systems will be evaluated on resilience. This needs to be done for miscellaneous disruptions, in order to get a representative and statistically significant comparison. The disruptions combined with simple traffic management strategies, such as short turning of trains, are evaluated for all the protection systems on one or more performance indicators, such as recovery time, the sum of the delays or punctuality.

Table 1: Example of an evaluation table of one specific disruption

| Configuration | 3-min punct | 5-min punct | Recoverytime | $\sum$ Delay | Max delay |
|---|---|---|---|---|---|
| NS'54/ATB | | | | | |
| ETCS L2 existing blocks | | | | | |
| ETCS L2 reduced blocks | | | | | |

## 3.4. Characteristics of signalling systems

Section 3.4.1 describes how railway safety systems are built up, what the functions of its subsystems are, and how NS'54/ATB and (the levels of ETCS) differ on these subsystems. Section 3.4.2 defines which of the previously mentioned subsystems are relevant in relation to resilience.

### 3.4.1. Subsystems of railway signalling

Railway safety systems can be divided into several subsystems, automatic train protection, track-free detection, interlocking and signalling (Goverde et al., 2013). The following paragraphs explain the function of each of these subsystems, and how these subsystems work for different signalling systems. A more detailed explanation of NS'54/ATB and the levels of ETCS, can be found in Appendix 7.2.

**Automatic Train Protection (ATP)** The transmission of information and supervision of the speed and braking curve can either be intermittent or continuous, or a combination of both, where the transmission is intermittent but the supervision continuous Coenraad (2008). Continuous systems have shown to reduce ATP capacity penalties and improve headways. A disadvantage of an intermittent ATP is that a train is being restricted to the braking curve imposed by a distant signal or similar, even after the signal approaching has been cleared. This, however, can be overcome by infill information, so that a train does not have to creep up to the balise at that signal before its ATP information can be updated. This can be helped by early signal, which update the information, or a semi-continuous system near the signal can help overcome this, by almost instantaneously providing an update to the train on changes to the signal aspect. Disadvantages sticking to continuous ATP systems depending on speed codes can be that the worst braking train defines the braking profile, that the brake supervision is coarse due to limited frequencies in track circuits, that bi-directional running requires a switch-over of transmitter and receiver, and that modern traction system can disrupt the transmission.

The way the braking curves is supervised, varies per signalling systems. For example, in NS'54/ATB, the driver has to start braking at the distant signal and continue at 40 km/h up to the red signal. In ETCS, the train calculate its braking curve up to the end of the movement authority and thus postpones the braking until necessary. ETCS can calculate the braking curves based on numerous input parameters to perform its supervision and advisory functions in real-time (ERA, 2016). The input data consist of four categories, namely (i) physical parameters from real-time measurements, such as position, speed and acceleration; (ii) ETCS fixed values, such as driver reaction times; (iii) ETCS track-side data, such as signalling data (target speed/locations) and infrastructure data (slopes); and (iv) On-board parameters before the start of the mission as part of the ETCS train data. The indication of when to brake is thus no longer fixed for all trains, but moreover variable and real-time calculated by each individual train.

**Movement authority communication** Related to the ATP is whether the movement authority is communicated to the driver via line-side signals or via in-cab signalling. Driving with optical signals implies that the location where to start braking is fixed for all trains, while the initial speed is adapted in order to match the braking performance of the train, or in other words, poor braking trains must drive slower to still have the same block length (ERA, 2016). From the point of train separation, the essential benefit of a system with cab signalling compared with a system with line-side signalling is the independence of the cab signals from the approach distance of the line-side signal system, which allows trains to run a higher speeds locally (Hansen & Pachl, 2014).

The block signalling can be divided in three principles (Goverde, 2018): (i) train separation with fixed block distance, as in most signalling systems such as NS'54 and ETCS level 2, (ii) train separation with absolute braking distance as in ETCS level 3, or (iii) train separation with relative braking distance, which is not yet anywhere in use though. The difference between principle i and ii has been captured in Figure 5. In a moving block system, the movement authority of the following train is up to tail of the preceding train, with some safety measures included, whereas in a fixed block system the movement authority is up to the beginning of the block the preceding train is in.

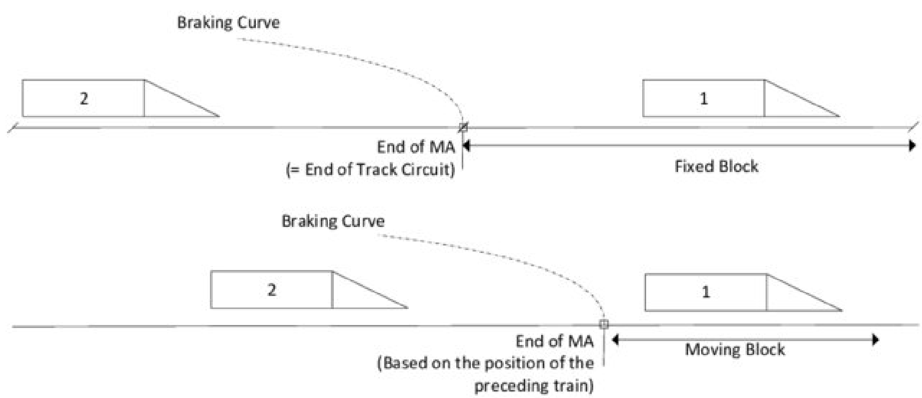



Figure 5: Braking curve difference between fixed and moving block (Ferrari et al., 2012)

**Train detection** On the infrastructural side there are two main methods to detect the presence of a train in a section and separate trains, namely via track circuits and axle counters. Track circuits are based on a electrical circuit using the rails, to detect the presence of a train in a section. Axle counters detect train traffic by counting train axles crossing the border of a section using a counting head. The main advantages of axle counters over track circuits are that the length of the sections are virtually unlimited, that no insulating joints between sections are needed, and generally less installation and maintenance costs (FERSIL, 2019). Which has been implemented is mostly a historical precedent.

It is also possible to separate train safely using only train-based systems, as in ETCS L3. In this signalling system, the train itself send its position, length and integrity status to the interlocking system to release (digital) track sections (ERTMS.be, 2017). This thus requires an extra train integrity monitoring (TIM) system.

A combination of both can also be used. The hybrid ETCS L3 is an example for this, where the train integrity status can also be provided by track-based train detection. This allows to have trains with TIM as well as trains without TIM running at the same tracks (ERTMS.be, 2017).

**Interlocking** Interlocking is a system composed by a set of signal apparatus that prevents trains from conflicting movements through only allowing trains to receive authority to proceed, when routes have been set, locked and detected in safe combinations. Its main function is to set and lock routes related to each train located in an area under its responsibility, in order to ensure safe movements along the track.

The influence of different configurations of the interlocking in NS'54/ATB compared to ETCS Level 2 falls out of the scope of this project, as no report has been found that has treated the interlocking as an input parameter to a capacity analysis.

In Appendix 7.2, one can find how various existing signalling systems work on the previously mentioned characteristics.

### 3.4.2. Signalling characteristics related to resilience

There are several characteristics of the railway signalling systems that have influence on resilience. One of those is the way the movement authority is communicated to the driver, which can be either via line-side signals or via in-cab signalling, and the latter could be discrete (ETCS L1) or continuous (ETCS L2).

Another characteristic is the length of the blocks in the interlocking. This can either be determined by the worst braking train on that line, as in NS'54, or that the blocks are independently set from the trains and the train itself determines its braking distance over one or more blocks, such as in the levels of ETCS. Related is how many aspects the signalling system has; the more aspects, the shorter the blocks can be as the braking distance can be divided over more blocks, and the closer trains can run after each other, such as the 4-aspect signals in the UK.

Benefits of ETCS can be obtained especially in switchover situations, as shown in Figure 6.

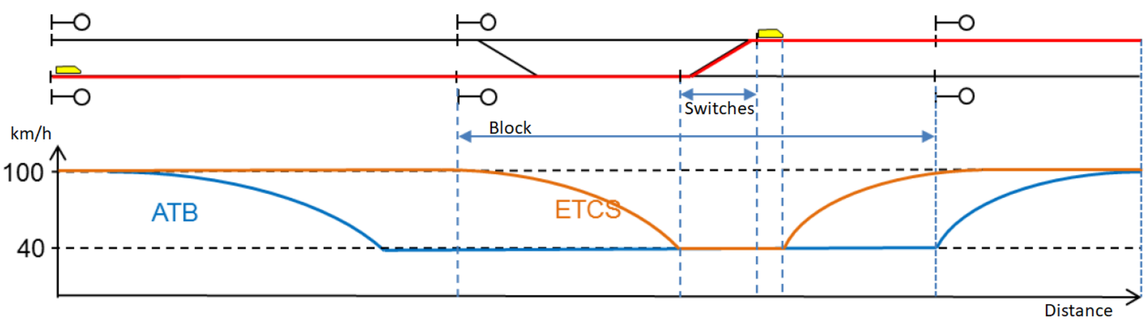



Figure 6: Speed restriction comparison near switches for ATB and ETCS (Goverde, 2012)

An ATB train has to start braking at the approach signal, whereas an ETCS-equipped train will calculate its braking curve right up to the point of danger where the lower speed is required. After the switch, the ETCS-equipped train is able to accelerate earlier, as soon as it clears the switch.

## 3.5. Decision process of implementing and designing ERTMS

This section will investigate the role that simulation and modelling play in decision making, how the decision making process for the implementation of ERTMS is organized in the Netherlands, and how this research fits in that decision making process.

The decision making process of when and where to implement ERTMS is not only a technical process but also a political one with high involvement of the government and parliament, in the Netherlands as well as in other European countries.

This is substantiated by Albrechts (2003), who states that the planning of infrastructural projects has become partly a political choice. This planning is no longer an abstract analytical concept made by engineers but moreover a concrete socio-historical practice. As such, planning is part of politics, and cannot escape politics, but it is still different from politics. Albrechts concludes from a case study that plan-making and political decision-making are dealt with in different arenas and that in both arenas different actors are involved. Both sides do need to know the consequences of their decisions on the other parties. An interview was held with Hugo Thomassen from the ministry to examine the role that the university/engineers play versus the role of decision makers, the findings of which to be found in Appendix 7.4. Mr. Thomassen has had several management functions inside ProRail, Keyrail and NS, and therefore has a broad, nuanced knowledge and overview of the rail sector as a whole.

### 3.5.1. Role of simulation in the transport field

Brömmelstroet & Bertolini (2011) state that the ways of dealing with transport issues in daily urban planning practice are facing several transitions worldwide, from a relatively simple institutional context to a complex one with multiple participating stakeholders, holding multiple values and having multiple conflicting goals. This transition set new requirements on transport knowledge to support planning. The role that models play in the planning process can be very different: instrumental, symbolic and conceptual, although the weight of this role should be seen relative to many other influences on the planners. They conclude that in order to technologically move forward, we need to test solutions in a rigorous way. One way to do this is testing the central heuristics in highly controlled experiments and see if these interventions have the expected results compared to a control group.

Another reason for simulation is that decisions regarding urban transportation investments such as building a new light rail systems or changes in land use policies, have significant and long-term economic, social, and environmental consequences (Borning et al., 2008). Simulation models can help government agencies and citizens make more informed decisions about such issues.

Kavicka & Klima (2000) conclude that the substantial decisions in the field of transport should not be adopted without the modelling of their consequences. They state that due to the complexity of the railway transportation systems and its stochastic behavior, the application of exact mathematical solutions is very limited. An objective tool is needed to make decisions easier for the management. The findings of an objective model can provide the basis for many decisions using output statistical data as well as animated operational processes and events. The simulation of systems is a research method supporting the analysis, design and optimization of real systems in the following three steps:

1. Replacement of the real system by a simulation model
2. Experimentation with the simulation model with the aim to determine its properties, behavior and reactivity to changed conditions
3. Application of the results obtained to real system

However, there is a limit to the truth to reality of the model. Simulation is considered an approximate method, as the model is simulated in an experimental environment and simulation is an experimental method. Furthermore, an invalid simulation model may result with confidence in wrong results. The validation of a simulation can be the one of the pitfalls, as this requires various data, which in itself can be a pitfall. Another pitfall is the interpretation of results from a model, which may only provide valid answers for the context in which it has been designed (Koivisto, 2017). One should thus be careful on how to interpret and use the results of a simulation study in the right way.

### 3.5.2. The ERTMS implementation process

The "Dossier Programmabeslissing" from the ERTMS program elaborately describes the actors involved in the ERTMS decision making process and how this process has been set up. The Dutch House of Representatives has marked the ERTMS program as a 'big project', due to the size, length and cost of the implementation of the ERTMS program. This makes that the secretary of state from the ministry of Infrastructure and Water (I&W) is the client of the program, who has put the program management in the hands of ProRail.

ProRail has had the task, in collaboration with the ministry, NS, and other actors to balance the costs and the benefits and decide which routes will be upgraded and when. The benefits of ERTMS are increased safety, increased interoperability, better use of capacity and speed, and increased robustness. An important part of the costs goes to the investments in ERTMS itself: objects such as trains and the signalling, user training, test facilities and adaptions to the IT.

The current strategy is captured in ten migration steps to reduce the risks. Also, the routes to be adapted have been chosen, based on the costs and benefits. However, the ERTMS program has to have an adaptive nature due to the nature, size and duration of the program. New developments, insights or possibilities could lead to reconsiderations between the replacement need, international obligations and the need for capacity. Mr. Thomassen states that, although several collaborations have been made in recent years, here the role of the university should and can be bigger, in testing the impact on capacity and traffic management of e.g. ETCS L2/HL3 in combination with Automatic Train Operation.

### 3.5.3. The role of this research in the ERTMS program

The influence of ERTMS on capacity is thus only a small part of the decision making process. Furthermore, the "Dossier Programmabeslissing" states that ERTMS has only a limited positive impact on the robustness of the travel time. The lost travel time due to the extra margin in the timetable is 0,2 mln hours less in the situation with ERTMS than in the situation without ERTMS, compared to a standard reduction in travel time of 2,24 mln hours. This 0,2 mln hours is only based on robustness, which is, as described in Section 3.1, the ability of a timetable to cope with small deviations. This research will moreover try to find the benefits of ETCS during disrupted situations, which has been the initial question of Arcadis for this research. A positive outcome of this research may strengthen the case for a faster or wider implementation of ERTMS, although robustness is not one of the most important criteria for this.

However, it was suggested by mr. Thomassen to use a stakeholder based approach to see which stakeholders have most interest in an improved robustness of the train traffic and what their respective power is in the decision making process. This stakeholder may then use this research in its advance.
The ERTMS program has divided the 190 stakeholder organizations into several groups:

- Operators (freight and passengers)
- (Local) governments
- Companies in the harbour
- Passenger organizations (e.g. Rover)
- Infrastructure managers (e.g. ProRail)
- European instances (e.g. ERA)
- Market parties, such as contractors, suppliers, and engineering firms
- ILT (Inspectie Leefomgeving & Transport)

Each of these groups have been placed in a power/interest-diagram, shown in Figure 7, showing

how much power a group has in the decision making process versus its interest in increased robustness of the train traffic.
The parties that have the most power are the instances responsible for the ERTMS program, the ministry and ProRail; they have to oversee the whole program make the decisions. For the ministry is robustness not one of the main criteria, as stated before. ProRail has the most power, as they set out the criteria for the design of future signalling lay-outs and have to approve those designs. Their interest, however, is more aimed at realizing extra capacity, than at improving the resilience. Besides the signalling, ProRail is also responsible for the allocation of the capacity on the tracks, and is thus closely concerned with the timetable of operators.
The suppliers of the ERTMS-products have some power in the form of the products they offer and develop.
The two groups with the highest interest, (passenger/freight) operators and Rover, are parties that represent the interest of the passenger/users, as they in the end are the ones that benefit most from improved robustness. Passenger operators, such as NS and Arriva, also have another interest in an increased robustness: in their concession to drive trains they are scored and evaluated on their punctuality, although their punctuality is likely to decrease during the transition period, something for which they may not be blamed in the beginning (Thomassen, 2019). NS is the main operator in the Netherlands, and therefore later on the main users of ERTMS. This gives them the position to negotiate in the decision making.
Rover has interest on behalf of the passenger. However, they only have limited power in the decision making, so they are limited to giving advice and feedback on plans made by the operators.
The interest of engineering firms in increased resilience small. However, they will be designing the future lay-outs of the signalling system and thus have influence on how the signalling can be optimized with respect to resilience, being it under the supervision of ProRail.
Presumably, the power of the suppliers is higher in the Netherlands than in larger countries. The smaller the market, the larger is their influence on what they develop and sell.

This research may thus be of most importance to NS, ProRail and engineering firms, and not so much to the ministry, as their focus is more on safety and replacing old infrastructure than on robustness. For NS and ProRail, however, it will be interesting to find out the effect of ETCS on the resilience of their timetable and their contingency plans. For engineering firms, the research may be of influence to how they will design future signalling lay-outs, further described in Section 3.5.4.

### 3.5.4. Design process of ETCS

A method to show the relation between resilience and signalling characteristics is investigated in this research. It can help designers and decision makers on several levels, based on the need and the available information. Three levels can be distinguished:

1. Full comparison: fully developed method in which alternatives can be compared between themselves as well as with other factors. The benefits and costs can be expressed such that non-experts are able to include resilience in the decision-making.
2. Comparison between alternatives: results are complex to translate to monetary terms in the

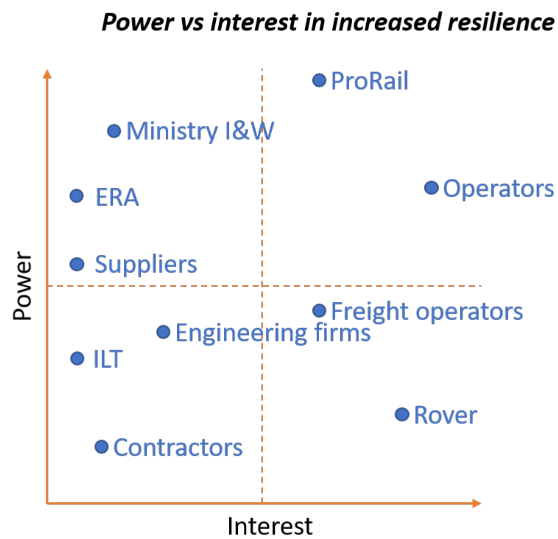



Figure 7: Stakeholder involvement diagram in ERTMS w.r.t. resilience, where 'power' is the influence of a stakeholder in the ERTMS-decision making process, and 'interest' its interest in improved resilience

method and can therefore only be compared among themselves but not with other factors.

3. Expert judgement needed: the method is to be used as a tool to find the relation between alternatives. The results are not convertible to monetary terms and decisions have to be based on expert judgement of the results.

During the design of the lay-out of the infrastructure and signalling of a corridor by an engineering company, many criteria have to be taken into account and have to be traded off. These are criteria such as functionality, costs, planning and risks. The designs of the engineering companies are examined by and under the responsibility of ProRail, who manages and finances the infrastructure. Whilst it is ProRail who pays for adaptions of the infrastructure, it is the operators on that infrastructure who are the ones benefiting from the adaptions. The costs of the adaptions thus do not lie with the parties that will benefit, creating conflicting interests between the developers and the users.

Resilience is not yet a significant part of the trade-off during the design phase of the signalling lay-out. However, conflicting interests do exist when it comes to resilience. Operators would like more infrastructure and contingency options for a more robust and resilient operation. Infrastructure managers, on the other hand, try to limit the number of switches, tracks, sections and signals, to reduce maintenance and the number of failures.

Robustness is usually a less crucial criterion in the design phase. More importantly, the design has to meet the requirements of the main functions, such as driving- and headway times and simultaneities, and moreover is the capacity analysis an important criterion. Resilience and robustness are smaller criteria, but can be used to assess design choices and to underpin the choice for one.
This research aims to provide a method to express the benefits in time savings to evaluate the benefits of resilience. However, it is ambiguous how to normalise the benefits to compare it with

other factors. Not all time savings are worth the same. Five minutes delay at a passengers final destination is not comparable to five minutes delay at a transfer station. The same holds when comparing savings in delay. A quarter of an hour saved during a disruption of two hours may be worth less than a quarter saved during a disruption of half an hour. Other indicators should thus be used as well to weigh designs properly, such as the potential in-/decrease of the appreciation of travellers or number of missed connections.

## 3.6. Disruptions

The purpose of a resilient system is to overcome disruptions. This section will therefore discuss which typical disruptions normally have to be overcome on a railway, and especially on the Dutch network.

A disruption can have a plethora of reasons. Xu et al. (2016) have differentiated between eight sources of disruptions that take place at the Chinese high-speed rail network, a categorization which is likely to be also valid for other railways:

1. Bad weather: snow, rain, fog, etc.
2. Vehicle on-board equipment failure: smoking alarm, failure of train control system, etc.
3. Train body failure: bogie or wheelset failure, etc.
4. Communication equipment failure: GSM-R failure, transponder failure, etc.
5. Track system failure: switch or rail failure, etc.
6. Electric related failure: pantograph breakdown, overhead wire failure, etc.
7. Dispatching human interface failure or alarm
8. Other: e.g. collisions

The severeness and impact of disruptions and the needed disruption management can be classified in two categories according to Blenkers (2015) : partial blockage and full blockage. In the first case, partial train traffic is possible, but balancing of the trains is needed. Examples of this are that from two tracks only one can be used, or that the the speed is restricted. In the full blockade case, no traffic at all is possible across a certain point in the line.

### 3.6.1. Typical Dutch disruptions

To see which disruptions are most common on the Dutch railway network, one can use the statistics of *Rijdendetreinen.nl*, an independent website which collects, stores and analyses all the information of the NS about the disruptions on the whole Dutch network. The NS considers a situation a disruption if no trains are driving on a trajectory or with serious delays, not just when some trains experience delays. The severity of a disruption cause is the combination of the frequency of the disruption and how long it takes for the disruption to be resolved. Explanation on how defects can occur and their impact, can be found on *prorail.nl/storingen*. From this first source it can be seen that the most common disruptions in the Netherlands are the following:

- **Train defects**: meaning a train with a defect, which cannot drive any further due to e.g. an electrical failure or problems with the doors, and which hinders the other train traffic, until the defect is fixed or the train is towed away. Although this disruption happens quite frequently, five times a day on average, it only lasts 50 minutes on average;

- **Signal failures**: a disruption where traffic controller cannot control on or more signals anymore, due to e.g. during an incorrect track occupation, or failing track circuit. During such a defect or disruptions, signals automatically turn to the safest state, the red aspect, thus forcing trains to come to a standstill. In some cases when e.g. only one signal is failing, trains may pass these red signals at 30 km/h with an explicit permission of the dispatcher. If more signals, especially around station areas, are failing, train traffic is completely blocked. This type of disruption happens approximately once a day, with an average length of 2,5 hours;
- **Switch failures**: a situation in which trains cannot cross one or more switches, because it failed. In less severe situations, trains are diverted via other tracks or can use the switch in only one direction, leading to only small delays, but in the worst case a whole yard or corridor is blocked. The length of this disruption is on average 4,5 hours, but it happens quite infrequently, once or twice per month;
- **Collision with a person**: a train having a collision with someone in or near the tracks, either deliberately or by accident. This causes the tracks to be fully blocked for an average of 3,5 hours, on average almost once a day, until the tracks are cleared by the emergency services;
- **Deployment of emergency services**: meaning one (or more) emergency service is sent to a train. E.g. an ambulance when someone in the trains becomes unwell, or the police for irregularities in the train, such as aggression. The length of these disruptions varies between a few minutes to one or two hours, with an average of 50 minutes, once every two days;
- **Repair works**: the infrastructure may have to be repaired unplanned. To ensure safe work conditions, the tracks are taken out of service. Although this type of disruption happens only once every four days, it takes more than 3 hours on average to be resolved;
- **Level crossing failure**: meaning a level crossing is not functioning properly. Due to the fail-safety approach of the railways, all trains have to drive slowly if it cannot be assured that the crossing is safely functioning, leading to delayed or cancelled trains. Depending on the system, several interdependent level crossings may fail simultaneously. This happens on average only once every three days, but can lead to disruptions of on overage 2,25 hours.

This list is not exhaustive, as there are many more sources of disruptions, which will be mentioned but not discussed in detail. Other frequent/severe sources of disruptions defects to the catenary, power failures, collisions with vehicles or animals, delayed works to the tracks, unauthorized persons along the tracks, or weather conditions.

### 3.6.2. Disruptions related to resilience

As this research aims at comparing railway resilience for different signalling systems, only disruptions that are likely to have a significantly different outcome under different signalling systems are included. The first type of disruption which may be interesting to look at are disruptions in which the tracks have been blocked for some time, such as a pulled emergency brake. While recovering the service from standstill, several trains have to follow each other closely, a situation for which a cab-centered, continuous signalling system is beneficial. The second type of disruptions which likely have a different outcome are those in which the speed is restricted someplace on the line. In these situations, a postponed braking curve is helpful, as shown in 6.

The following sources of disruptions have been identified as disruption where different signalling systems are likely to give a different outcome:

- **Level crossing failure**: The failing of a level crossing has the effect that trains have to pass the crossing at a restricted speed until the crossing is safe. This will shorten headways between trains. From safety perspective, a benefit of ETCS L2 is that the location of the speed restriction can be communicated via the movement authority directly to the train and driver instead of instruction via radio communication between the dispatcher and driver, as current practice.
- **Signal failure**: Two types of failures can be distinguished: one where the traffic is completely blocked for some time, and one where the speeds is restricted on one or more block section.
- **Faulty train**: In case of a broken train which is standing still along the corridor or driving at a limited speed, a few other trains will flock behind this train, until the faulty train is able to drive 'normally' again or is towed away. It is also possible that the other trains are directed via other tracks, such as the track in the opposite direction.
- **Switch failure**: A switch failure will lead to some routes that can not be set over that switch, and trains thus have to find other routes over other tracks if possible. This creates higher occupation of the other tracks.

### 3.6.3. Disruption management

The EU ON-TIME project states that resilience requires knowledge of the traffic control measures. The most common dispatching measures are therefore described here:

For the real-time rescheduling of the railway traffic during disruptions, by dispatchers as well as rescheduling tools, three main actions can be taken, according to Hansen (2010): re-timing, reordering and rerouting. In case of a full blockage, some extra actions or decisions are possible for the dispatcher, namely short-turning of services, cancelling services, extra stopping and stop skipping ((Blenkers, 2015) and (Goverde, 2018)). This is done to isolate the impact of the disruption to adjacent areas. In short, all rescheduling measures will be explained:

- Re-timing: By either expediting or delaying a service, route conflicts can be prevented, thus increasing safety and energy efficiency and possibly even reducing travel times;
- Reordering: Services can be swapped to allow overtaking of (delayed) slower services by faster ones, so that the delay of the faster service is minimized;
- Rerouting: By giving a service a different route than intended, delays may be prevented. This can be done either on macroscopic scale, leading a train onto a different corridor, or microscopically, by changing its route in e.g. a station yard;
- Short turning: Services may be short turned to opposite services at strategic stations, so that the timetable outside the disrupted area can remain intact;
- Cancelling: Services may have to be cancelled due to in-availability of infrastructure, rolling

stock or crew. Passenger of the cancelled service have to find a different route or other services which are not cancelled;

- Extra stop: an additional stop for a service will delay the service with a few minutes, but allows reduction of delays for passengers at the additional stop and decreases for the next train calling at that stop;
- Skip stop: to reduce its delay, a train service may skip one or more of its stations, which can be done if that service would hinder other services by stopping and a next train which does stop at the skipped station is following soon, so that passengers are not too negatively affected.

## 3.7. Conclusion

Goverde (2012) has evaluated the robustness of different railway signalling systems, which is the ability of the system to cope with small delays. Resilience, to be evaluated in this research, is the ability of a system to recover from a disruption given a control management plan.
The performance indicators that are used to measure resilience are punctuality, time until recovery, the total and the maximum delay, and a weighted resilience index, which is the weighted sum of the benefits based on the duration of the disruption. Furthermore, graphical aids such as the deviation area and time-space diagrams are added for simple interpretation of the results.

Two signalling characteristics have been identified to have the largest effect on resilience: Firstly, whether the movement authority is communicated via line-side signals or via in-cab signalling. The latter is expected to have a positive impact on resilience; And secondly, the length of the block sections. The shorter the blocks sections, the more resilient the railway system will be.

In the decision making process of ERTMS in the Netherlands, resilience is an insignificant decision factor. Factors such as the need for replacement and international relevance are more important motives. However, during the design phase of ETCS infrastructural- and signalling lay-outs, resilience can play a role. This role can vary from comparing alternative designs, to adapting a specific design.

Several distinctive types of disruptions have been distinguished that are likely to be resolved differently under different signalling systems and configurations. Moreover, a handful disruptions management measures have been determined to be used in the case study of this research.

# Chapter

# 4. Case study

For the case study on the Utrecht - Den Bosch corridor, data has been collected for four aspects of the railway: rolling stock, infrastructure, signalling and interlocking, and scheduled operations (reference to slides railway traffic management). Firstly in Section 4.1, a corridor is selected, for which all the required data is gathered. The operations on this corridor and the characteristics of the rolling stock are described in Section 4.2 and 4.3, respectively. Lastly, in Section 4.4, the disruption scenarios to be simulated are chosen and described in Section 4.5.

## 4.1. Corridor selection

Not every railway corridor in the Netherlands is suitable. The corridor to be chosen should have certain attributes so that the effect on resilience can be properly compared. Therefore, a set of criteria has been composed in consultation with a former dispatcher working at Arcadis:

- The number of trains per hour should be sufficiently large. The effect between different signalling systems is better visible with short headways and trains closely interacting
- Relatively high infrastructure occupancy, with small buffers between services
- Infrastructure should allow traffic readjustments, such as facilities to short turn services and switchover possibilities
- Somewhat heterogeneous train traffic, not fully homogeneous
- Sufficiently long sections of track between (main) station

The Utrecht - Den Bosch corridor fulfills the specified criteria satisfyingly, due to the variety of number of tracks, the close interaction between following trains and the possibilities for disruption management. That this corridor is appropriate, is moreover substantiated by the fact that this corridor is used in several similar, renowned simulation studies, such as Goverde et al. (2013) and Hansen (2010). The current layout of the corridor can be seen in Figure 8.

The corridor is approximately 50 kilometers long, starting and ending at the main stations of Utrecht CS and Den Bosch, where the intercity trains stop, and with seven intermediate sprinter stations. Four of the sprinter stations (Vaartsche Rijn, Lunetten, Houten and Houten Castellum) are along the four track section, separated from the intercity trains. Culemborg and Zaltbommel are placed at the double track section, and Geldermalsen is a larger hub for sprinters, where overtaking is possible.

The configuration of the infrastructure, e.g. switch, signalling and speed profile characteristics,

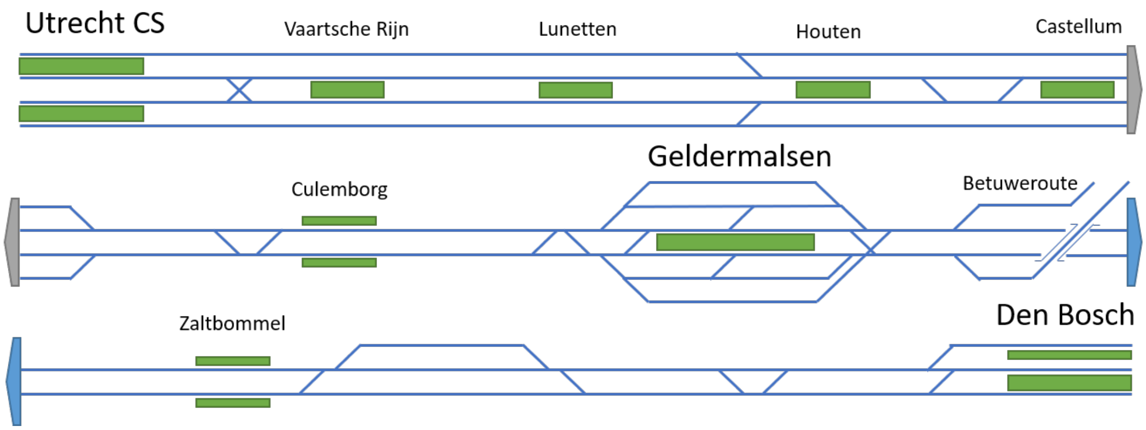



Figure 8: Infrastructure layout as used in this research

Table 2: Station overview

| Station | Abbr | Type |
|---|---|---|
| Utrecht CS | Ut | Intercity |
| Vaartsche Rijn | Utvr | Sprinter |
| Lunetten | Utln | Sprinter |
| Houten | Htn | Sprinter |
| Castellum | Htnc | Sprinter |
| Culemborg | Cl | Sprinter |
| Geldermalsen | Gdm | Sprinter |
| Zaltbommel | Zbm | Sprinter |
| Den Bosch | Ht | Intercity |

as well as the interaction between signals, is based on the design drawings of Arcadis and ProRail referring to the year 2017. The first is imported from the so-called *OBE-bladen* (overview tracks and yards) and the latter from the *OS-bladen* (overview signals)

## 4.2. Scheduled operations

On this corridor, there are three half-hourly intercity (IC) services, two half-hourly local trains (sprinters) and an hourly freight train. The intercity service consists of three lines, each with a frequency of two services per hour. All six IC services connect Utrecht and Den Bosch without intermediate stops. The sprinter service consists of two lines, also with a frequency of two per hour. Both lines go from Utrecht to Geldermalsen. There, one line continues to Den Bosch, and the other diverts to Tiel. The freight trains, coming from and going to the Betuweroute, join the corridor near Meteren, just south of Geldermalsen,

The local train coming from and to the direction Dordrecht/Gorinchem joining just before and ending at Geldermalsen is neglected in this research, as this service is to be physically separated in the conversion of the station Geldermalsen in the coming years, thus having no effect anymore on the rest of the train traffic near Geldermalsen.

A distinction is made between opposite services of the same line, by using even numbers for all southbound services and odd numbers for the northbound services. The last two digits of the train line number indicate the time of operation for said service. Each line operates with a frequency of two per hour per direction, thus the train line number increases by 2 for each consecutive service of a line.

Table 3: Timetable of the passenger services between Utrecht and Den Bosch and v.v.

Ut-Ht

| Station | 800 | 3900 | 3500 | 6000 | 6900 |
|---|---|---|---|---|---|
| Ut | v :03 | v :14 | v :24 | v :11 | v :22 |
| Utva | — | — | — | :13 | :24 |
| Utln | — | — | — | :16 | :27 |
| Htn | — | — | — | :20 | :31 |
| Htnc | — | — | — | a :22<br>v :23 | a :33<br>v :34 |
| Cl | — | — | — | :30 | :40 |
| Gdm | — | — | — | a :36 | a :46<br>v :51 |
| Zbm | — | — | — | | :57 |
| Ht | a :30 | a :40 | a :52 | | a :07 |

Ht-Ut

| Station | 801 | 3901 | 3501 | 6001 | 6901 |
|---|---|---|---|---|---|
| Ht | v :28 | v :18 | v :08 | | v :22 |
| Zbm | — | — | — | | :30 |
| Gdm | — | — | — | v :23 | a :37<br>v :42 |
| Cl | — | — | — | :29 | :48 |
| Htnc | — | — | — | :36 | :54 |
| Htn | — | — | — | :39 | :57 |
| Utln | — | — | — | :43 | :01 |
| Utva | — | — | — | :46 | :04 |
| Ut | a :56 | a :45 | a :35 | a :49 | a :07 |

In Figure 9, the basic hour pattern of the line between Utrecht and Den Bosch is shown. In Appendix 7.3 an enlarged basic hour pattern is added. The red, green, and orange lines represent the trains in the 800-, 3500- and 3900-series, respectively. The black and blue lines are for the trains in the 6000 and 6900-series. The pink line is the hourly freight train.

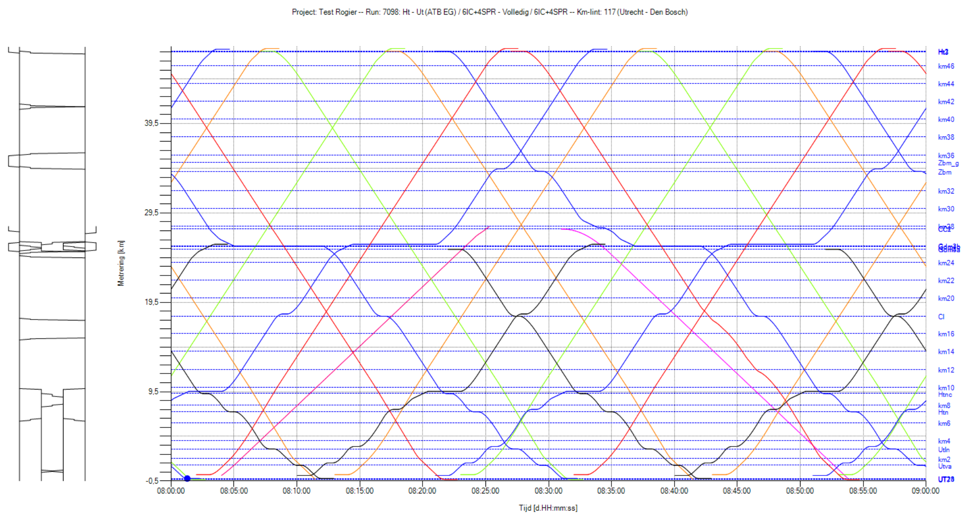



Figure 9: Basic hour pattern between Utrecht (below) and Den Bosch (above)

## 4.3. Rolling stock characteristics

This section provides an overview of the rolling stock that will be used for the simulation. Three types of rolling stock will be used: double-deck trains for the intercity services, low-floor trains for the sprinter services, and freight trains pulled by one or more locomotives.

For the intercity services, a VIRM IV configuration has been selected, as shown in Figure 10a. This configuration is quite common outside peak hours. Inside peak hours, multiple train sets can be coupled. The sprinter services use instead a SLT, shown in Figure 10b. A configuration of a SLT4+SLT4 is a common configuration for this line. A normative freight train configuration on the corridor Den Bosch-Utrecht is a coal train, coming from the Betuweroute, heading towards the harbour of Amsterdam and v.v. These coal trains are pulled by two BR189-locomotives, as shown in Figure 10c, followed by several hundred meters of coal wagons. In Table 4, the characteristics these three train types are reported.

Table 4: Rolling stock characteristics

| Setting | VIRM IV | SLT 4+4 | Freight |
|---|---|---|---|
| Length | 108 meter | 138 meter | 720 meter |
| Nominal deceleration | 0,66 $m/s^2$ | 0,8 $m/s^2$ | 0,33 $m/s^2$ |
| Max deceleration | 2 $ms/^2$ | 2 $m/s^2$ | 1 $m/s^2$ |
| Max acceleration | 1 $m/s^2$ | 1,2 $m/s^2$ | 0,6 $m/s^2$ |

(a) VIRM-IV as intercity (b) SLT-VI as sprinter (c) 2x BR189 pulling a coal train

Figure 10: The three trains

## 4.4. Disruptions on the corridor

By observing historic data on disruptions, the most frequent type of disruptions can be determined. The website *Rijdendetreinen.nl* has collected and categorized for every trajectory in the Netherlands which disruptions have taken place in the last 24 months. From there it can be seen that on the corridor Utrecht - Den Bosch the following most common disruptions have taken place, in order of frequency (data retrieved in April 2019):

1. Train defect: 78 (31%)
2. Switch defect: 22 (9%)
3. Track defect: 18 (7%)
4. Stranded train: 13 (5%)
5. Collision with a person: 10 (4%)
6. Signal failure: 9 (4%)

If these statistic are combined with the theory in Chapter 3.6.2 on which disruptions are interesting to look at regarding resilience, it shows that all these most common disruptions can be used for simulating either a partial blockage or the recovery from a full blockage.

In the following paragraphs, the terms train-, switch, and track defects, and a collision and signal failure are clarified in more detail.

**Train defect/stranded train** This consists of a train coming to a standstill on either an open or closed track section or on a station yard. This disruption could last either only a couple minutes in case of something small as someone pulling the emergency brake after which it can drive further on its own, or up to one or more hours due to e.g. a technical failure, after which the train has to be towed away. In Figure 11a, an example of a train which has stranded with a broken pantograph and cannot drive further on its own. Figure 11b shows a receiver for the ATP signal. Some trains, such as in Figure 11b, have protection in front of the receiver, but others not, making the receiver vulnerable to objects on the tracks.

A train stranded on the open tracks will cause more trouble than a train stranded around a closed track in a controlled station area area where the signals can be controlled by the dispatcher. In the first case, the consecutive trains that are in the same open track will come to a standstill


(a) A broken pantograph will lead to a serious disruption (Brouwer, 2013)

(b) An object on the tracks could damage the ATP receiver (Sweekhorst, 2015)

Figure 11: Two examples of train defects

behind the faulty train. In the latter two cases, it is easier for the train dispatcher to redirect the other trains around the faulty train, or to let them travel backwards.

**Switch defect** If the mechanical moving parts of a switch cannot be set faultlessly and completely in the right direction, a switch is defect. As it can not be assured that the switch is fully safe, trains are not allowed to cross the switch. Causes of a switch defect could be snow, stones or small objects obstructing the moving parts, a faulty cable or motor, or a fault in the controlling computer.

How bad the train traffic is disrupted by a switch defect depends mostly on the location of the switch. If routes can be set easily around the faulty switch, delays can be minimized. However, if the switch is on the entrance of an station yard, it may block all train traffic. Another solution to minimize the impact is to clam the switch in the normal direction, so that train traffic can pass the switch in that direction.

Figure 12: The switch rails can freeze to the stock rails if the heating is not functioning properly (Wikipedia, 2019)

**Track defect** There are several cases of track defects where the speed on the track has to be limited. First, a level crossing could be failing, so that trains have to cross the crossing at 15 km/h for some time. On the corridor Ut-Ht, several level crossings can be found, which thus could fail. Second, there could be an incorrect track occupation due to e.g. a broken cable or rail. After identification that the failing block is most probably safe and unoccupied, the block is to be crossed at max 30 km/h, the maximum speed when 'driving on sight' by all trains, until the problem is fixed.

Figure 13: Example of a broken rail (ProRail, 2019)

**Collision with a person or vehicle** When a train has collided with either a vehicle, or person, all tracks have to be taken out of service until the emergency services have cleared the tracks and the train. Depending on the impact damage to the train and the object of collision, or severity of the injuries, this may take from one hour to almost 3 hours to be resolved.

Figure 14: After a collision, the track is blocked for several hours (ANP, 2018)

**Signal failure** A signal failure occurs when one or more signals keep showing the 'stop'-aspect, although it should not. This can be due to a failure inside the signal, a damaged cable, or when the bulb is broken.

If only one signal is failing, the dispatcher may decided to allow trains through the section the signal is controlling, being it at lowered speed. The more signals are failing, the harder it is to dispatch trains safely, and the longer it takes before the disruptions will be resolved.


Figure 15: A signal failure will bring the signals to the most restrictive aspect (Freepik.com, 2018)

## 4.5. Disruption scenarios

Seven distinct disruptions have been chosen to be simulated, one collision, two train defects, two switch defects, and two track defects. These are used to evaluate the two earlier mentioned types of disrupted situations this research is investigating, partial blockade and full blockade, or more precise, the recovery from a standstill situation and a speed restricted situation.

The paragraphs below will describe in more detail what the reason for the disruptions could have been, which trains are initially disrupted, what the consequence is for the train traffic and what dispatching measures are taken to mitigate the delays. An overview of the disruption scenarios is given in Table 5.

Table 5: Overview of the chosen disruption scenarios

| # | Type | Cause | Consequence | Duration |
|---|---|---|---|---|
| 1 | Train defect | ATP defect | Continuing to end station reduced speed | 15-30 min. |
| 2 | Train defect | Engine defect | Complete standstill, until towed away | 30 minutes |
| 3 | Switch defect | Power failure | Switch is fixed in normal direction | 1 hour |
| 4 | Switch defect | Damaged switch rod | Switch cannot be crossed | 1 hour |
| 5 | Track defect | Level crossing failure | 15 km/h speed limit over level crossing | 30 min. |
| 6 | Track defect | Incor. track occupation | 30 km/h speed limit over a block | 20 min. |
| 7 | Collision | Collision with a vehicle | Standstill, then all trains closely following | 2 hours |

### Scenario 1 - ATP defect

The receiver of the ATP signal, located underneath the train in front of the first axle, can be damaged by small objects on the tracks, as mentioned in Section 4.4. This makes that the ATP is no longer able to receive codes from the track circuits, thus restricting the train to a maximum speed of 40 km/h, which is the maximum allowed speed by the Dutch ATP without a code from the track circuits. This speed needs to be continued to a place where the train can be taken aside without hindering other traffic, in this case places as Geldermalsen, Utrecht or Den Bosch.
With ETCS L2, this can be compared to a loss of radio communication, thus not receiving movement authorities. After contact with the dispatcher, it is allowed to continue at reduced speed. Since train separation is still managed via the track circuits in ETCS L2, the successive trains can follow at block distance.

**Delay specifications** The fourth northbound intercity, the IC807 in the simulation, comes to a standstill south of the station of Zaltbommel as it has hit an object that has damaged the ATP receiver, or it has lost its radio communication with the RBC. It then remains standstill for 10 minutes to identify the problem and to contact the dispatcher on how to proceed. Then, it continues to Geldermalsen at 40 km/h, where it can be sidetracked.

**Effect** Several consecutive northbound sprinter and intercity services will get stuck behind this intercity and will have to follow as close as possible until Geldermalsen. From there on, they can continue unhindered by the faulty train. All southbound services are unhindered by this disruption.

**Scenario disruption management** The disruption is not long and disturbing enough to cancel trains. In Geldermalsen, the faulty train can be sidetracked and be overtaken. The order of the sprinter and intercity services may be changed in Geldermalsen, with possible re-timing, but sprinter still goes in front of several intercity services, otherwise the extra delay of the sprinter would become out of proportion.

Table 6: Overview of scenario 1

| Scenario 1 | |
|---|---|
| Type of defect | Train defect |
| Cause | ATP defect/loss of radio contact |
| Affected train | IC 807 |
| Duration | 10 minutes standstill |
| Location of defect | Near Zaltbommel |
| Effect | IC807 comes to a full standstill for 10 minutes. Then, with 40 km/h to Geldermalsen |
| Hindrance | Four northbound services are stuck behind it |
| Disruption management strategy | Reordering in Geldermalsen |

### Scenario 2 - Engine defect

All mechanical and electrical components of a railway vehicles are prone to failures, such as an electrical relay problem, which can cause a partial or complete engine defect. In case of a complete engine defect, the train is stranded and is not be able to drive further on its own. A tow locomotive or second train is needed to tow the stranded train at a restricted speed of max 80 km/h to Utrecht or Den Bosch The duration depends on the availability of a towing locomotive/train and how fast it can reach the stranded train.

**Delay specifications** The third northbound sprinters, the SP6003, remains stagnant after its stop in Culemborg for 40 minutes. A tow locomotive then comes from Utrecht, couples to the sprinter, and returns to Utrecht at 80 km/h.

**Effect** Several northbound services will get stuck behind this train but wait for the train to be towed away. Trains in the other direction has to pass with 40 km/h for safety reasons

**Scenario disruption management** For the northbound trains stuck between Geldermalsen and Culemborg, there are no suitable dispatching possibilities anymore: the faulty train is blocking the switchover possibilities in Culemborg, and passing the faulty trains via the opposite track once it is driving again, is counter-effective, since the southbound would be hindered more than that delay is reduced for the northbound trains.
For trains in or before Geldermalsen, there are more possibilities: the order of the trains can then still be changed. The intercity and sprinter services are sent alternately, firstly because this is most fair for all passengers, secondly because the extra stop of the sprinter in Culemborg is of little impact compared to the impact of the faulty train, and lastly, because the platform capacity at Utrecht does not allow multiple intercity trains following each other closely. All southbound services have to drive past the failing train at reduced speed for safety reasons, but not dispatching measures have to be taken.

Table 7: Overview of scenario 2

| Scenario 2 | |
|---|---|
| Type of defect | Train defect |
| Cause | Engine defect |
| Affected train | SP6003 |
| Duration | 40 minutes |
| Location of defect | Near Culemborg |
| Effect | SP6003 remain stagnant for 40 minutes |
| Hindrance | Northbound services are stuck behind it and southbound services are speed restricted |
| Disruption management strategy | Reordering of sprinters and intercities |

### Scenario 3 - Switch power failure

A switch can fail by the effect of a broken cable, so that the switch is powerless and cannot be steered. If the dispatcher cannot control the switch, no traffic is allowed past that switch. In this scenario, this is the case for an entrance switch on the south side of Geldermalsen.

**Delay specifications** A northbound entrance switch in Geldermalsen has been damaged and may not be crossed, starting at 7:45, the time at which it was notified, up to an hour later, when it has been repaired.

**Effect** The through track for intercity and freight services in Geldermalsen cannot be used for one hour. The intercity and freight services have to be rerouted via the tracks along the platform, which in normal conditions are used only by the sprinters. In Figure 16, an overview is given of the situation. The intercity services that normally overtake the sprinters in Geldermalsen, are now stuck behind it until Houten Castellum.

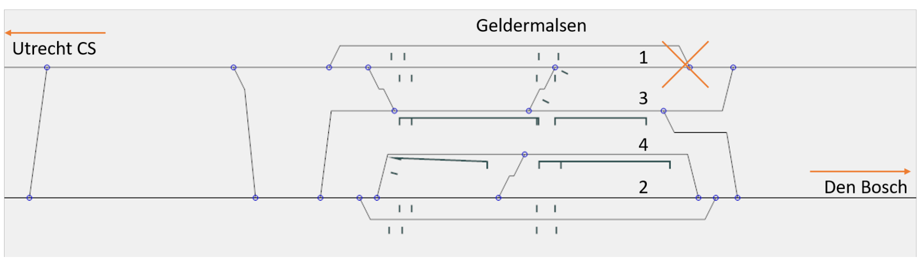



Figure 16: An overview of the layout of Geldermalsen with the defect switch

**Scenario disruption management** The intercity and freight trains which normally use the through tracks are diverted via track 3 in Geldermalsen. As this is the departure track for the sprinters, the sprinters have to go in front of the intercities. The sprinter service ending and starting in Geldermalsen (the 6000-series), can be diverged to another track not to be in the way of a rerouted intercity by which it would be overtaken in Geldermalsen. The running time of this sprinter will slightly increase as it has to go via the track in opposite direction and through extra switches to get to the right track, but this will be less than the delay incurred by an intercity being stuck behind it. Southbound traffic is not hindered by this redirecting of the 6000-series, as the departure time of this sprinter series fits perfectly between all southbound services.

Table 8: Overview of scenario 3

| Scenario 3 | |
|---|---|
| Type of defect | Switch defect |
| Cause | Power failure |
| Time of defect | 07:45 |
| Duration | 1 hour |
| Location of defect | Entrance switch of Geldermalsen |
| Effect | Switch is clammed in one direction |
| Hindrance | Intercity services cannot use intended route and cannot overtake the sprinter in Gdm |
| Disruption management strategy | Rerouting of intercity and sprinter services |

## Scenario 4 - Damaged switch rod

The rod of a switch can be damaged due to vandalism, wear, or objects falling of passing trains. This may cause malfunctioning of the switch, as it cannot be guaranteed that the switch is able to safely The switch then may not be crossed in any direction. The train traffic is only partially disrupted if trains can take other routes around this switch, otherwise all train traffic is obstructed until reparation.

**Delay specifications** The northbound entrance switch unto Castellum, where the sprinter and intercity services normally are separated unto separate tracks, is failing and cannot be controlled by the dispatcher. The disruption starts at 7:40 and takes approximately an hour to be solved.

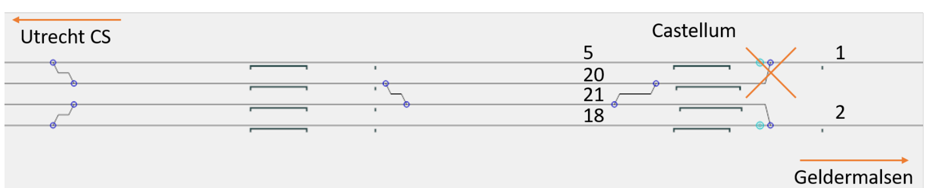



Figure 17: An overview of the layout of Castellum with the defect switch

**Effect** No traffic is allowed for an hour via the northbound entrance switch of Castellum, shown in Figure 17. In Culemborg there are switchover possibilities to switch between the two tracks. This makes that the northbound track between Culemborg and Castellum is out of service, but that the southbound section can still be used, even in both directions. In Houten there are again switchover possibilities for the northbound sprinter and intercity services to return to their original tracks.

**Scenario disruption management** With 11 services per direction per hour, traffic in both directions would not fit onto a 8-kilometer long one-track section. To prevent a complete blockage of the tracks, several services have to be cancelled. One intercity and one sprinter service are cancelled in both directions to make room on the tracks for the other services.

Table 9: Overview of scenario 4

| Scenario 4 | |
|---|---|
| Type of defect | Switch defect |
| Cause | Damaged switch rod |
| Time of defect | 07:40 |
| Duration | 1 hour |
| Location of defect | Entrance switch of Castellum |
| Effect | Switch cannot be crossed for an hour |
| Hindrance | All traffic has to go via one track between Culemborg and Castellum |
| Disruption management strategy | Rerouting of services between Culemborg and Castellum, cancellation of services |

### Scenario 5 - Level crossing failure

If there is a problem to a level crossing, which can be caused by power outage, a broken cable, or problems to the track, it will automatically close, as it would if a train is approaching, in accordance to the fail-safety design rules of the railways. When a level crossing is closed longer than 5 minutes, it is said to be out of order and deemed unsafe. Train drivers then are warned by the dispatcher and have to approach the level crossing at 15 km/h and honk repeatedly as warning, to ensure that no one is passing the level crossing. Meanwhile, a mechanic is called to the spot to repair the failure. After a confirmation and/or that the level crossing is functioning properly, train traffic can pass the crossing at normal speed.

**Delay specifications** The first level crossing after Geldermalsen in northern direction, in Tricht, is failing for half an hour, starting at 7:10.

**Effect** The first ten minutes, no trains can pass the level crossing. After that, trains have to pass the level crossing with 15 km/h for 20 minutes after which the crossing is expected to function safely again, and trains can pass unhindered. This applies to traffic from both directions. In the first ten minutes of stagnation, several services will be queued up in front of the failing level crossing. After this, the services will follow each other as close as possible.

**Scenario disruption management** Services are not cancelled, but could be re-timed and reordered in Geldermalsen to let the intercity services overtake the sprinters. However, in the gap, emerging from a sprinter service accelerating earlier out of the speed-restricted zone than the successive intercity service, a stop in Culemborg does not affect the successive intercity. The sprinter and intercity are therefore not reordered. The same holds for the sprinter and intercity service following a delayed freight train. The sprinter is sent in front of the intercity to minimize their combined delay.

Table 10: Overview of scenario 5

| | Scenario 5 |
|---|---|
| Type of defect | Track defect |
| Cause | Level crossing failure |
| Time of defect | 07:10 |
| Duration | 30 minutes |
| Location of defect | Level crossing in Tricht, north of Geldermalsen |
| Effect | Trains have to slow down<br>for the crossing |
| Hindrance | 10 minutes no traffic over the level crossing<br>Then 20 minutes speed restriction of 15 km/h |
| Disruption management strategy | Reordering |

### Scenario 6 - Track circuit defect

A track circuit can be shortcut by the effect of a broken cable or a rail fracture, which gives an incorrect track occupation notification to the train dispatcher, to whom it looks as if a train is present in that track section. Trains approaching the section will encounter a stop signal. Only when the dispatcher is certain that the section is in reality unoccupied and safe, trains can be allowed over the section, being it at a maximum speed of 30 km/h, the speed for driving on sight.

**Delay specifications** The northbound section of the station Zaltbommel gives an incorrect track occupation, shortly after 7:00.

**Effect** For 30 minutes train drivers have to stop in front of the signal leading up to that block and may only cross the section after receiving approval from the dispatchers to drive through that red signal with a speed restriction of 30 km/h. At the end of the section, they can speed up to the maximum allowed speed. For safety reasons, traffic in the other direction is also slowed down.

**Scenario disruption management** In normal conditions, the 801- and the 3501-series overtake the 6901- and 6001-series, respectively, in Geldermalsen. However, the first delayed 801-service is delayed that much that the sprinter does not have to wait for it to pass, and can depart at its usual departure time, without disturbing the already delayed intercity even more. The same goes for the 6901-sprinter, although then the successive intercity is slightly delayed. But still, the combined delay would be smaller than letting the sprinter wait until the delayed intercity has passed. In southern direction, no disruption management is needed, what happens after Den Bosch is outside this scope.

Table 11: Overview of scenario 6

| Scenario 6 | |
|---|---|
| Type of defect | Track defect |
| Cause | Incorrect track occupation |
| Time of defect | 07:05 |
| Duration | 30 minutes |
| Location of defect | Zaltbommel |
| Effect | For 30 minutes, the block with the incorrect track occupation can be crossed at maximum 15 km/h |
| Hindrance | All services have to stop in front of the signal and will follow each other more closely |
| Disruption management strategy | Reordering of services in Geldermalsen, |

### Scenario 7 - Collision

In this scenario, the robustness of the signalling systems will be tested. This means that no dispatching measures are taken, but that the system has to resolve itself. Delays in this scenario will be bigger than with dispatching measures, stipulating the differences between the signalling configurations even more.

An object, such as a bike or scooter, has stranded on a level crossing and was not able to be cleared from the tracks before the train arrives, and thus come into collision with the approaching train. The consequence is that all train traffic has to be stopped, until the emergency services have arrived and cleared the tracks.

**Delay specifications** The fourth northbound intercity, the IC 3503 in the simulation, has collided with an object, and comes to standstill at emergency braking distance after the level crossing of collision in Schalkwijk, 12 km south of Utrecht, between Houten and Culemborg.

**Effect** The track between Culemborg and Houten cannot be used for one hour to wait for the emergency services and investigation of the collision, and all trains, in both directions, are stuck behind one another during that time.

**Scenario disruption management** No disruption management will be applied, that is why this scenario is testing robustness and not resilience. Disruption management would have been to cancel the intercity services in both directions and to short-turn the sprinter at Castellum and Geldermalsen.

Table 12: Overview of scenario 7

| Scenario 7 | |
|---|---|
| Type of defect | Collision |
| Cause | Collision with a vehicle |
| Affected train | IC 3503 |
| Duration | 1 hours |
| Location of defect | Schalkwijk, between Culemborg and Houten |
| Effect | IC 3503 comes to standstill for 1 hour at the level crossing |
| Hindrance | No services can pass that parts of the tracks |
| Disruption management strategy | None |

## 4.6. Signalling variants

In the theory, two signalling characteristics have be distinguished which are most potentially relevant for railway resilience: the length of the block section and whether the movement authority is sent via line-side signals or via in-cab signalling. The disrupted scenarios will be simulated for three different signalling lay-out configurations: the current configuration of NS'54/ATB-EG, and two different configurations of ETCS L2, one of which is a likely configurations to be implemented as replacement of current infrastructure and one as a hypothetical configuration, with shorter blocks.

### 4.6.1. NS'54/ATB-EG configuration

This configuration is an exact copy of the infrastructure as it lays 'outside'. The data about the location of stations, switches, (relation between) signals, and speed limit markers is imported via the drawings made by Arcadis of the infrastructure layout. Further on, this configuration will be referred to as the 'ATB-configuration'.

### 4.6.2. ETCS L2 existing blocks

This configuration is a copy of the ATB-configuration, using the same blocks for track detection and the same placement of signals. It is a realistic signalling configuration, as in the short term it is the cheapest solution to replace a legacy protection system. This on the premise that the same track detection is used currently as it is costly to adjust track detection block sections. Would axle counters be used then it is somewhat easier and cheaper to adjust the number of blocks and the length of the blocks.

With this configuration, the pure effect of in-cab signalling of ETCS L2 over the line-side signals of ATB can be measured, as no other factors are changed expect for the protection system.

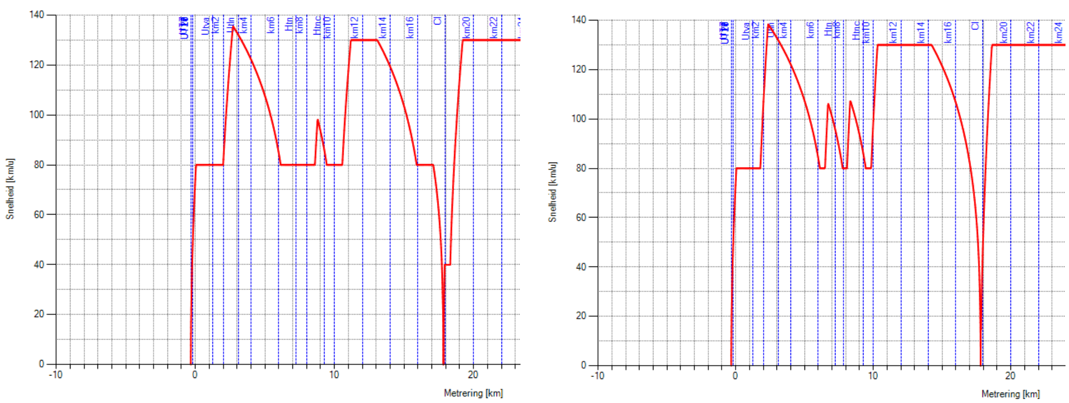
### 4.6.3. ETCS L2 reduced blocks

This is a mainly theoretical variant which approximates the benefits which could be achieved with moving blocks, by implementing block sections every 200 meter. In reality, it is too expensive and impractical to implement such short physical blocks. In ETCS L3, on the other hand, there are no physical blocks anymore, but virtual blocks, the number of which can be increased without costs for additional track-side devices. By using virtual block lengths of 200 meter, a moving block configuration can be approached. Such short block lengths allow that the follow-up times between trains can be minimized and movement authorities can be provided very close to the rear of the preceding train.
With this configuration the impact of short block sections can be measured, by comparing it to the other ETCS L2 configuration, as well as the impact of short block sections combined with in-cab signalling over 'normal' block lengths and line-side signals, by comparing it to the ATB configuration.

On the basis of this configuration, one could try to determine the optimal blocking configuration, by removing the signals that have no effect on the outcome of the simulation. However, as the outcome of the simulation is moreover dependent on other factors, such as the train types and the timetable, that this would be a research in itself and thus falls outside the scope of this research.

## 4.7. Validation of the model

To validate the train behaviour in the model, a comparison is shown in Figure 18 of a modelled train having to brake several times, either to stop or to enter a speed restricted situation. From here, it can be seen that the ETCS L2-train brakes later, in space as well as in time, and that is not effected by the 'yellow'-aspect, which can be seen for the ATB-train near km 18.
Only one ETCS L2 configuration is added in Figure 18, since there is no difference between the braking behaviour of the two ETCS L2 configurations. The reduced blocks are not of influence to the braking profile for this modelled train.




(a) ATB configuration

(b) ETCS L2 reduced blocks configuration

Figure 18: Time-space diagram showing the difference in braking behaviour between ATB and ETCS L2 of a modelled train in Xandra (movement is towards the left)

## 4.8. Conclusion

The corridor Utrecht - Den Bosch has been selected for the case study of this research, based on its appropriate length, its heterogeneous and intensive traffic, and the available spatial traffic management possibilities. Traffic on this corridor consists in both directions of three half-hourly intercity services, two half-hourly sprinter service, and an hourly freight train. One sprinter service and the freight train go as far Geldermalsen/Betuwelijn, all other traffic continues onto Den Bosch. The corridor Utrecht - Den Bosch does not stand alone, all services have their start and/or final destination outside this corridor.

Several types of disruptions have been distinguished that are both likely to occur on this corridor, based on historical data, and that are relevant with respect to resilience. Based on these disruptions, seven scenarios, varying in cause, length, location and severeness, have been constructed to test resilience.

Three signalling configurations are to be tested on resilience. Firstly, the current NS'54/ATB configuration; secondly, an ETCS L2 configuration based on the existing blocks; and lastly, an ETCS L2 configuration with decreased block section lengths.

# Chapter

# 5. Simulation of case study

This chapter will explain for all the scenarios that have been created in Section 4.5, how it has been set up in the simulation model, what the results were, and which conclusions can be drawn.

The simulation tool Xandra calculates the delay of a train at every station or timetable point for which a planned time of arrival has been entered. This will give only a very discrete overview of the delay and the delay cannot be tracked between stations. A more continuous overview of the delay, however, is preferred and will improve the analysis. As a solution, timetable points have been added such that each train passes a timetable point, either a station or a point along the open tracks, nearly every two kilometers. As the trains have a speed of about 130 to 140 km/h, they will pass a timetable point surely each minute. The delays of trains that are standing still is only measured as soon as that train moves again and passes a timetable point.

The maximum delay measured in the results is dependent on the timestep set in the processing of the data. The punctuality and sum of the delays are not affected by this as they are measured per train per timetable point, not per minute. The delays are binned in timesteps of two minutes, as larger timesteps make the graph of the deviation area, delay per moment in time, too coarse for a solid interpretation of the results, and smaller timesteps also make it too unclear.

The indicator "$\sum$ delay" is not the summed area under the deviation graph, but the summed delay of all trains at their final destination. This way, the indicator is operator oriented, for whom only the delay at certain larger stations is measured for punctuality, and useful for expressing the benefits of ETCS L2 in monetary values. The other way, the summed delay would only be useful for comparative conclusions between the configurations and not for absolute conclusions.
The "Recoverytime" indicator is measured as the time it takes from the moment the disruptions starts to be resolved, i.e. when the first train can resume after standstill, until the moment when the total delay in the system is less than five minutes.

Each scenario is simulated from 6:00 AM to 11:00 AM. The scenarios have been set up such that the timetable has fully started up by the time the disruption starts, and that all train series continue for some time after the disruption has been solved.

## 5.1. Simulated scenarios

### Scenario 1 - ATP defect

In this scenario, a faulty train has been standing still for 10 minutes near Zaltbommel and then proceed at 40 km/h to Geldermalsen, to be sidetracked. In Figure 20a, 20b and 20c, the time-space diagram is shown of the trains following the faulty train, in the ATB-, ETCS L2 existing-, and ETCS L2 reduced configuration respectively. These time-space diagrams show what happens when the faulty train, the orange line, can proceed, starting in the upper left corner. Three services have already queued up behind this train, and several more have been queued up when arriving Geldermalsen.

From the deviation area in Figure 19 as well as from the delay statistics in Table 13 it can be seen that the benefits of ETCS L2 are, similar to case 1, significant. The delay at the final stations is, compared to the ATB case, 9% lower in the L2 existing configuration of ETCS L2, and as much as 13% lower in the L2 reduced configuration.

The reason the delay almost triples somewhat after 8 o'clock is that the trains from then can drive at full speed again, 130 km/h instead of 40 km/h, thus passing three times as many measuring points per unit of time.

Figure 20 provides a visualization of and explanation for the difference in delays between the configurations.

The difference in headways and consequently the difference in the final delays is partly caused when the other trains have to follow to faulty train, but is moreover caused when the faulty train finally can be overtaken. Trains with ATB are restricted by the yellow aspect, which forces them to 40km/h until the next signal is a 'better' aspect than yellow. In the ETCS L2 cases, the meaning of the yellow signal is put away. The approach to the end of the movement authority (the 'red signal') is now only restricted by the braking curve, which updates instantly after the previous train has cleared its block, due to the continuous communication of the movement authority. This reduces the headway between trains, and could mean that with ETCS L2 the buffer between trains can be smaller than with ATB.

In sections where the blocks of the NS'54/ATB signals have been reduced, the blocks being shorter than the normative braking length and the ATB-coding thus being incremental, is the effect of the yellow signal smaller, as a driver now receives more distinctive information. However, these optimized configurations can usually only be found on a few sections near stations areas or switches, and not on the open tracks. It does raise the question how effective a fully incremental ATB configuration would be, with the smallest signalling and coding steps as possible. The distance between the signals then could be reduced to a minimum of 970 meters, which is the most restrictive braking distance between two adjacent ATB-codes (130 to 80), whilst in the current configuration the signals on the open track are placed around 1,4 km apart.

From the time-space diagrams, it can be seen that especially in the ETCS L2 existing blocks, the driver behaviour is not fully realistic. Each time a train is allowed in the next block, it accelerates until the start of its braking curve, which is not comfortable and energy efficient. In the L2 reduced configuration, this happens too, but on a much smaller scale, as the blocks are much smaller.

The conclusion of this scenario is that the in-cab signalling has had the most impact on mitigating the delays, more than the decreased blocking lengths, and that the additional benefit of the smaller blocks is limited by other factors, such as platform capacities.

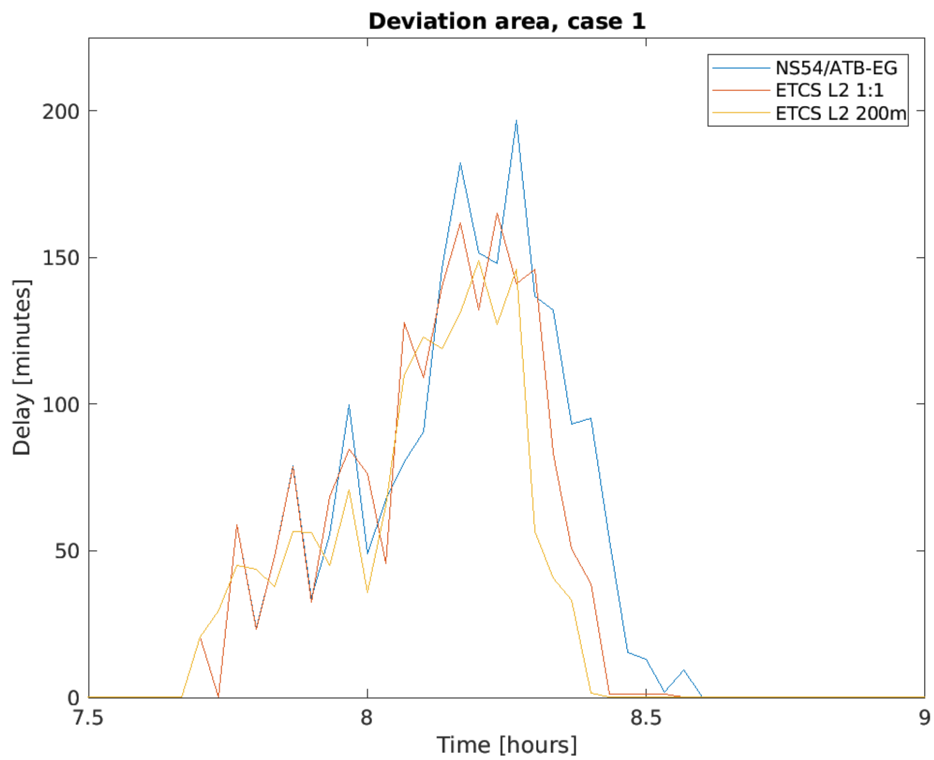
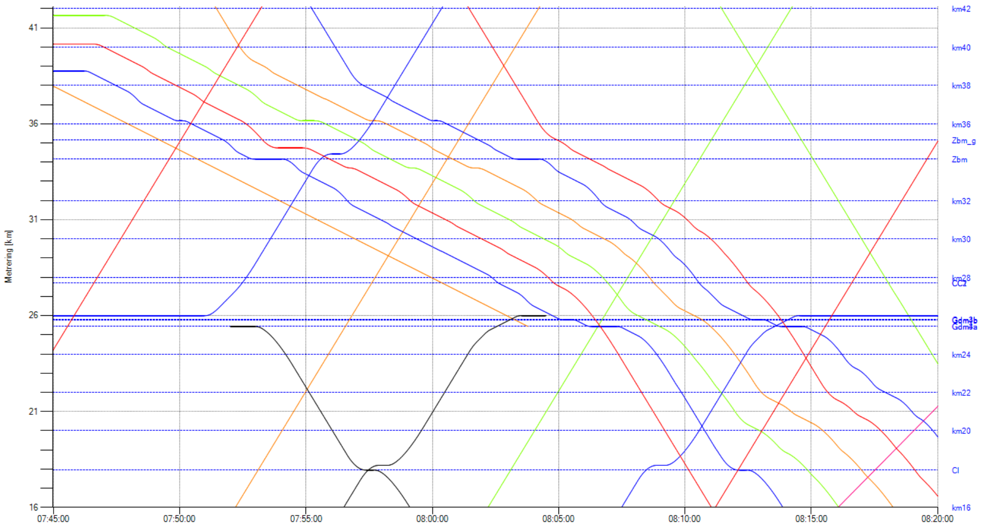


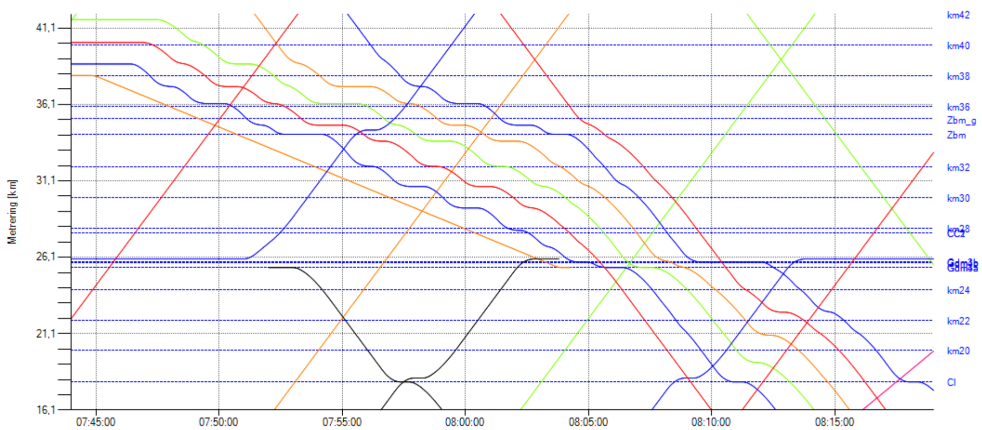

Figure 19: Deviation area of case 1

Table 13: Delay statistics of case 1

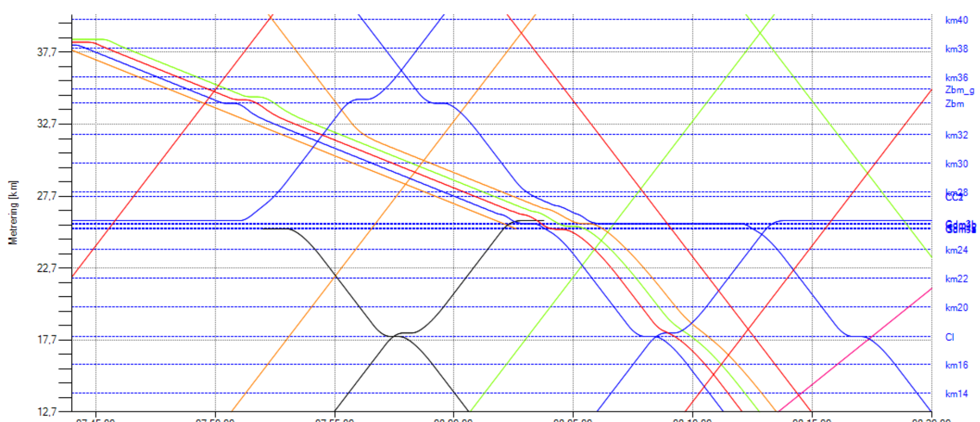
| | 3 min punct (%) | 5 min punct (%) | Recoverytime (h:mm) | $\sum$ delay (h:mm:ss) | Max delay (min) |
|---|---|---|---|---|---|
| NS'54/ATB | 93% | 94% | 0:55 | 2:11:53 | 197 |
| ETCS L2 existing | 94% | 95% | 0:45 | 1:59:50 | 165 |
| ETCS L2 reduced | 95% | 95% | 0:43 | 1:54:52 | 149 |




(a) ATB configuration




(b) ETCS L2 existing configuration




(c) ETCS L2 reduced configuration

Figure 20: Time-space diagrams of scenario 1

## Scenario 2 - Engine defect

In this scenario, a train has broken, and is therefore standing still for 40 minutes until either it is towed away or removed using its own power. However, its speed until Utrecht is restricted to 80 km/h. In the time-space diagrams in Figure 22, it is not shown how the traffic in the other direction has to slow down when passing the faulty train. This because it has no effect worth mentioning on the train traffic, and no difference between the configurations.

As argued in the explanation of this scenario, it is indeed useful to send the intercity and sprinter alternately. Even with the sprinter in between two intercities, the platform capacity in Utrecht for the intercities is still restrictive, yet only in the ETCS L2 configurations. In the L2 existing one, there is only 1 minute buffer between the intercities at the platform. In the L2 reduced one, the successive intercity has to wait.

Compared to the ATB configuration, the ETCS L2 existing and reduced configuration save 14% and 25%, respectively, of the total delay at the end stations, as can be seen from Table 14. From the deviation area in Figure 22, it can be seen how the configurations differ in the handling of the train traffic in the aftermath of the disruption.
The L2 reduced configuration has a lower peak, and it reaches this peak earlier because all trains are moving sooner, which are to the credit of the small blocks. The L2 existing configuration is similar to the ATB one, in terms of the peak, meaning the in-cab signalling is of no influence here. In the aftermath of the disruption, this does matter apparently, as here the delay is significantly different.

Table 14: Delay statistics of case 2

| | 3 min punct (%) | 5 min punct (%) | Recoverytime (h:mm) | $\sum$ delay (h:mm:ss) | Max delay (min) |
|---|---|---|---|---|---|
| NS'54/ATB | 90% | 91% | 0:46 | 5:23:08 | 406 |
| ETCS L2 existing | 91% | 92% | 0:40 | 4:39:02 | 404 |
| ETCS L2 reduced | 94% | 94% | 0:30 | 4:02:10 | 377 |

The large profit of the ETCS L2 reduced block configuration is twofold: firstly, as in both the previous cases, all trains are moving quicker after the disruption has resolved, thus the delay decreases sooner, leading to a significantly smaller total delay; secondly, where the L2 existing configuration also makes a difference compared to the ATB, a slower delayed freight train disrupts the other train traffic far less. In the normal schedule there is large enough buffer between the freight train and the following train(s) to cope with the speed difference. Now there is no buffer, so they are stuck behind the freight train. This has little impact in the L2 reduced configuration, whilst in the L2 existing and ATB ones, more trains are hindered by the freight train, seeing from the peaks around 8 o'clock, compared to the horizontal line of the L2 reduced one, another reason for the delay being dissolved later.
A dispatching measure could have been to sidetrack the freight train in Geldermalsen until the next hole in the timetable, so that it can run without causing conflicts to successors. However, then it would have to wait half an hour, which is less than the delay it is causing in its initial path.

The conclusions of this scenario is that the benefit of the in-cab signalling and the shorter blocks is equally substantial, with both a significant amount of delay can be saved. However, the platform capacity is again so restrictive that the benefits of the smaller blocks is done away with, as trains have to wait to enter the station. This is also the case with L2 existing configuration, so just with the in-cab signalling and not the smaller blocks, but when the sprinters and intercities are sent alternately after Geldermalsen, this can be prevented. It has to be noted that this is only functional when the sprinters have only one extra stop compared to the intercities.

The alternate sending of sprinter and intercity is not just most fair, but will also only increase the delay of the intercity slightly, but only when the sprinters is stopping once or twice, not more often.

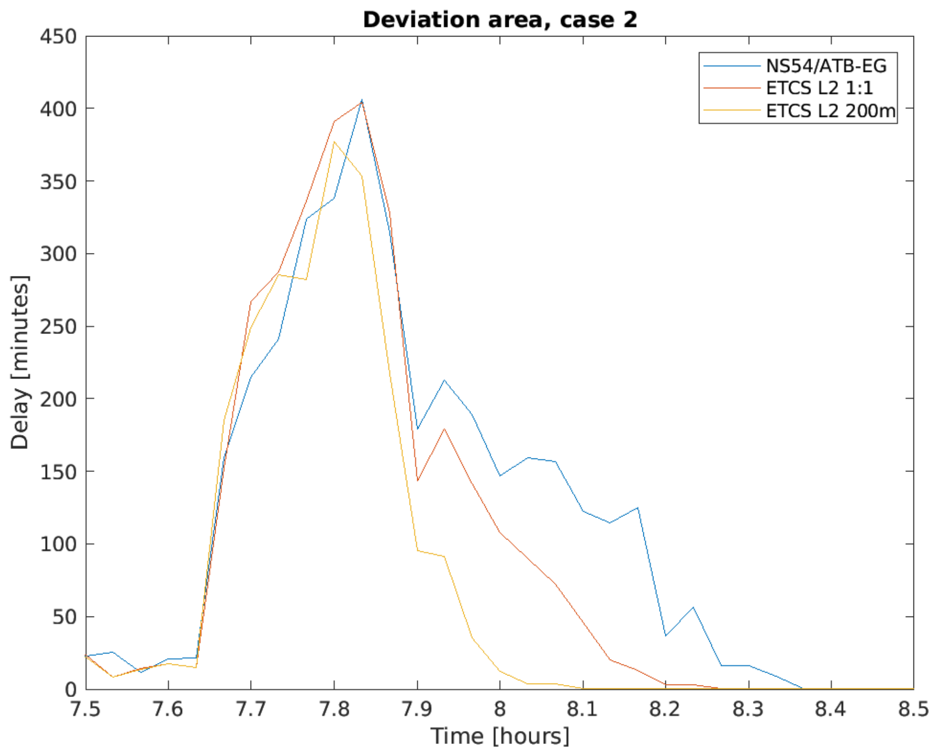
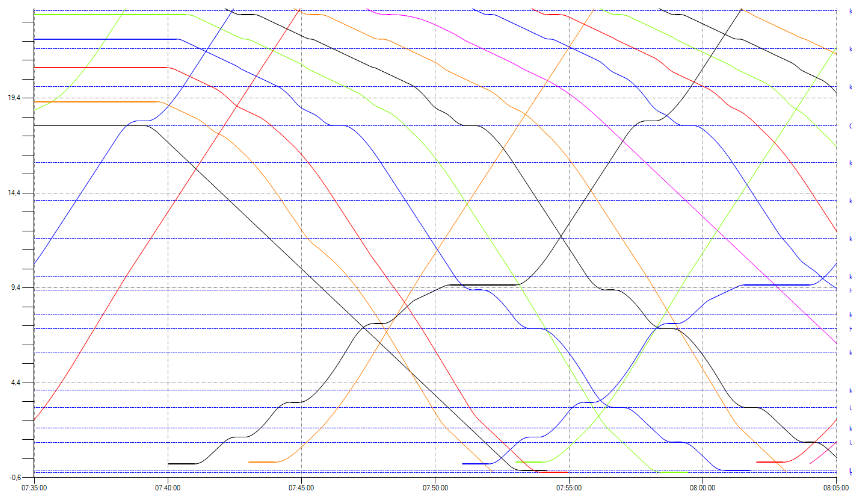
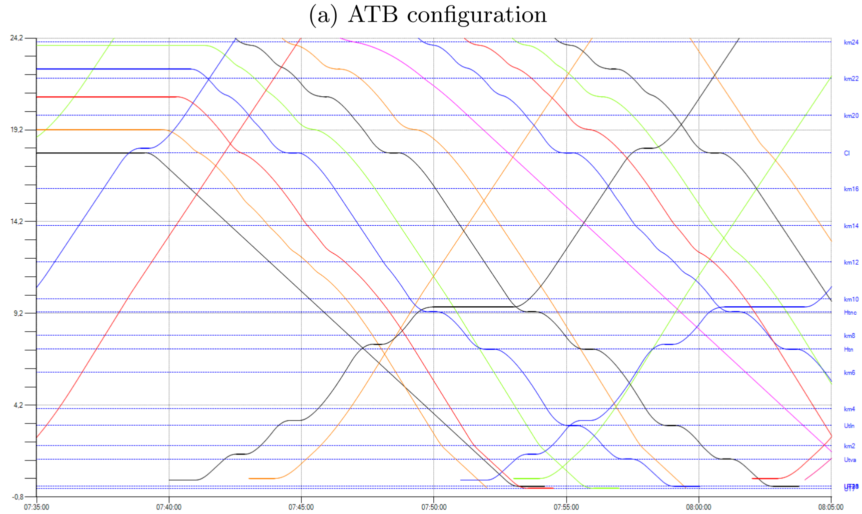


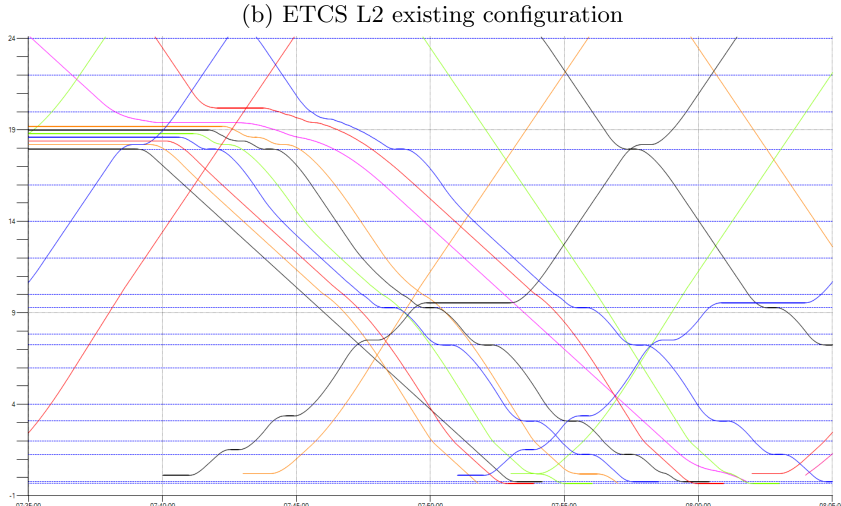

Figure 21: Deviation area of case 2



(a) ATB configuration



(b) ETCS L2 existing configuration



(c) ETCS L2 reduced configuration

Figure 22: Time-space diagrams of scenario 2

### Scenario 3 - Switch power failure

Due to a switch power failure, the northbound intercity and freight services cannot use the through track in Geldermalsen, by which they normally overtake the sprinter services having a long stop there. The effect of this can been seen in the time-space diagrams in Figure 24 as the 'red' service having to wait in Geldermalsen to go behind the 'blue' sprinter service. Would it not be for the disruption measure to relocate the 'black' sprinter to another track, then this would also happen for the 'green' intercity and the 'black' sprinter. This sprinter does have to return to Utrecht partly via the track in the opposite direction to get to switches bringing it back to its normal track. However, this does not bring any delays to opposite trains. The effect of this dispatching measure is therefore nearly the same in all configurations.

Both ETCS L2 configurations score equally well in terms of delays at the final stations. Both save up to 19% of the delay incurred in the ATB configuration, as can be calculated from the statistics in Table 15. In the deviation area as shown in Figure 23 it can be seen that the lines of the ETCS L2 configurations are very close and overlap very much, but that they are both significantly different from the ATB configuration.

Table 15: Delay statistics of case 3

| | 3 min punct (%) | 5 min punct (%) | Recoverytime (h:mm) | $\sum$ delay (h:mm:ss) | Max delay (min) |
|---|---|---|---|---|---|
| NS'54/ATB | 99% | 99% | 1:06 | 0:26:13 | 27 |
| ETCS L2 existing | 99% | 99% | 1:06 | 0:21:11 | 22 |
| ETCS L2 reduced | 99% | 99% | 1:04 | 0:21:10 | 18 |

One difference between the configurations is when the sprinter is making a stop and how close the successive intercity, which is stuck behind the sprinter, can follow. The difference in final delay of the successive intercity is 6:47 with ATB, 4:59 with the L2 existing blocks, and 4:43 with the ETCS L2 reduced blocks. This large difference is due to a combination of the blocking distance and the in-cab signalling. Another difference is made when the intercities are redirected via the sprinter track in Geldermalsen. To go the the other tracks, they have to lower their speed to go through a diverging switch. The in-cab signalling makes that the braking curves of the ETCS L2 trains are more optimized than those of the ATB-trains. E.g.: the delay of the 'orange' intercity, a service which is only hindered by the diverging switches and not by other trains, is 1:20 in the ATB-case, but 1:06 in both the ETCS L2 configurations. The in-cab signalling thus saves 14 seconds in this situation. However, for the relocated sprinter in Geldermalsen, which has to go via the opposite track and thus some extra switches, there is no significant difference in its arrival delay between the configurations. The benefit of the train-specific speed profile is thus only useful for situations in which the speed decreases, not where it increases. For the latter, the ATB-coding is just as advantageous, or in practice even more advantageous, as the ATB does not know when the last part of the train has left the speed-restricted section. This is for the driver to estimate, but nothing stops him for accelerating too early. The benefit of ETCS here is then more in the safety aspect.

In this scenario, the difference between the ETCS L2 configurations has been negligible, thus the benefit compared to the ATB configuration is mainly due to the in-cab signalling and not so much the smaller blocks.

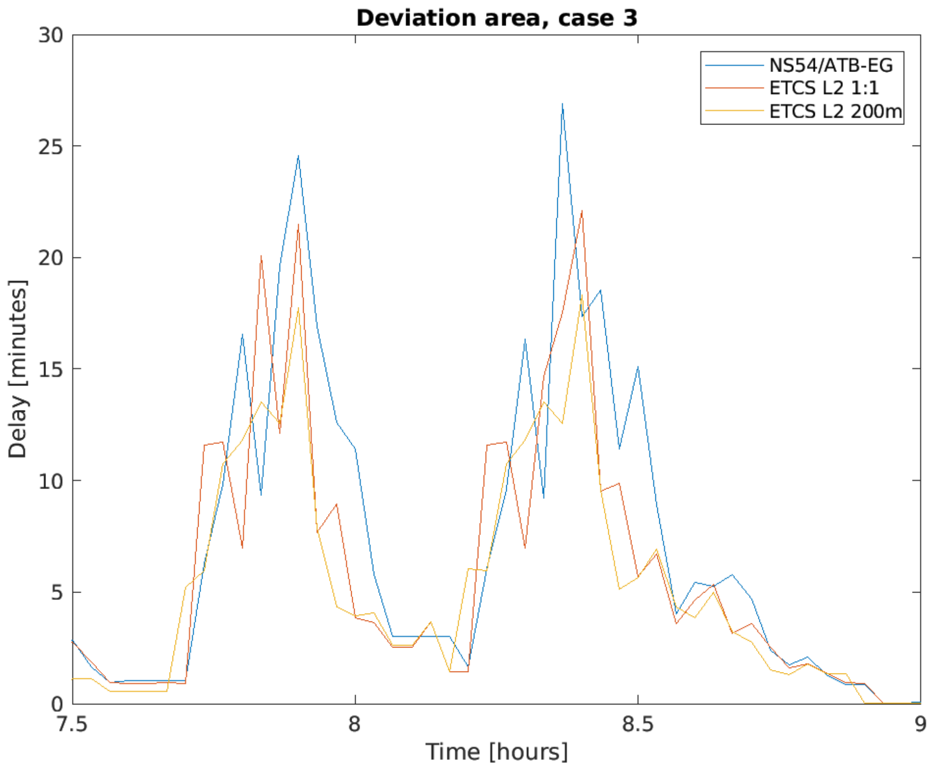
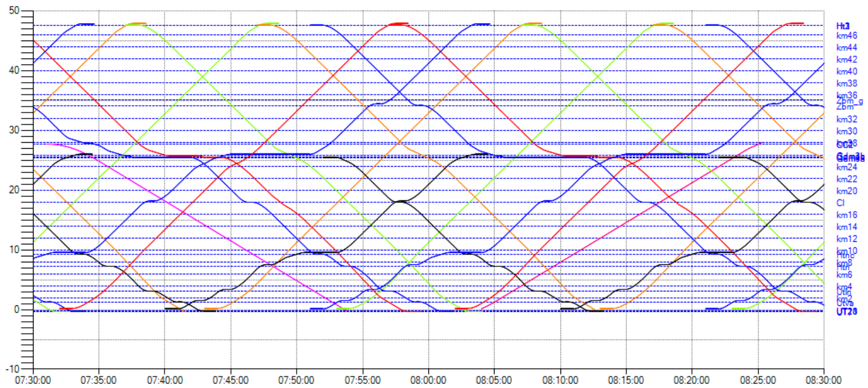
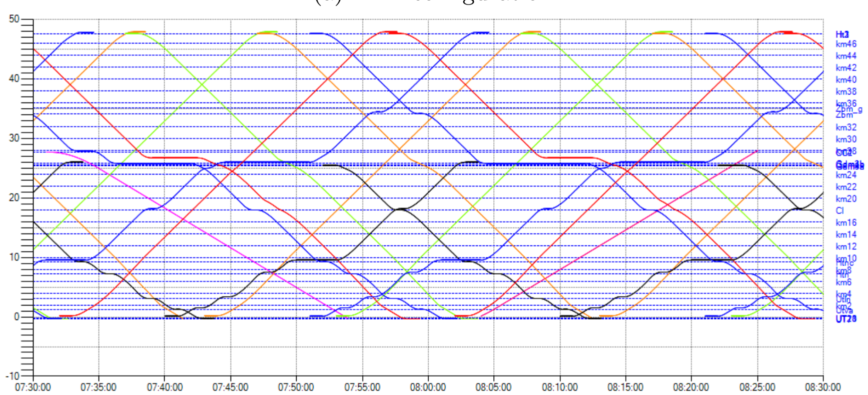


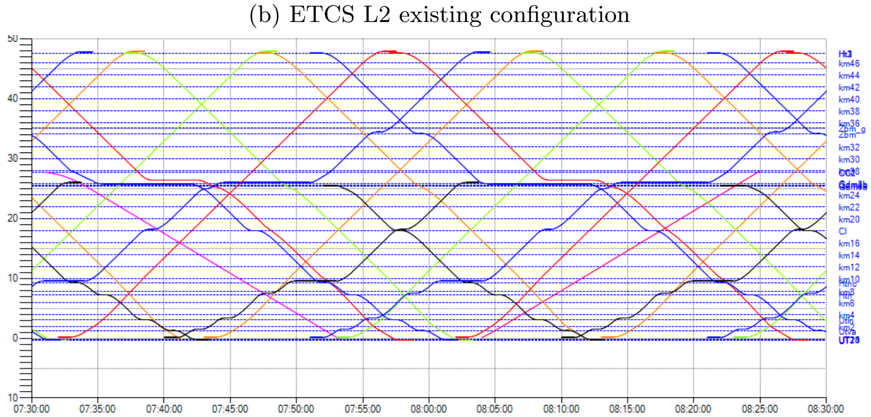

Figure 23: Deviation area of case 3




(a) ATB configuration




(b) ETCS L2 existing configuration




(c) ETCS L2 reduced configuration

Figure 24: Time-space diagrams of scenario 3

## Scenario 4 - Damaged switch rod

The northbound entrance switch of Castellum, where the slower and faster traffic normally splits, cannot be crossed for an hour. All northbound traffic therefore has been redirected via the opposite track from Culemborg to Castellum. After Castellum and Houten, all northbound traffic can go back to their initial tracks. To be able to fit all train on a 8-km one-track section, one intercity and one sprinter service have been cancelled per direction. Alternately, a couple of trains in the same direction can go in a group across the one-track section. This has worked remarkably well, since the trains barely have to wait at trains from the other direction.

In the time-space diagrams in Figure 25, it can be seen that at 07:40, a couple services have flocked, waiting for the 'red' intercity to Den Bosch to clear the one-track section. They then run to Castellum together. The first intercity the other direction, the green line, is slightly delayed due to the oncoming traffic. This phenomenon happens again around 08:10.

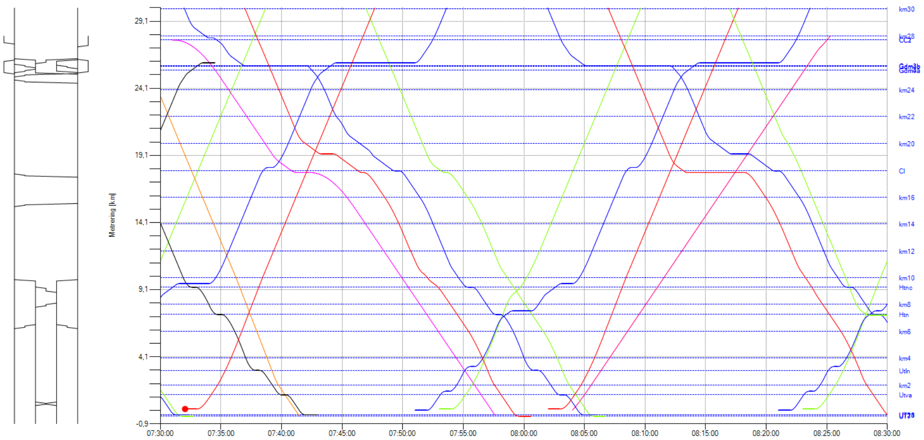
The L2 existing blocks configuration is able to reduce the delay with 9%. For the L2 reduced configuration the delay is reduced by 13%. This reduction is not as big as seen in the second and third scenario. Apparently, those scenarios have been set up such that the potential of ETCS L2 could be used better. From the deviation area in Figure 26 it can be deduced that the benefit of ETCS L2 lies in the aftermath of the disruption, as the peaks during the disrupted phase are similar. Both the ETCS L2 configurations, though, cause the delay to decline faster.

Table 16: Delay statistics of case 4

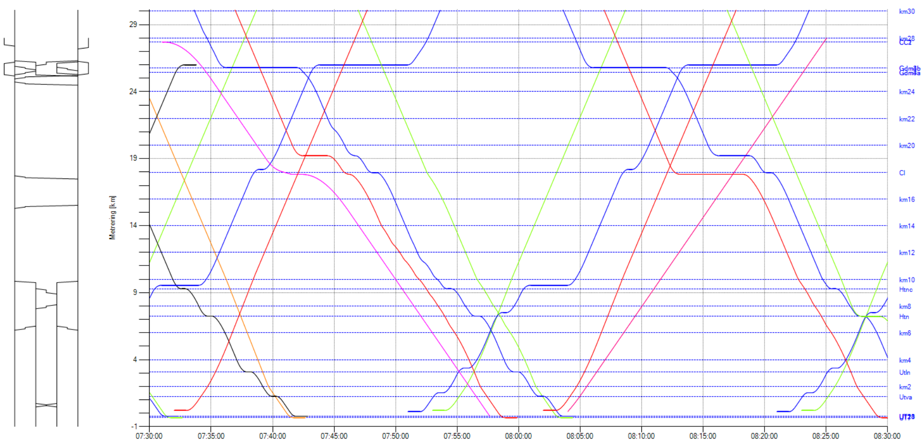
| | 3 min punct (%) | 5 min punct (%) | Recoverytime (h:mm) | $\sum$ delay (h:mm:ss) | Max delay (min) |
|---|---|---|---|---|---|
| NS'54/ATB | 98% | 99% | 1:05 | 0:30:52 | 26 |
| ETCS L2 existing | 98% | 99% | 1:03 | 0:28:05 | 19 |
| ETCS L2 reduced | 99% | 99% | 1:01 | 0:26:58 | 18 |

One cause for the difference is because of the train-specific braking curves in ETCS L2: the time loss of the northbound 'green' intercity having to brake at the beginning of the section in front of the switchover possibility, and then having to drive 80 all across that section, directly impacts the delay of the southbound 'green' intercity. The impact of this is thus felt twice.
Another cause is how the in-cab signalling allows the successive intercity and sprinter to follow the slower freight train more closely, saving one minute delay per train. In the L2 reduced configuration, an extra minute for the intercity and even two for the sprinter, can be saved.

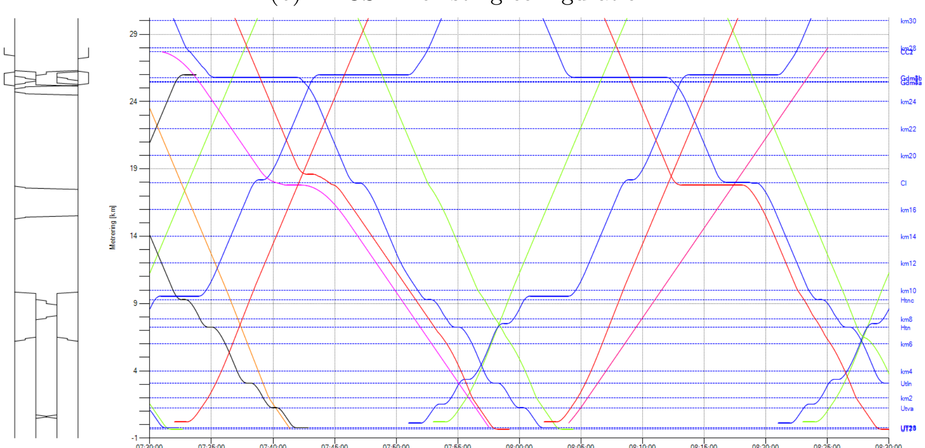
The conclusion of this scenario is that both the in-cab signalling and the smaller blocks cause a significant decrease of the delays incurred in this scenario. The effect of the in-cab signalling and the smaller blocks became clear when an intercity and sprinter had to follow a slower freight train.



(a) ATB configuration



(b) ETCS L2 existing configuration



(c) ETCS L2 reduced configuration

Figure 25: Time-space diagrams of scenario 4

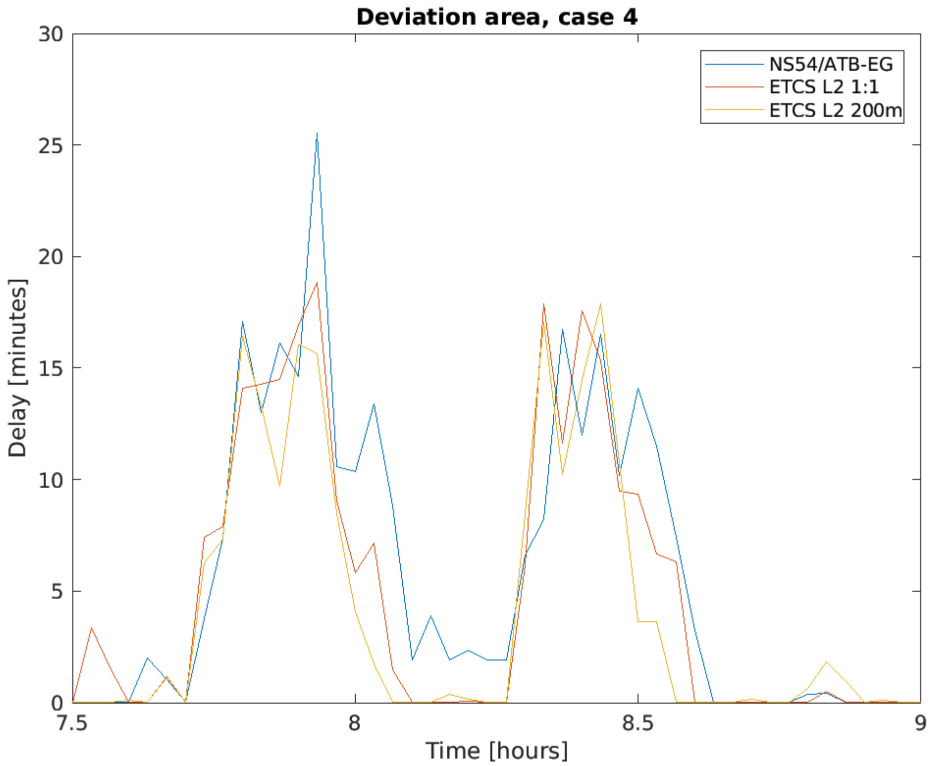



Figure 26: Deviation area of case 4

## Scenario 5 - Level crossing defect

A level crossing north of Geldermalsen is failing in this scenario. This makes that the first 10 minutes after the failure, in both directions no trains can pass the level crossing, as it cannot be confirmed that the crossing is safe to pass. Thereafter, it may only be crossed at 15 km/h for 20 minutes.

The dispatching measure in this scenario, to send the sprinter in front of the intercity service twice, instead of letting the intercity overtake the sprinter in Geldermalsen, seems counter intuitive at first, but has shown to reduce the combined delay of the sprinter and intercity. The low speed across the level crossing causes a headway large enough between the sprinter and the consecutive intercity that the stop of the sprinter at Culemborg does not impact the consecutive intercity. And the second time, the sprinter and intercity are following a delayed freight train. Because of the speed difference between the freight train and the sprinter and intercity services, the stop of the sprinter in Culemborg has no influence on the headway time of the successive intercity service, which is still determined by the slower freight train.

The difference between the ATB and L2 existing configuration is minimal, but the difference to the L2 reduced configuration is quite large. However, the L2 reduced configuration saves almost 28% of the delay incurred in the ATB configuration.

The disruption at the level crossing in itself is not handled differently in the different configurations, proven by the first delayed train which is only hindered by the speed restriction and no other

Table 17: Delay statistics of case 5

| | 3 min punct (%) | 5 min punct (%) | Recoverytime (h:mm) | $\sum$ delay (h:mm:ss) | Max delay (min) |
|---|---|---|---|---|---|
| NS'54/ATB | 93% | 93% | 0:49 | 2:12:51 | 148 |
| ETCS L2 existing | 92% | 93% | 0:49 | 2:09:55 | 160 |
| ETCS L2 reduced | 93% | 93% | 0:43 | 1:35:50 | 126 |

trains, and arrives at the same time in Utrecht in all configurations.

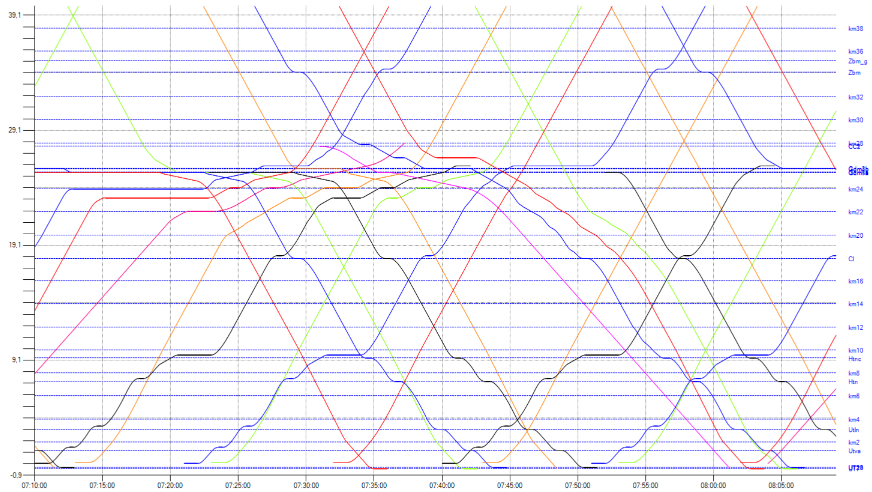
Between the ATB and the L2 existing configuration, there is hardly any difference in delay. When following the slower delayed freight train, the L2 existing configuration saves no time for the successive sprinter and only around a minute for the successive intercity, compared to the ATB configuration.

Note to this is that the platform capacity plays a restrictive role in arrival times at the final destination. This can be seen in Figure 27b at 8:00 in Utrecht, where the 'green' intercity has to wait for the delayed 'red' one to clear the platform to enter Utrecht, doing away with the headway improvement priorly realized by ETCS L2. In the L2 reduced configuration this problem does not occur as the delay of the 'red' intercity is small enough not to hinder the 'green' intercity.

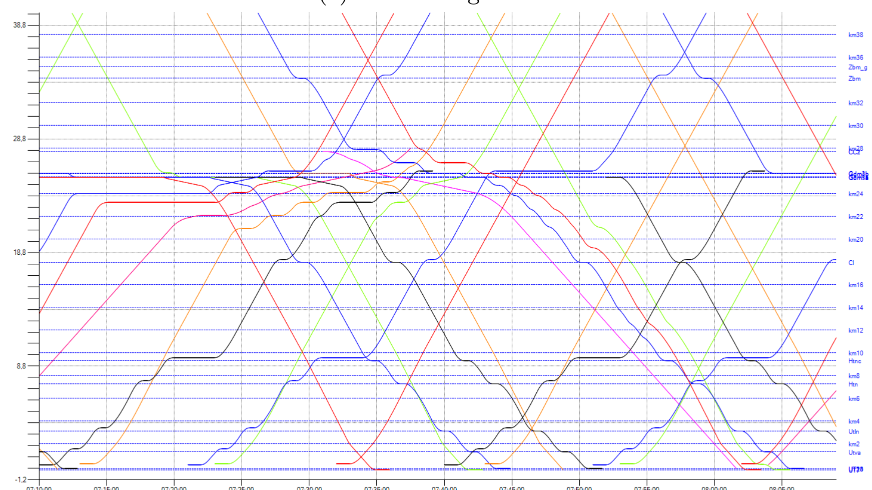
Since their is almost no improvement between the ATB and L2 existing configuration, the significant improvement of the L2 reduced configuration has to be written to the credit of the shorter blocks, not the in-cab signalling, otherwise the L2 existing should have shown further improvements too.

The improvement of the L2 reduced configuration is largely due to the closeness of the successive trains passing the level crossing at low speed. For example: the time headway after the level crossing between the first 'red' intercity and the 'black' sprinter (third train after the intercity) is 10 minutes in the L2 existing blocks, and less than 5 minutes in the L2 reduced configuration. In the ATB configuration, this is similar to the L2 existing blocks.

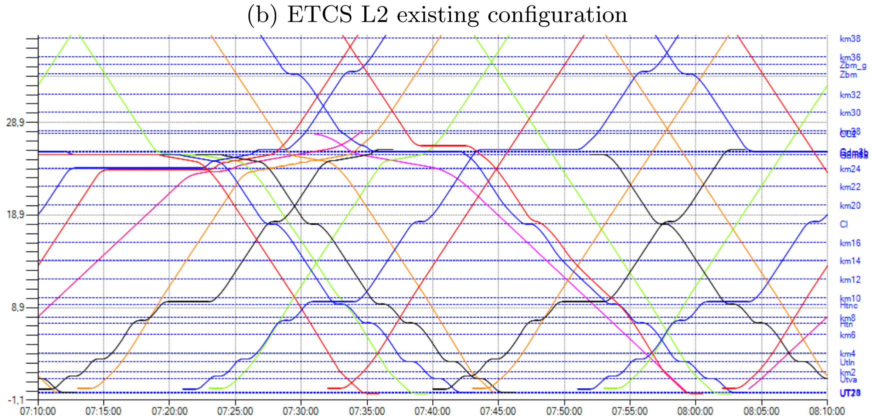
To conclude the findings of this scenario: especially the shorter blocks have been of influence to the advantages of the L2 reduced configuration.



(a) ATB configuration



(b) ETCS L2 existing configuration




(c) ETCS L2 reduced configuration

Figure 27: Time-space diagrams of scenario 5

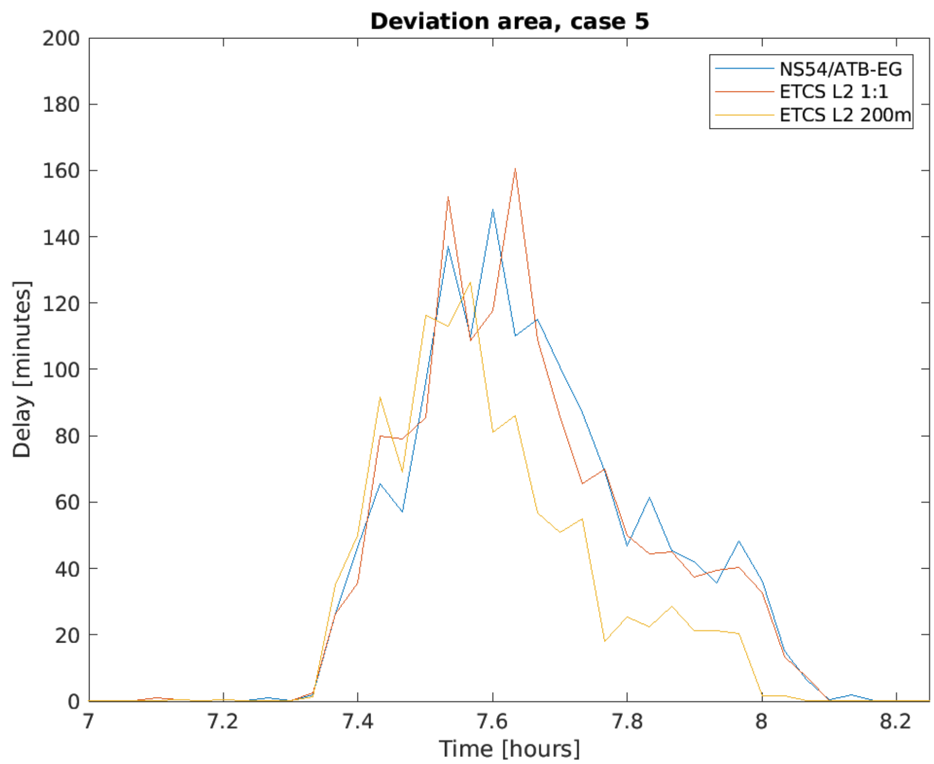



Figure 28: Deviation area of case 5

## Scenario 6 - Track circuit defect

In this scenario, a track circuit is failing. For the sake of comparison to the other configurations, it is assumed that in the ETCS L2 reduced configuration the same block length is failing as in the other two configurations.

The applied dispatching measures here are that in Geldermalsen the sprinters are sent in front on the delayed intercities, by which they normally would be overtaken.

The difference in total delay at the end stations of the ATB and the L2 existing configuration is relatively small: the delay is 3% less with the L2 existing configuration. With the L2 reduced configuration, this is 11,5% less though.

Table 18: Delay statistics of case 6

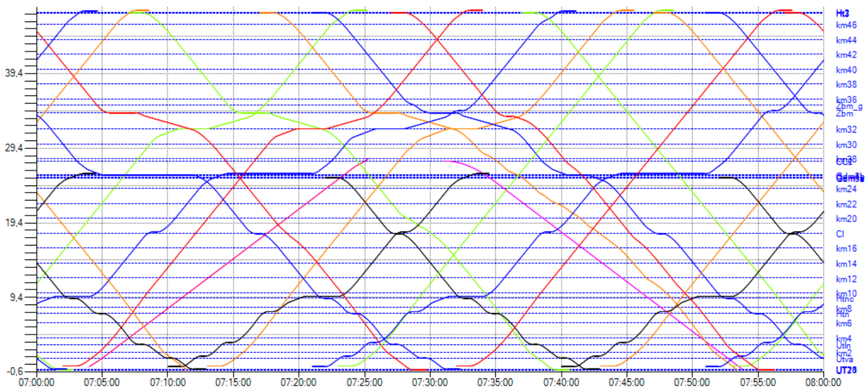
| | 3 min punct (%) | 5 min punct (%) | Recoverytime (h:mm) | $\sum$ delay (h:mm:ss) | Max delay (min) |
|---|---|---|---|---|---|
| NS'54/ATB | 96% | 96% | 0:51 | 0:53:56 | 61 |
| ETCS L2 existing | 97% | 97% | 0:51 | 0:52:27 | 67 |
| ETCS L2 reduced | 96% | 96% | 0:51 | 0:47:45 | 53 |

The first 'red' intercity that is not hindered by the disruption itself but only by other impacted trains, can follow and overtake the sprinter much closer between Zaltbommel and Geldermalsen. In the L2 reduced configuration, it barely has to brake, whilst in the ATB and L2 existing blocks, it comes to a standstill (7:35, Zbm), which in the end costs 2 minutes.

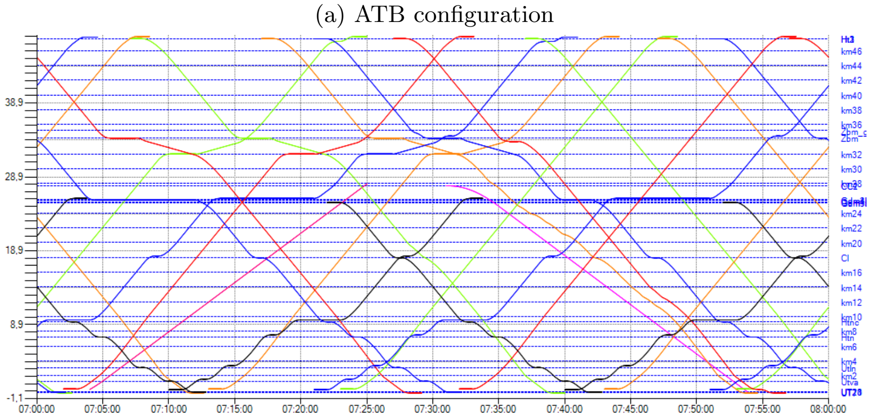
The main differences between the configurations is that in the ATB configuration, the 'red' and 'green' delayed intercities are more impacted by the sprinter sent in front of them. This can be concluded by the arrival times of these intercities in Utrecht: the arrival of 'red' intercity in Utrecht is 7:28:47 with ATB, whilst the arrival time for the L2 existing and reduced configuration is at 7:28:11. For the arrival of the 'green' intercity, the difference is larger: 7:39:40 for the ATB configuration, 7:39:03 with the L2 existing blocks, and 7:38:22 with L2 reduced blocks. The arrival times for the sprinters are the same in each configurations. Half of the benefit can be credited to the in-cab signalling and the corresponding braking curve protection, and the other half to the smaller blocks.

There are two notes to this scenario: Firstly, the delay of the 'blue' delayed sprinter is not properly measured: the buffer at Geldermalsen makes that it can recover from its delay before Geldermalsen and arrive on time in Utrecht, where its delay is measured. This would speak even more in the benefit of the L2 reduced configuration (1 minute saved). However, this benefit is also measured by the following intercity, which is the same amount of time faster as saved by the sprinter. And secondly, the platform capacity at Utrecht is again a restrictive component. This does not say, however, that is not useful to improve the headway on the open track, as the shorter headway in the ETCS L2 configurations makes that the 'blue' sprinter towards Utrecht, following the two intercities at 7:42 in Geldermalsen, is not impacted by the intercities. In the ATB configuration, it is slightly impacted.

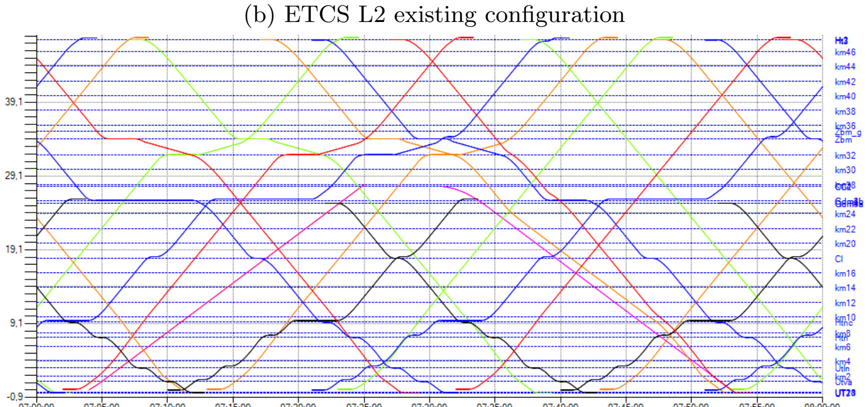
Since the ETCS L2 existing blocks configurations is not significantly better than the ATB configuration but the L2 reduced blocks one is, it has to be concluded that the smaller blocks are what makes the difference in this scenario, not the in-cab signalling.




(a) ATB configuration




(b) ETCS L2 existing configuration




(c) ETCS L2 reduced configuration

Figure 29: Time-space diagrams of scenario 6

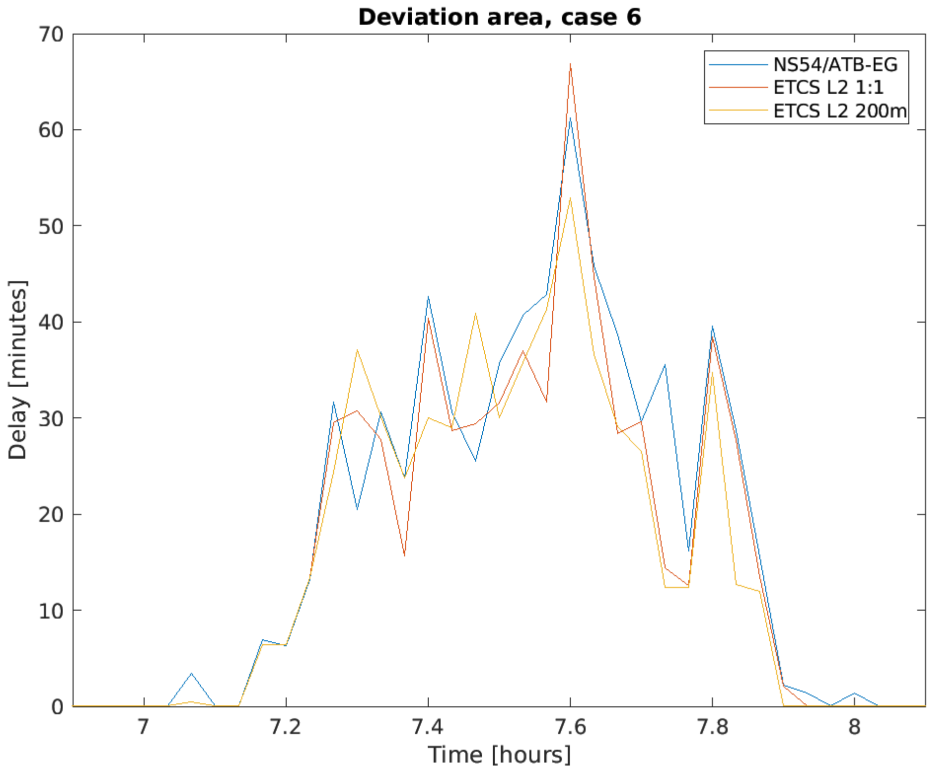



Figure 30: Deviation area of case 6

## Scenario 7 - Collision

A comparison of the delay per moment in time of the first case can be seen in Figure 31. Table 19 summarizes several performance indicators. The delays can be only be measured when trains pass timetable points, thus the measurement of the delay only starts after the disruption has been resolved and trains can drive again, 7:53 in this case.

No traffic management has been applied in this case to find pure effect of the signalling configurations when the train of a full hour are standing behind one another. In reality, several trains would have been cancelled until the disruption is (nearly) solved. Still, some trains would be stuck behind one another, having the same effect as simulated in this case, but smaller. However, the effect will become clearer by increasing the number of disrupted trains.

The main gain of ETCS L2 is achieved in the aftermath of the disruption, as the maximum delay, shortly after the blockage has been cleared, does not differ significantly between the configurations. The maximum delay is reached earlier in the ETCS L2 configurations by cause of trains being on the move earlier, thus passing timetable points quicker.
The L2 existing blocks configuration, however, causes the disruption to be resolved 15 minutes earlier than the current ATB configuration, and the L2 with the reduced blocks resolves the disruption yet another 15 minutes earlier.

Compared to the ETCS configurations, there are more delays in the ATB configuration on trains departing from Utrecht of Den Bosch after the blockage has been cleared, secondary delays, whilst the primary delays are fairly the same. This can also be seen in the sum of the delays. The total delay of the ETCS L2 reduced configuration is 29% lower than the ATB configuration in this case; the L2 existing configuration saves 12% of the total delay.

Figure 33 shows a visual comparison on how close trains can follow one another in the ETCS L2 reduced configuration and in the ATB configuration. The green train is the first train to depart after the disruption has resolved. It can be seen that the so-called 'accordion-effect' is considerably smaller: after the disruption has been resolved, approximately 2,5 trains fit in the same headway in the ETCS L2 reduced configuration as 1 train in the ATB-configuration. Therefore, it is logical that the peak of the maximum delay is reached earlier and is slightly higher, and that the total delay is that much smaller in the ETCS L2 configurations.

In the theory it was said that the main impact of ETCS L2 was mainly expected around switches and stations, but it has shown that the impact for resilience is in this case also significant on the open tracks. Still, the impact of the L2 reduced configuration could have been even bigger, would there be more platforms be available at the end stations to divide the trains. Even when alternating the slower and faster traffic, the benefits of the smaller blocks are softened. Nevertheless, the potential benefit with in-cab signalling combined with smaller block is significant in situations like these.

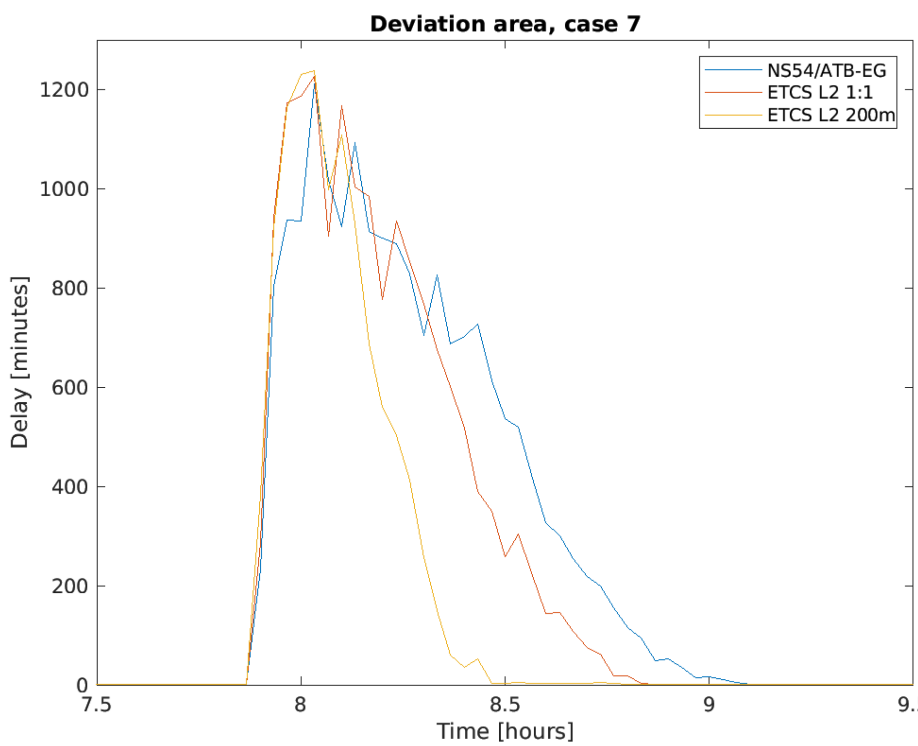
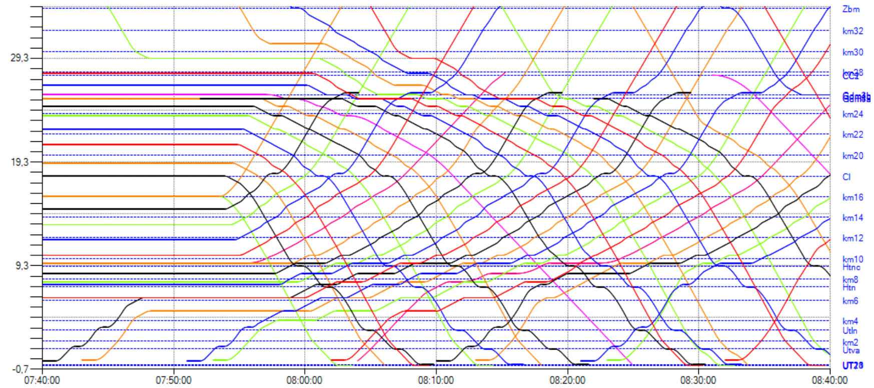
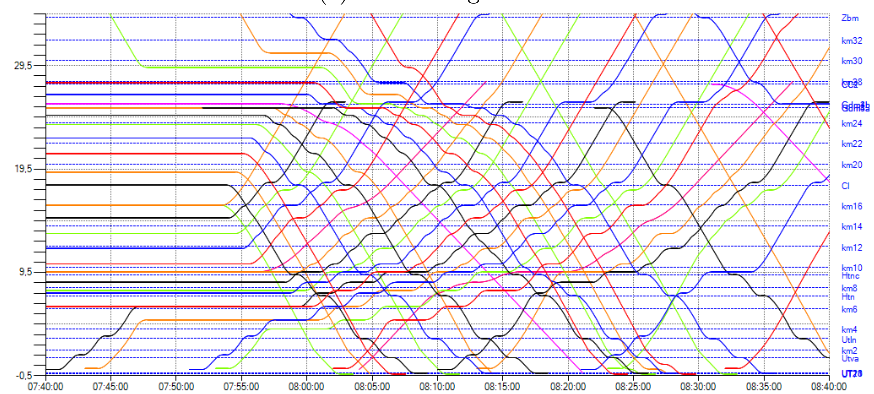


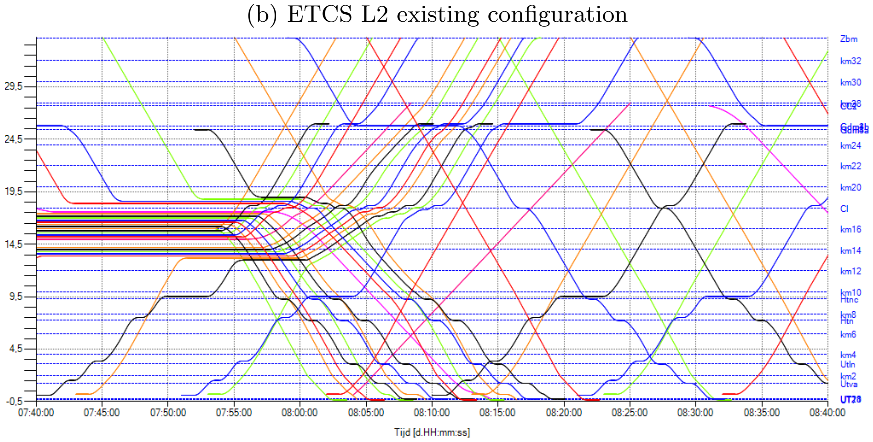

Figure 31: Deviation area of case 7




(a) ATB configuration



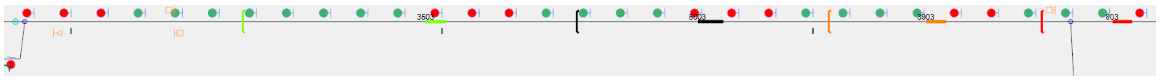

(b) ETCS L2 existing configuration




(c) ETCS L2 reduced configuration

Figure 32: Time-space diagrams of scenario 7

(a) ATB configuration



(b) ETCS L2 reduced blocks configuration

Figure 33: Difference in the headway between the configurations, shown visually

Table 19: Delay statistics of case 7

| | 3 min punct (%) | 5 min punct (%) | Recoverytime (h:mm) | $\sum$ delay (h:mm:ss) | Max delay (min) |
|---|---|---|---|---|---|
| NS'54/ATB | 72% | 73% | 1:13 | 22:01:03 | 1210 |
| ETCS L2 existing | 75% | 77% | 0:59 | 19:24:56 | 1225 |
| ETCS L2 reduced | 85% | 85% | 0:45 | 15:42:41 | 1238 |

## 5.2. Results

Table 20 gives an overview of how quickly the delay has been decreased to less than 5 minutes in each configuration for every scenario. In Table 21, a summary is shown of the potential gains of ETCS L2 in each scenario on the most significant and clarifying performance indicator, the sum of the delay of each train at its final destination.

Table 20: Overview of how quickly the disruption has resolved in each configuration per scenario (h:mm)

| # | NS'54/ATB | L2 existing | L2 reduced |
|---|---|---|---|
| 1 | 0:55 | 0:45 | 0:43 |
| 2 | 0:46 | 0:40 | 0:30 |
| 3 | 1:06 | 1:06 | 1:06 |
| 4 | 1:05 | 1:03 | 1:01 |
| 5 | 0:49 | 0:49 | 0:43 |
| 6 | 0:51 | 0:51 | 0:51 |
| 7 | 1:13 | 0:59 | 0:45 |

Table 21: Overview of how much delay is saved compared to the ATB configuration per scenario

| # | NS'54/ATB | L2 existing | L2 reduced |
|---|---|---|---|
| 1 | 100% | -9.1% | -12.9% |
| 2 | 100% | -13.6% | -25.1% |
| 3 | 100% | -19.2% | -19.3% |
| 4 | 100% | -9.0% | -12.6% |
| 5 | 100% | -2,2% | -27.9% |
| 6 | 100% | -3.0% | -11.5% |
| 7 | 100% | -11.8% | -28.6% |

The outcomes of the total delays of the scenarios are in a wide range, from -2,2% to -19% for the L2 existing blocks and from 12% to 28% for the reduced blocks. This is logical as the scenarios have been varied in length and severeness. However, although the large differences, they do indicate that significant reductions of delay can be achieved by firstly upgrading to ETCS Level 2, and secondly by reducing the block sections for train detection. This is in accordance to the hypothesis in the theory that both of the signalling characteristics would increase the resilience.

Now that all scenarios have been modelled, the resilience index for both the ETCS L2 configurations can be calculated. There are seven scenarios, and the ATB configuration is the reference configuration for the calculations:

$$Resilience\ Index = \frac{7}{\sum_{n=1}^{7}\left[\frac{\sum_{t=1}^{T} Final\ Delay_{t,L2}}{\sum_{t=1}^{T} Final\ Delay_{t,ATB}}\right]_n}$$

For the ETCS L2 with existing blocks configuration, the resilience index is 1.10. For the L2 reduced blocks configuration, the resilience index is 1.25, as shown in Table 22. This means that with the L2 existing configuration 9% ( 1-(1/1.1) ) of the delay has been saved in the modelled scenarios on average, and the L2 reduced one has saved 20% ( 1-(1/1.25) ) over all scenarios.

Table 22: Resilience index of the three configurations

| NS'54/ATB | ETCS L2 existing | ETCS L2 reduced |
|---|---|---|
| 1 | 1.10 | 1.25 |

In scenarios 1, 3 and 4, the largest part of the reduction of the delay, compared to the ATB-configuration, was already ensured by the in-cab signalling part. The small blocks in the ETCS L2 reduced blocks configuration have only added a small extra reduction on top of the in-cab signalling. In the other four scenarios, on the other hand, the smaller blocks were responsible for a major part of the reduction, compared to ATB-configuration. In some situations it is thus already quite beneficial to change from line-side signals to in-cab signalling, but that even more benefits can be gained by reducing, or at least optimising, the block section lengths.

From the simulations, it has shown that especially in short speed-restricted disruptions, such as switching over between tracks back and forth, the in-cab signalling part of ETCS L2 can make a difference. So, the more heterogeneous the train traffic, the more the traffic would have to brake in disrupted situations, and the more the traffic would profit from ETCS L2. Also, the short block sections are of great importance around switches and station areas, which are situations where traffic merges and splits or where the headways between trains become small. On the open track, blocks of 200 meter will be superfluous for resilience with respect to the associated cost, since the open tracks is not the most restrictive bottleneck.

However, in-cab signalling does not automatically mean time gain in all situations. In decelerating situations it does, but in accelerating situations, e.g. driving through switches with 60 km/h and thereafter 130 km/h on the open track, it is not faster than how ATB works. Yet, in-cab signalling is safer in a situation like that, as a driver cannot accelerate when the last coaches are still in the speed restricted area, which is possible with ATB.

A conclusion for traffic dispatching is that with very small blocks, or later on, moving blocks, it is not advised to send fast and slow train traffic in groups, but to send them alternately. Most likely, capacity is restricted somewhere further on in the network, for example at a platform, making it superfluous to send traffic in pairs. So, unless the trains can also be handled with short headways upstream, the total delay in the network is likely to be reduced more by alternating the traffic.

In-cab signalling in combination with small blocks has shown to lead to a significant increase of resilience in these scenarios. This strengthens the case of adjusting the current blocking lay-out to smaller blocks, when converting to ETCS, increasing the capacity as well as the resilience.

However, this always needs to be weighed against the costs of increasing the number of block sections. Designers can either use the results of this research or an appliance of this method to the concerned design area to underpin their design choices. The method used is also useful to assess the benefits of spatial dispatching possibilities, such as side-tracks and switches. The delay saved for passengers and freight, due to the increased resilience, then can be weighed off against costs and other criteria.

### 5.2.1. Discussion

How the train traffic is handled, and can be handled, at the end of this corridor has shown to be of importance to the results. Firstly, it has been of influence on which dispatching strategies should be taken, and secondly on how shortly trains can be sent out of the corridor after each other. If short headways cannot be handled further upstream, then there is no need to make short blocks on this corridor. During the build-up of this research, most of the focus has been on the right simulation of all (spatial) dispatching possibilities 'in the middle' of the network, and less on the boundaries of the network. For example, it may be that in Utrecht and Den Bosch, the intercities can be sent to another platform if the normal platform is occupied, but this has not been taken into account.
Furthermore, now only this one line has been considered, leaving few rescheduling possibilities, so little traffic management was possible. If a larger network would have been considered then there would have been more dispatching possibilities. Then, also more distinctive dispatching possibilities between the safety systems potentially, now there was little space for different dispatching measures.

This was a very busy corridor with small headways, where each small delay almost meant directly a delay for a successive train. The potential benefit from these scenarios is likely to be only applicable for other busy lines where the headways are around or below the 5 minutes. On corridors with less traffic and larger headways, and thus more buffers, there will still be gain, but probably not as much as found in this research.

Moreover, the results are dependent on the time and location of the disruption, the length of the trains, which trains has been chosen to be disrupted, the timetable on the corridor, all of which have been chosen on own insight.

A configuration with ATB and the smallest blocks possible therewith would also have been interesting to include in the simulation. This configuration would have given the opportunity to estimate the difference between the signalling characteristics even better.

The time-space diagram of an intercity having to follow a slower freight train, shown in Figure 34, shows that the driving behaviour of the drivers in Xandra is somewhat unrealistic. It assumes that trains drive at full power. In reality, a driver would not accelerate and brake so much, but he/she would maintain a lower, constant speed, which is better for passenger comfort as well as for energy efficiency. Especially with the current so-called driver assistance systems, drivers can see the traffic before and behind them, and can anticipate better on which speed to maintain for an optimal traffic flow. This deviation makes that trains are somewhat faster in the simulation than they would be in reality. This deviation from reality is present in all three configurations, so it is assumed that it has no influence on the comparison between the configurations.

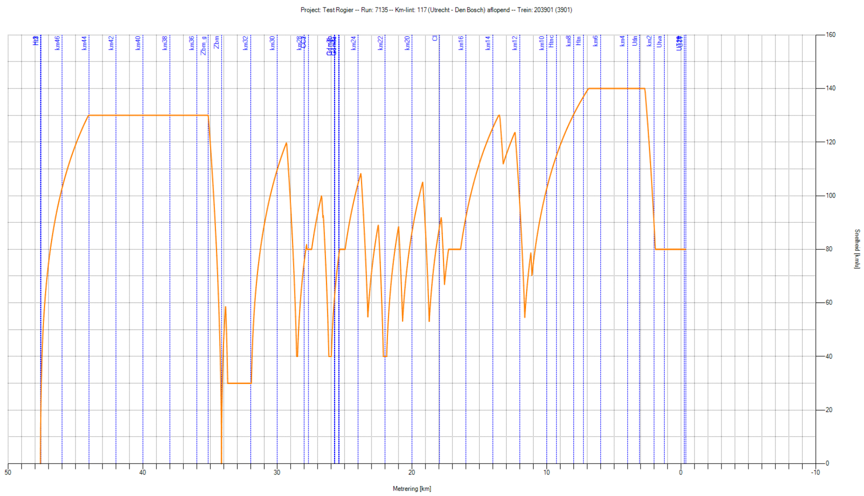



Figure 34: A time-speed diagram showing unrealistic driver behaviour

# Chapter

# 6. Conclusion

Currently, the railway signalling system ETCS is being implemented Europe wide. This is expected to bring benefits with respect to safety, capacity and interoperability. Resilience is an factor that is likely to be increased with the coming of ETCS. Resilience is the ability of a system to recover from a disruption given a predetermined control management plan. The design of the signalling system affects the resilience, next to the available infrastructure, the timetable and the quality of the contingency plans. This research focuses on the resilience of railway signalling characteristics and its impact on the design and implementation of future ETCS projects.

The main research question of this research has been the following:

***What is the influence of railway signalling system characteristics on resilience?***

This question can be split into several components which all consider a part of the main question. The main research question is therefore subject to the following sub-questions:

- *How does resilience influence the implementation and design of ERTMS in the Netherlands?*
- *Which signalling system characteristics have effect on resilience?*
- *What difference in resilience does ETCS level 2 have compared to NS'54/ATB-EG?*

## 6.1. Key findings

Literature review has been used (1) to identify the proper performance indicators for resilience, (2) to find the relevant signalling characteristics with respect to resilience, (3) to see the role of simulation in decision making, and (4) to find common disruptions and disruption management strategies in the Netherlands. Interviews and expert knowledge have been used (i) to find the role of resilience in the Dutch ETCS-decision making and designing process, (ii) to create a power-interest diagram with respect to resilience, and (iii) to discuss the set-up of simulation, the disrupted scenarios and the applied dispatching strategies.
The simulation tool Xandra, developed within Arcadis, has been used to simulate seven disrupted scenarios over three signalling configurations, the current NS'54/ATB signalling lay-out, ETCS L2 using the existing block sections, and ETCS L2 with reduced block section lengths. The corridor Utrecht - Den Bosch and the related traffic have been modelled using microscopic, deterministic simulation.

The following conclusions have flown from the literature review, interviews and simulations:

- The resilience of the signalling system is just one of the many aspects to be considered in the infrastructural decision process of ERTMS. More important criteria are safety, the need for replacement, interoperability and cost;
- Increased resilience is most interesting for operators and passenger organizations; the power, on the other hand, lies in the hands of ProRail and engineering firms, as they are the ones responsible for the design of the lay-out of the signalling and infrastructure;
- There are three levels in the design phase of ETCS where resilience can be of support: full comparison between alternatives on all factors, including resilience; comparison between alternatives, but only on resilience; and to find the relation between design choices and resilience, based on expert judgement of the results of a resilience study;
- Two signalling characteristics are of relevance for resilience: the placement and length of block sections and whether the movement authority of a train is provided to the driver via line-side signals or via in-cab signalling;
- The modelled disruption scenarios has shown that compared to the current Dutch NS'54/ATB signalling configuration, an ETCS L2 configuration that uses the same block distances is 10% more resilient; An ETCS L2 configuration with reduced block section lengths of 200 meters is up to 25% more resilient than the current NS'54/ATB signalling configuration. The effect of reducing the block sections lengths in ETCS L2 to an approximation of the moving-block principle is thus 13%;
- With very short block section lengths, the potential benefits is restricted by factors other than the signalling component, such as platform capacity or the speed difference between trains.

## 6.2. Context

This research has indicated that with the implementation of ETCS L2, the delays on a railway network can be reduced significantly. This can in the first instance already be achieved by overlaying the current signalling configuration, by using the same block lengths for train detection, leading for a reduction of delays up to 10%. Then, by decreasing these block lengths, an even larger reduction of the delays can be realized, up to 20%. This holds for corridors that have a similar intensity as the corridor Utrecht-Den Bosch and similar small headways between services.

These results are consistent with the findings of Goverde et al. (2013) that there is a significant benefit for ETCS L2 in disturbed situations. This research has proven that this benefit is also valid and significant for disrupted situations.

However, when decreasing the block lengths to an approximation of Level 3 moving block principle, other factors will restrict the potential reduction of delays. Platform capacity is one of these factors that can undo the reduction. The headways on the open track, realized by the small blocks, therefore does not need to be smaller than what can be handled at junctions and stations in the network.

The results of this research are related to the characteristics of the simulated corridor and set of disruptions, and thus cannot be generalized for other lines. Nonetheless, they can be used as an indication for the potential benefit of ETCS L2 on a line. On lines that are less busy than the corridor Utrecht - Den Bosch, the benefit of ETCS L2 will be smaller. On the other hand, it is likely that the more heterogeneous the traffic on a corridor, the larger the benefits will be. To get a better indication of this, either more disrupted scenarios should be simulated or the current scenarios could be applied to other corridors with different characteristics. For an even better comparison between the relevant signalling characteristics, it would have been useful to have included an NS'54/ATB configuration with the blocks as small as possible.

No data on the exact duration of real-life disruptions was available to translate the potential benefits of ETCS L2 on resilience to monetary terms. This would also require additional research on the value of time of passengers during delays, which has been out of the scope of this research.

## 6.3. Recommendations

Up to 10 or even 20% of passengers delay can be saved by installing ETCS L2, and by reducing the block sections lengths. This means that installing ETCS L2 not only increases the safety and capacity, but also helps to reduce the effect of disruptions, thus cutting down the costs for operators and increasing passengers satisfaction.

Although it is not required to take resilience into account since many other factors are of importance too, it is advised to include it to help and clarify the decision making or designing process. The burden of proof regarding resilience can be reversed from proving the benefits or resilience towards proving the extra costs. Operators can use these findings to calculate the estimated reduction of cost and increased passenger satisfaction when having shorter disruptions, and strengthen their negotiation point to have a more resilient infrastructure. Then, infrastructure managers will have to show whether this outweighs the cost of the extra infrastructure or not.

In the design phase of any future ETCS L2 lay-out, there can be a more transparent trade-off between the costs for extra infrastructure or signalling components, and the benefits that come with increased resilience. Although not one optimal signalling lay-out for resilience can be determined by this simulation approach, as this is too dependent on the location, length and impact of the disruption chosen in this research, it has shown that shorter block sections around switches and stations are beneficial for the resilience in each scenario. Implementing reduced block sections along the whole corridor, which provides maximum resilience as well as capacity, is not advised, since it is not likely that the benefits of this will outweigh the costs of designing and implementing the extra blocks.

## 6.4. Further research

In this research, only simple traffic management has been applied, and even the same strategy in the ATB and ETCS L2 configurations to give a fair comparison. Research can or must be done on how traffic management will change when ETCS L2 will be implemented on more parts of the Dutch network. For this, a larger network should be simulated than this one corridor, as then more distinctive dispatching possibilities can be simulated.

It was beyond the scope of this research to find an optimal blocking lay-out strategy for the ETCS L2 configuration. However, this is interesting for further research to help the design of future ETCS L2 routes. It could be done with the simulation approach used in this research. In the ETCS L2 reduced blocks configuration, per case one could remove the blocks that have no influence on the outcome of the simulation of that case. Doing this for several cases may give an optimal blocking configuration strategy with respect to resilience.

This simulation approach could also be used to find which spatial dispatching possibilities, such as switches and side tracks, are relevant for the resilience of the train traffic and which are not and thus may be removed.

Further research is also recommended on the effect of combining ETCS with Automatic Train Operation. Research may be done to see what blocking distance strategy is most optimal for ATO over ETCS, what the expected capacity and robustness increase is with ATO, how the standards for headways between trains can change, how traffic management can or should change, and how the (specification and requirements for) the on-board and track-side systems are affected.

# References


Albrechts, L. (2003). Reconstructing decision-making: Planning versus politics. *Planning theory*, *2*, 249–268.

Aven, T. (2010). On how to define, understand and describe risk. *Reliability Engineering & System Safety*, *95*, 623–631.

Baroud, H., Ramirez-Marquez, J. E., Barker, K., & Rocco, C. M. (2014). Stochastic measures of network resilience: Applications to waterway commodity flows. *Risk Analysis*, *34*, 1317–1335.

Barter, W. (2008). Ertms level 2: Effect on capacity compared with\“best practice” conventional signalling. *WIT Transactions on The Built Environment*, *103*, 213–222.

Binder, S., Maknoon, Y., Sharif Azadeh, S., & Bierlaire, M. (2015). *Generation and evaluation of passenger-oriented railway disposition timetables in case of severe disruptions*. Technical Report Transport and Mobility Laboratory, Lausanne.

Blenkers, L. (2015). *Railway disruption management*. Ph.D. thesis TU Delft.

Borning, A., Waddell, P., & Förster, R. (2008). Urbansim: Using simulation to inform public deliberation and decision-making. In *Digital government* (pp. 439–464). Springer.

Brömmelstroet, M. t., & Bertolini, L. (2011). The role of transport-related models in urban planning practice.

Chen, L., & Miller-Hooks, E. (2012). Resilience: an indicator of recovery capability in intermodal freight transport. *Transportation Science*, *46*, 109–123.

Coenraad, W. (2008). Intermittent and continuous automatic train protection. *IRSE*, .

Corman, F., D’Ariano, A., & Hansen, I. A. (2014). Evaluating disturbance robustness of railway schedules. *Journal of Intelligent Transportation Systems*, *18*, 106–120.

Corman, F., Quaglietta, E., & Goverde, R. M. (2018). Automated real-time railway traffic control: an experimental analysis of reliability, resilience and robustness. *Transportation planning and technology*, *41*, 421–447.

Dicembre, A., & Ricci, S. (2011). Railway traffic on high density urban corridors: capacity, signalling and timetable. *Journal of Rail Transport Planning & Management*, *1*, 59–68.

ERA (2016). Introduction to etcs braking curves. URL: `https://www.era.europa.eu/activities/european-rail-traffic-management-system-ertms_en`.

ERTMS.be (2017). Benefits of axle counters over track circuits. `http://www.ertms.be/sites/default/files/2018-03/16E0421_HL3.pdf/`. Accessed: 2019-05-20.

Ferrari, A., Spagnolo, G. O., Martelli, G., & Menabeni, S. (2012). Product line engineering applied to cbtc systems development. In *International Symposium On Leveraging Applications of Formal Methods, Verification and Validation* (pp. 216–230). Springer.

FERSIL (2019). Benefits of axle counters over track circuits. `https://www.fersil-railway.com/en/systems-softwares/train-detection/sil4-axle-counters/`. Accessed: 2019-05-17.

Fisher, L. (2015). Disaster responses: More than 70 ways to show resilience. *Nature*, *518*, 35.

Freeman, R. E., & Reed, D. L. (1983). Stockholders and stakeholders: A new perspective on corporate governance. *California management review*, *25*, 88–106.

Goverde, R. (2018). Block signalling [cie4872 powerpoint slides].

Goverde, R. M. (2012). Robuust spoor met ertms. In *Colloquium Vervoersplanologisch Speurwerk CVS, Amsterdam, Nederland, 22-23 Nov. 2012*. CVS.

Goverde, R. M., Corman, F., & D'Ariano, A. (2013). Railway line capacity consumption of different railway signalling systems under scheduled and disturbed conditions. *Journal of rail transport planning & management*, *3*, 78–94.

Goverde, R. M., & Hansen, I. A. (2013). Performance indicators for railway timetables. In *2013 IEEE International Conference on Intelligent Rail Transportation Proceedings* (pp. 301–306). IEEE.

Haimes, Y. Y. (2009). On the definition of resilience in systems. *Risk Analysis: An International Journal*, *29*, 498–501.

Hansen, I. A. (2010). Railway network timetabling and dynamic traffic management. *International Journal of Civil Engineering*, *8*, 19–32.

Hansen, I. A., & Pachl, J. (2014). Railway timetabling & operations. *Eurailpress, Hamburg*, .

Hofman, M., Madsen, L., Groth, J. J., Clausen, J., & Larsen, J. (2006). Robustness and recovery in train scheduling-a simulation study from dsb s-tog a/s. In *University of Denmark*. Citeseer.

Kavicka, A., & Klima, V. (2000). Simulation support for railway infrastructure design and planning processes. *WIT Transactions on the Built Environment*, *50*.

Koivisto, M. (2017). Pitfalls in modeling and simulation. *Procedia computer science*, *119*, 8–15.

Nicholson, G. L., Kirkwood, D., Roberts, C., & Schmid, F. (2015). Benchmarking and evaluation of railway operations performance. *Journal of Rail Transport Planning & Management*, *5*, 274–293.

ON-TIME (2014). Methods and algorithms for the development of robust and resilient timetable. URL: `http://www.ontime-project.eu/deliverables.aspx`.

ProRail (2019a). Ertms. URL: `https://www.ertms-nl.nl/default.aspx`.

ProRail (2019b). Over punctualiteit. `https://www.prorail.nl/over-punctualiteit?` Accessed: 2019-05-17.

Quaglietta, E. (2014). A simulation-based approach for the optimal design of signalling block layout in railway networks. *Simulation Modelling Practice and Theory*, *46*, 4–24.

Slootjes, E. G. (2013). *Vergelijkend onderzoek naar ERTMS invoering en gebruik scenario's*. Master's thesis University of Twente.

Vromans, M. J., Dekker, R., & Kroon, L. G. (2006). Reliability and heterogeneity of railway services. *European Journal of Operational Research*, *172*, 647–665.

Xu, P., Corman, F., & Peng, Q. (2016). Analyzing railway disruptions and their impact on delayed traffic in chinese high-speed railway. *IFAC-PapersOnLine*, *49*, 84–89.

# 7. Appendices

## 7.1. Appendix A- Paper

# The influence of railway signalling characteristics on resilience

Rogier Simons, Master student, Transport Infranstructure and Logistics, TU Delft

***Abstract*—Seven disrupted scenarios have been simulated on the railway corridor Utrecht-Den Bosch, each over three signalling configurations to find the effect of in-cab signalling and reduced block section lengths on resilience. Furthermore, the place of resilience in decision making and design of railway signalling has been investigated.**

**Although resilience is only one of the factors in the decision making and design of ETCS-projects, besides factors such as capacity, safety and interoperability, a quantification of resilience can help to either compare alternatives on all factors including resilience, to compare alternatives on resilience only, or to find the relation between design choices and resilience.**

**Simulation of seven disrupted traffic scenarios has shown that using in-cab signalling with ETCS L2 compared to line-side signalling with NS'54/ATB has saved 10% of the delay in the scenarios on average. By using reduced block sections lengths in combination with in-cab signalling, 20% of the delay has been saved on average. For absolute resilience, short block sections should be placed along the whole track. As this is not realistic from cost perspective, decreased block sections are in each case advised nearby yards and switches.**

***Index Terms*—ETCS, ATB, simulation, disruptions**

## I. Introduction

VARIOUS legacy protection systems in Europe have been developed mid-previous century and are due for replacement in the coming years. The European standard for railway signalling and communication, ERTMS, is the designated standard to replace these legacy systems. The signalling and control component of ERTMS is ETCS, besides components for the communication and management. ETCS is specified at four levels, in which the detection and signalling technique gradually move from the track to the train. Further explanation on the ETCS levels can be found in [1].

In the Netherlands, ETCS Level 2 is the standard level when ETCS is being implemented. At this level, train detection is still on the track-side, as with most legacy protection systems, but the movement authority is now communicated to the driver via in-cab signalling instead of line-side signals. ETCS has already been implemented on some corridors, and in the coming years, several other corridors where the NS'54/ATB is at the end of its lifetime and needs to be overhauled, ETCS will be installed.

The discussion about the planning and locations of the replacements has grown to a political and organizational debate, as a result of the amount of capital at stake and the long time span of the replacement. Now, seven corridors have been designated to be overhauled before 2030. In the decision making of implementing ETCS, several factors have been of relevance, such as the need for replacement on a corridor, the cost of the replacement, the capacity and safety benefits, and international interoperability. On the more practical side, during the design and engineering phase of an ETCS-corridor, most of the same factors are of importance on how to convert from the legacy system to ETCS. For example, a larger budget can lead to a design where the block sections lengths can be adapted, creating more capacity and resilience.

Resilience is another factor which is potentially relevant in the decision making. Resilience is the ability of a system to recover from a disruption given a predetermined control management plan.

Both in the decision making process as in the design process of ETCS, resilience is of importance: firstly on high level decision making, of where and when to convert from the legacy system to ETCS. If it is demonstrated that the latter has a large positive impact on resilience, this is in favor of converting to ETCS sooner or on more corridors. Secondly, on the level of designing the lay-out of the signalling and infrastructure. Namely, in the design of the signalling lay-out, different objectives can lead to friction between the operators and the infrastructure managers. An infrastructure manager wants less infrastructure to reduce maintenance and the amount of disruptions, meanwhile operators would like more infrastructure for their operations. A design focused on resilience can save delay for passengers during disruptions, thus saving money, and increasing passenger satisfaction.

To allow comparison of alternative designs on this factor, resilience needs to be quantified. If a method can be provided to evaluate both the cost and the benefits of a more resilient infrastructure/signalling in monetary terms, then resilience can be weighed against the other factors involved, such as cost.

This has led to the following main research question:

***What is the influence of railway signalling system characteristics on resilience?***

This questions is subject to several sub-questions, to split the main question in multiple components:

- *How does resilience influence the implementation and design of ERTMS in the Netherlands?*
- *Which signalling system characteristics have effect on resilience?*
- *What difference in resilience does ETCS level 2 have compared to NS'54/ATB-EG?*

Thesis to be defended on 25 november 2019 at the TU Delft

## II. Methodology

To answer the research questions, several distinctive methods have been applied.

Firstly, a literature review has been used to identify the proper performance indicators for resilience, to find the relevant signalling characteristics with respect to resilience, to see the role of simulation in decision making, and to find common disruptions and disruption management strategies in the Netherlands. Secondly, interviews and expert knowledge have been used to find the role of resilience in the Dutch ETCS-decision making and designing process, to create a power-interest diagram with respect to resilience, and (iii) to discuss the set-up of simulation, the disrupted scenarios and the applied dispatching strategies.

And lastly, simulation has been used to find the impact of signalling characteristics on resilience. The corridor Utrecht - Den Bosch and the related traffic have been modelled using microscopic, deterministic simulation. Microscopic, as this provides accurate modelling of the signalling and interlocking, and deterministic, so that all other factors, such as rolling stock, timetable and running times can be controlled for. A suitable tool for this purpose is the simulation tool Xandra, developed within Arcadis. This tool has been used for the simulation of the case studies.

Seven distinctive disrupted scenarios have been set up such that they vary in duration, location and severeness, to get a representative outcome. Typical disruptions in the Netherlands and more specific to this corridor, have been distinguished via historic data on disruptions. These scenarios are then modelled over three signalling configurations to find the difference in resilience between these configurations, and hence the influence of the signalling characteristics.

Besides the typical disruptions, four types of data are needed to set up the simulation: (i) infrastructure data, (ii) interlocking data, (iii) rolling stock characteristics and (iv) scheduled operation data.

The outcome of the simulation, the delay of trains, has been processed using the computing environment MATLAB.

This approach is the most valid to find the impact of signalling characteristics on resilience since other methods are not able or suitable to capture this effect. A mathematical approach does not suffice since the effects happening in a disrupted situation that are hard to capture mathematically, opposed to that of an undisturbed operation. Comparison of historic delays at ETCS corridors and legacy system corridors would require an extensive amount of disruptions to find the pure effect of the signalling system.

## III. Literature review

### A. Resilience

Several definitions of resilience are being used in different sectors. They all have one factor in common: the mitigation of the impact and the focus to return to the original state In the railway sector, this mitigation is specified as rescheduling actions [2] & [3]. Resilience is for this researched defined as *the ability of a system or timetable to recover from a disruption given a predetermined control management plan*.

### B. Prior work

One of the studies that this research builds upon is the one from Goverde, et al. (2013) [1], who made a capacity assessment of a Dutch corridor of different signalling configurations under both normal and disturbed railway conditions. Their results showed that in delayed operations, there is a considerable gain for ETCS compared to NS'54/ATB, since the braking distances decrease when delayed trains run at lower speeds, having a stabilizing effect on headway times, delay propagation and throughput. This research will focus on larger disruptions, instead the smaller disturbed situations in [1]. It is shown how shortening the block sections can reduce the headway and the energy consumption on that line, but that it comes at extra costs [4]. It shows the force field of interests when designing a signalling lay-out. In an analysis of technical railway characteristics, it was found that the influence of block section lengths is higher for homogeneous train traffic and lower for heterogeneous traffic [5]. Furthermore, they found that reduced sections lengths around stations become more useful when the number of trains with stops at these stations is increased.

### C. Performance indicators

Several performance measures have been suggest for resilience [6]: punctuality, average or maximum secondary delay, as well as average track occupation. Another three measures of resilience have been suggested, captured in a deviation area diagram. A deviation area diagram graphically shows three performance indicators: (i) the maximum deviation during a time period, meaning the maximum of the sum of the delays per moment in time, (ii) the time to recover to an acceptable threshold and (iii) the deviation area, or the sum of the delay of all trains [7].

On top of the found performance indicators, this research introduces a new indicator for resilience, the Resilience Index (RI), which is the ratio between the total train delay of a given signalling system over all scenarios, and the delay of a benchmark signalling system, multiplied by the number of scenarios.

$$ResilienceIndex = \frac{N}{\sum_{n=1}^{N}\left[\frac{\sum_{t=1}^{T}\textit{Final Delay}_{t,SignSys}}{\sum_{t=1}^{T}\textit{Final Delay}_{t,RefSys}}\right]}$$

where *N* is the number of scenarios, *T* is the number of trains, *Final delay* is the delay of a train at its final destination, and *SignSys* is the given signalling system which is compared to the reference signalling system, *RefSys*. The resilience index of the reference configuration is set to 1 by multiplying the ratio of the delays by the number of scenarios. An index larger than 1 indicates a more resilient signalling configuration, and an index smaller than 1 a less resilient configuration.

Taking the reciprocal of the resilience index gives the average percentage of delay a configuration can save compared to the base configuration. This delay saving can be used to translate the benefit of a configuration for an economic evaluation and identify the corresponding monetary savings.

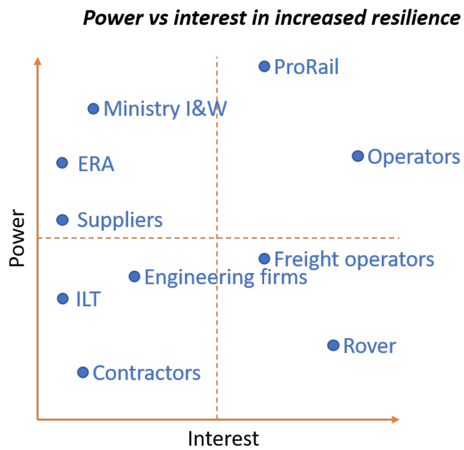



Fig. 1. The power vs interest of stakeholders in increased resilience

### D. ETCS implementation and design process

A stakeholder based approach has been used to see which stakeholders have most interest in an improved robustness of the train traffic and what their respective power is in the decision making process. It can then be seen which stakeholders benefit most from increased resilience.
The Dutch ERTMS program has divided the 190 stakeholder organizations into several groups: Operators, governments, companies in the harbour, passenger organizations (e.g. Rover), infrastructure managers (ProRail), European instances, market parties, such as contractors, suppliers, and engineering firms, and inspectors (e.g. ILT).

Each of these groups have been placed in a power/interest-diagram, with help of an interview and expert knowledge [8], shown in Figure 1. It shows how much power a group has in the decision making process versus its interest in increased robustness of the train traffic.

Research into resilience may be of most importance to NS, ProRail and engineering firms, and not so much to the ministry, as their focus is more on safety and replacing old infrastructure than on robustness. For NS and ProRail, however, it will be interesting to find out the effect of ETCS on the resilience of their timetable and their contingency plans. For engineering firms, quantification of resilience may be of influence to how they will design future signalling lay-outs.

A method to show the relation between resilience and signalling characteristics can help designers and decision makers on several levels, based on the need and the available information. Three levels can be distinguished:

1) Full comparison: fully developed method in which alternatives can be compared between themselves as well as with other factors. The benefits and costs can be expressed such that non-experts are able to include resilience in the decision-making.
2) Comparison between alternatives: results are complex to translate to monetary terms in the method and can therefore only be compared among themselves but not with other factors.

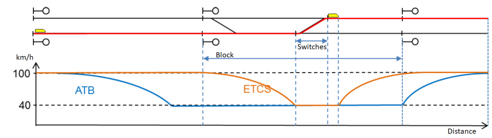



Fig. 2. Comparison of braking curve between ATB and ETCS

3) Expert judgement needed: the method is to be used as a tool to find the relation between alternatives. The results are not convertible to monetary terms and decisions have to be based on expert judgement of the results.

### E. Signalling characteristics

A railway safety system can normally be divided into four sub-components [1]:

(i) *Automatic train protection*, which supervises the speed and braking curves of a train. The communication and supervision can be intermittent or continuous, or a combination. Continuous systems have shown to reduces ATP capacity penalties and to improve headway. Another disadvantage to intermittent systems is that a train is being restricted to the braking curve imposed by a distant signal. The way the braking curves is supervised, varies per signalling systems. In NS'54/ATB, the driver has to start braking right at the signal with the 'yellow' aspect and continue at 40 km/h up to the red signal. In ETCS, however, the train calculate its braking curve up to the end of the movement authority and is thus able to postpone the braking until necessary.

(ii) *Movement authority communication*: Related to the ATP is whether the movement authority is communicated to the driver via line-side signals or via in-cab signalling. Driving with optical signals implies that the location where to start braking is fixed for all trains, while the initial speed is adapted in order to match the braking performance of the train, or in other words, poor braking trains must drive slower to still have the same block length [9]. From the point of train separation, the essential benefit of a system with cab signalling compared with a system with line-side signalling is the independence of the cab signals from the approach distance of the line-side signal system, which allows trains to run a higher speeds locally [10].

(iii) *Train detection*: On the track-side there are two main methods to detect the presence of a train in a section and separate trains, namely via track circuits and axle counters. Track circuits are based on a electrical circuit using the rails, to detect the presence of a train in a section. Axle counters detect train traffic by counting train axles crossing the border of a section using a counting head. The main advantages of axle counters over track circuits are that the length of the sections are virtually unlimited, that no insulating joints between sections are needed, and generally less installation and maintenance costs. Which system has been implemented is mostly a historical precedent. It is also possible to separate train safely using only train-based systems, as in ETCS L3. In this signalling system, the train itself send its position, length and integrity status to the interlocking system to release (digital) track sections [11]. This thus requires an extra train integrity monitoring (TIM) system.

(iv) *Interlocking* is a system composed by a set of signal apparatus that prevents trains from conflicting movements through only allowing trains to receive authority to proceed, when routes have been set, locked and detected in safe combinations. Its main function is to set and lock routes related to each train located in an area under its responsibility, in order to ensure safe movements along the track.

Various subsystems of railway signalling have been identified, but only two relevant signalling characteristics for resilience: the length and number of the block sections and in-cab signalling vs line-side signalling. Reduced block section lengths allow for smaller headways, especially in disrupted situations, when train are following one another closely at reduced speeds, this is of importance. In-cab signalling is important as this allows for more optimized braking curves, thus trains maintain higher speeds for a longer time and can follow one another better through speed restricted situations.

### F. Typical disruptions and dispatching

The purpose of a resilient system is to overcome disruptions. This section will therefore discuss which typical disruptions normally have to be overcome on a railway, and especially on the Dutch network.

A disruption can have a plethora of reasons. Eight sources of disruptions have been differentiated that take place at the Chinese high-speed rail network, a categorization which is likely to be also valid for other railways [12]:

- Bad weather: snow, rain, fog, etc.
- Vehicle on-board equipment failure: smoking alarm, failure of train control system, etc.
- Train body failure: bogie or wheelset failure, etc.
- Communication equipment failure: GSM-R failure, transponder failure, etc.
- Track system failure: switch or rail failure, etc.
- Electric related failure: pantograph breakdown, overhead wire failure, etc.
- Dispatching human interface failure or alarm
- Other: e.g. collisions

The severeness and impact of disruptions and the needed disruption management can be classified in two categories according to [13]: partial blockage and full blockage. In the first case, partial train traffic is possible, but balancing of the trains is needed. Examples of this are that from two tracks only one can be used, or that the the speed is restricted. In the full blockade case, no traffic at all is possible across a certain point in the line.

The following sources of disruptions have been identified as disruption where different signalling systems are likely to give a different outcome, and are thus interesting to be investigated: (i) collision, (ii) level crossing failure, (iii) signal failure, (iv) faulty train, and (v) switch failure.

The EU ON-TIME project states that resilience requires knowledge of the traffic control measures. The most common dispatching measures are therefore described here:

For the real-time rescheduling of the railway traffic during disruptions, by dispatchers as well as rescheduling tools, three main actions can be taken [14]: retiming, reordering and rerouting. In case of a full blockage, some extra actions or decisions are possible for the dispatcher, namely short-turning of services, cancelling services, extra stopping and stop skipping. This is done to isolate the impact of the disruption to adjacent areas.

## IV. Case study

The Dutch railway corridor Utrecht - Den Bosch was chosen for the simulation, since the number of trains on this corridor is sufficiently large, the headways and buffers between trains are small, physical traffic rescheduling measures are possible, and the train traffic is heterogeneously enough to provide interesting results.

The corridor is approximately 50 kilometers long, starting and ending at the main stations of Utrecht CS and Den Bosch, where the intercity trains stop, and with seven intermediate sprinter stations. Four of the sprinter stations (Vaartsche Rijn, Lunetten, Houten and Houten Castellum) are along the four track section, separated from the intercity trains. Culemborg and Zaltbommel are placed at the double track section, and Geldermalsen is a larger hub for sprinters, where overtaking is possible. Traffic on this route consists per direction of three

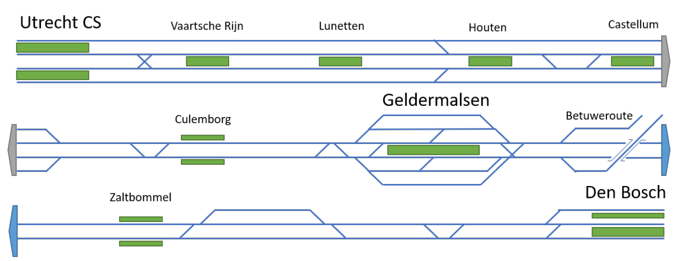



Fig. 3. Infrastructure layout as used in this research

half-hourly non-stop intercity services between Utrecht and Den Bosch, two half-hourly sprinter services between Utrecht and Geldermalsen, one of which continues to Den Bosch, and an hourly freight train from Utrecht to the branch to the Betuweroute.

The intercity services are all driven with VIRM train sets, the sprinters by SLT train sets, and the freight trains are commonly two BR189-locomotives pulling a coal train.

### A. Disrupted case scenarios

Based on the typical railway disruptions from the literature and historic data on disruptions on the corridor Utrecht - Den Bosch, seven disrupted case scenarios have been set up.

1) *ATP defect*: Due to damage to the ATP signal receiver, or loss of communication with ETCS, a service has been restricted to 40 km/h until a location where it can be taken out of service;
2) *Train engine defect*: Due to a mechanical failure, a service has stranded along the open tracks and needs to be towed away by another service or tow locomotive;

3) *Switch power failure*: the control of a switch can be lost due to a cable/power failure. It can be clammed in one direction so that traffic can pass in that direction anyhow
4) *Damaged switch rod*: when a switch rod has been damaged, no traffic is allowed over said switch until it has been repaired, and is thus redirected as much as possible via other tracks;
5) *Level crossing failure*: several reasons can lead to a failure of a level crossing, which then closes automatically. All services have to cross it at 15km/h to ensure no one is passing the level crossing;
6) *Incorrect track occupation*: A failing track circuit leads to an incorrect track occupation, meaning all drivers have to get explicit permission from the dispatcher to pass that section at reduced speed;
7) *Collision with a vehicle*: after a collision, the tracks are blocked until emergency services have cleared the tracks.

### *B. Signalling configurations*

In the literature review, two signalling characteristics have been distinguished which are most likely to be relevant for railway resilience, namely the block section lengths and whether the movement authority is provided via in-cab signalling or via line-side signals. To test the effect of both the characteristics separately as well as the combined effect, three signalling configurations will be modelled.

Firstly, the current NS'54/ATB-EG signalling configuration, as constructed outside, will be modelled. Data about the location of stations, switches, (relation between) signals, and speed limit markers is imported via the drawings made by Arcadis of the infrastructure layout.
Secondly, an ETCS Level 2 copy of the ATB configuration will be modelled. This configuration uses the existing blocks for track detection and the same placement of 'signals', which are now just marker signs. Lastly, an ETCS Level 2 with reduced block sections will be modelled. The blocks sections lengths are reduced to 200 meter, to approximate the effect of the moving block principle.

By comparing the second configuration to the first one, the potential effect of in-cab signalling over line-side signalling is tested, thus ETCS L2 over NS'54/ATB, whilst controlling for the block section lengths. Comparing the third configuration to the second configuration, gives the potential effect of short block section lengths over the current block section lengths in ETCS L2. The combined effect of both characteristics can be observed when comparing the third configuration to the first one.

To validate the train behaviour in the model, a comparison is shown in Figure 4 and 5 of a modelled train having to brake several times, either to stop or to enter a speed restricted situation. From here, it can be seen that the ETCS L2-train brakes later, in space as well as in time, and that is not effected by the 'yellow'-aspect, which can be seen for the ATB-train near km 18. Only one ETCS L2 configuration is added in Figure 5, since the blocking distance is not of influence to the the braking behaviour with ETCS L2 in such situations.


Fig. 4. Driving behaviour with NS'54/ATB

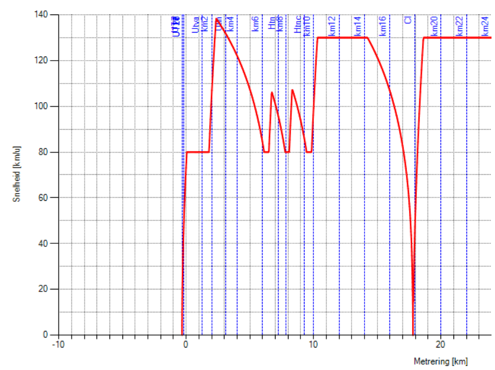



Fig. 5. Driving behaviour with ETCS L2

## V. RESULTS

Tables I and II show the outcome of the simulated scenarios on two performance indicators: (I) the time until the total delay in the system has decreased below five minutes, and (II) the total delay of trains at their final destination, compared to the base configuration NS'54/ATB.

TABLE I
OVERVIEW OF HOW QUICKLY THE DISRUPTION HAS RESOLVED IN EACH CONFIGURATION PER SCENARIO (H:MM)

| Scenario | NS'54/ATB | L2 existing | L2 reduced |
|---|---|---|---|
| 1 | 0:55 | 0:45 | 0:43 |
| 2 | 0:46 | 0:40 | 0:30 |
| 3 | 1:06 | 1:06 | 1:06 |
| 4 | 1:05 | 1:03 | 1:01 |
| 5 | 0:49 | 0:49 | 0:43 |
| 6 | 0:51 | 0:51 | 0:51 |
| 7 | 1:13 | 0:59 | 0:45 |

TABLE II
OVERVIEW OF THE DELAY PER CONFIGURATION PER SCENARIO COMPARED TO THE ATB CONFIGURATION

| Scenario | NS'54/ATB | L2 existing | L2 reduced |
|---|---|---|---|
| 1 | 100% | -9.1% | -12.9% |
| 2 | 100% | -13.6% | -25.1% |
| 3 | 100% | -19.2% | -19.3% |
| 4 | 100% | -9.0% | -12.6% |
| 5 | 100% | -2,2% | -27.9% |
| 6 | 100% | -3.0% | -11.5% |
| 7 | 100% | -11.8% | -28.6% |

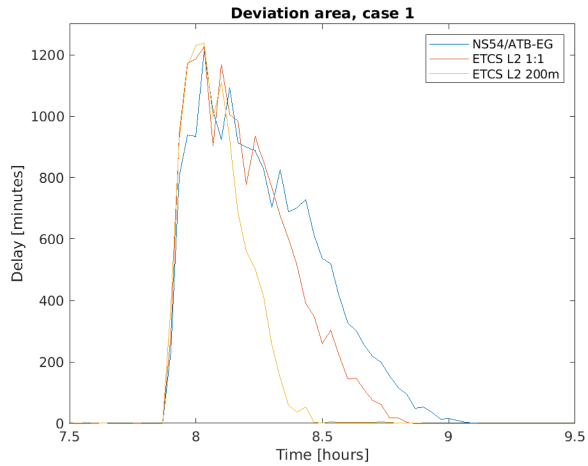



Fig. 6. Deviation area diagram of the $7^{th}$ scenario

Figure 6 shows the deviation area diagram of the $7^{th}$ scenario. Here, it is visible how the delay decreases quicker in the ETCS configurations than in the ATB configuration. The peak of the ETCS configuration are earlier in time and larger because the train services are restarted quicker, meaning that trains pass delay measuring points earlier in time.

The outcomes of the total delays of the scenarios are in a wide range, from -2,2% to -19% for the L2 existing blocks and from -12% to -28% for the reduced blocks. This is logical as the scenarios vary in length and severeness. In spite of the large differences in outcome, they do indicate that significant reductions of delay can be achieved by firstly upgrading to ETCS Level 2, and secondly by reducing the block sections for train detection. This is in accordance to the hypothesis in the theory that both of the signalling characteristics would increase the resilience.

The resilience index of the ETCS configurations can be calculated according to formula in the literature review. . For the ETCS L2 with existing blocks configuration, the resilience index compared to the base ATB-configuration is 1.10. For the L2 reduced blocks configuration, the resilience index is 1.25. This means that with the L2 existing configuration 9% of the delay has been saved in the modelled scenarios on average, and the L2 reduced one has saved 20% of the delay over all scenarios.

In the $1^{st}$, $3^{rd}$ and $4^{th}$ scenario, most of the benefits can be achieved by implementing in-cab signalling. Only a small extra benefit is obtained by reduced block sections length. In the other four scenarios, the reduced block section lengths were able to reduce the delay more than the in-cab signalling.

The simulation show that especially in short speed-restricted disruptions, such as switching over between tracks back and forth, the in-cab signalling part of ETCS L2 can make a difference. So, the more heterogeneous the train traffic, the more the traffic would have to brake in disrupted situations, and the more the traffic would profit from ETCS L2. Also, the short block sections show to be of great importance around switches and station areas, which are situations where traffic merges and splits or where the headways between trains become small.

In-cab signalling in combination with small blocks has lead to a significant increase of resilience in these scenarios. This strengthens the case of adjusting the current blocking lay-out to smaller blocks, when converting to ETCS, increasing the capacity as well as the resilience.

### A. Discussion

The results of this study are only applicable to corridors that have a similar intensity of the traffic. On corridors with less traffic and/or larger headways between trains the potential benefits will be smaller. For an even better comparison of the signalling characteristics, an extra configuration could have been simulated, namely the NS'54/ATB system with the smallest blocks as possible in that system. No differentiation has been made between the configurations to the used dispatching measures, to test the pure effect of the signalling characteristics and not that of the dispatching. For this, a larger network should be taken into account, as the dispatching measures are also of larger influence than just one corridor. Moreover, the results are dependent on the time and location of the disruption, the length of the trains, which trains has been chosen to be disrupted, the timetable on the corridor. Different input would have potentially led to a different outcome. The simulation tool Xandra assumes trains drive at full power and speed, a driving behaviour which is not applied by drivers in reality.

## VI. Conclusion

### A. Key findings

The following conclusions have resulted from the literature review, interviews and simulations:

- The resilience of the signalling system is just one of the many aspects to be considered in the infrastructural decision process of ERTMS. More important criteria are safety, the need for replacement, interoperability and cost;
- Increased resilience is most interesting for operators and passenger organizations; the power, on the other hand, lies in the hands of ProRail and engineering firms, as they are the ones responsible for the design of the lay-out of the signalling and infrastructure;
- There are three levels in the design phase of ETCS where resilience can be of support: full comparison between alternatives on all factors, including resilience; comparison between alternatives, but only on resilience; and to find the relation between design choices and resilience, based on expert judgement of the results of a resilience study;
- Two signalling characteristics are of relevance for resilience: the placement and length of block sections and whether the movement authority of a train is provided to the driver via line-side signals or via in-cab signalling;
- The modelled disruption scenarios has shown that compared to the current Dutch NS'54/ATB signalling configuration, an ETCS L2 configuration that uses the same block distances is 10% more resilient; An ETCS L2 configuration with reduced block section lengths of 200 meters is up to 25% more resilient than the current

NS'54/ATB signalling configuration. The effect of reducing the block sections lengths in ETCS L2 to an approximation of the moving-block principle is thus 13%;
- With very short block section lengths, the potential benefits is restricted by factors other than the signalling component, such as platform capacity or the speed difference between trains.

### B. Recommendations

Up to 10 or even 20% of passengers delay can be saved by installing ETCS L2, and by reducing the block sections lengths. This means that installing ETCS L2 not only increases the safety and capacity, but also helps to reduce the effect of disruptions, thus cutting down the costs for operators and increasing passenger satisfaction.

The benefits of increased resilience needs to be weighed against the costs of increasing the number of block sections. Designers can either use the results of this research or an appliance of this method to the concerned design area to underpin their design choices. The method used is also useful to assess the benefits of spatial dispatching possibilities, such as side-tracks and switches. The delay saved for passengers and freight, due to the increased resilience, then can be weighed off against costs and other criteria.

In the design phase of any future ETCS L2 lay-out, a trade-off should be included between the costs for extra infrastructure or signalling components, which increase the resilience and decrease the cost of disruptions. It can be decided to use the existing block sections or to use smaller block sections lengths, which costs extra but also increases the resilience. Furthermore, this simulation approach can be used to optimize the block sections lengths, and find which block sections are unnecessary regarding resilience. Operators can use these findings to calculate the estimated reduction of cost when having shorter disruptions, and strengthen their negotiation point to have a more resilient infrastructure. Although it is not required to take it into account since many other factors are of importance too, it can help the decision making or designing process.

### C. Further research

In this research, only simple traffic management has been applied, and even the same strategy in the ATB and ETCS L2 configurations to give a fair comparison. Research can or must be done on how traffic management will change when ETCS L2 will be implemented on more parts of the Dutch network. For this, a larger network should be simulated than this one corridor, as then more distinctive dispatching possibilities can be simulated.

It was beyond the scope of this research to find an optimal blocking lay-out strategy for the ETCS L2 configuration. However, this is interesting for further research to help the design of future ETCS L2 routes. It could be done with the simulation approach used in this research. In the ETCS L2 reduced blocks configuration, per case one could remove the blocks that have no influence on the outcome of the simulation of that case. Doing this for several cases may give an optimal blocking configuration strategy with respect to resilience.

This simulation approach could also be used to find which spatial dispatching possibilities, such as switches and side tracks, are relevant for the resilience of the train traffic and which are not and thus may be removed.

Further research is also recommended on the effect of combining ETCS with Automatic Train Operation. Research may be done to see what blocking distance strategy is most optimal for ATO over ETCS, what the expected capacity and robustness increase is with ATO, how the standards for headways between trains can change, how traffic management can or should change, and how the (specification and requirements for) the on-board and track-side systems are affected.

## REFERENCES


[1] Goverde, R. M., Corman, F., & D'Ariano, A. (2013). Railway line capacity consumption of different railway signalling systems under scheduled and disturbed conditions. *Journal of rail transport planning & management*, 3(3), 78-94.

[2] ON-TIME Consortium. Methods and algorithms for the development of robust and resilient timetables. *Deliverable D3*, 1.

[3] Corman, F., Quaglietta, E., & Goverde, R. M. (2018). Automated real-time railway traffic control: an experimental analysis of reliability, resilience and robustness. *Transportation planning and technology*, 41(4), 421-447.

[4] Quaglietta, E. (2014). A simulation-based approach for the optimal design of signalling block layout in railway networks. *Simulation Modelling Practice and Theory*, 46, 4-24.

[5] Dicembre, A., & Ricci, S. (2011). Railway traffic on high density urban corridors: capacity, signalling and timetable. *Journal of Rail Transport Planning & Management*, 1(2), 59-68.

[6] Goverde, R. M., & Hansen, I. A. (2013, August). Performance indicators for railway timetables. *In 2013 IEEE International Conference on Intelligent Rail Transportation Proceedings* (pp. 301-306). IEEE.

[7] Nicholson, G. L., Kirkwood, D., Roberts, C., & Schmid, F. (2015). Benchmarking and evaluation of railway operations performance. *Journal of Rail Transport Planning & Management*, 5(4), 274-293.

[8] Interview with dhr. Thomassen, H. *Director ERTMS program, Dutch ministry*, Utrecht, 2019

[9] ERA. Introduction to ETCS braking curves. *www.era.europa.eu, 2016*

[10] Hansen, I. A., & Pachl, J. (2014). Railway timetabling & operations. *Eurailpress*, Hamburg.

[11] ERTMS.be. Benefits of axle counters over track circuits. *www.ertms.be, 2017*

[12] Xu, P., Corman, F., & Peng, Q. (2016). Analyzing railway disruptions and their impact on delayed traffic in Chinese high-speed railway. *IFAC-PapersOnLine*, 49(3), 84-89.

[13] Blenkers, L., (2015). Railway disruption management. *TU Delft, thesis*.

[14] Hansen, I. A. (2010). Railway network timetabling and dynamic traffic management. *International Journal of Civil Engineering*, 8(1), 19-32.

## 7.2. Appendix B - Signalling systems

This appendix will give an overview of the principle and working of the Dutch NS'54/ATB signalling systems as well as ETCS levels 2 and 3. This overview is a combination of the explanations provided by Goverde et al. (2013) and Slootjes (2013).

### NS'54/ATB-EG

The signalling system NS'54 operates by light signals that give speed commands to drivers such that trains can always brake before a signal at danger. NS'54 uses fixed block sections with lengths that are at least equal to the maximum braking distance for a worst braking train plus a margin for the reaction time of the driver (and ATP system). The maximum braking distance depends on the entry speed into the block and is laid down in the Dutch railway law for a downhill gradient of 5%. A normal block has a length corresponding to the line speed in which case NS'54 is just a three-aspect two-block system with a clear (Green, G), approach (Yellow, Y) and stop (Red, R) aspect. A green signal indicates that the train may pass unhindered with the maximum track speed, a yellow signal orders to reduce speed to a restricted speed of 40 km/h and prepares to stop before a red signal, and a red signal orders to stop before the signal. As an example, for a line speed of 130 km/h the maximum braking distance is 1000 m and the corresponding block length with margin is 1181 m.

Near stations also short blocks are applied which have a length shorter than the maximum braking distance from the line speed by which trains can follow at a shorter headway. In this case, NS'54 indicates already one (or more) signals before that the train has to slow down to an indicated speed that must be reached before the next signal, so that the train will enter the short block with a lower speed associated to the short block length. This speed signalling is given by a Yellow signal plus a white numeral indicating the permitted speed at the next signal. For instance, 'Yellow 8' (Y8) indicates that the train has to reduce speed to 80 km/h before the next signal. This speed corresponds to a maximum braking distance of 800 m and block length of 918 m. For 40 km/h a maximum braking distance and block length of 400 m is defined. Consider for instance a block sequence from an open track to a station platform track with successive block lengths 1350 m, 1000 m and 400 m. Then a train is guided from the open track to standstill at the platform track by progressive speed signalling over a signal aspect sequence G-Y8-Y-R, assuming a line speed of 130 km/h. Local speed restrictions are indicated by flashing green with a white numeral indicating the permitted speed in the block.

The NS'54 wayside signalling system is complemented with the on-board automatic train protection system ATB to guard against errors of train drivers. ATB is a continuous ATP system that supervises overspeed and braking orders by the signalling system in rough steps (40,60,80,130,140 km/h). These speed steps are transmitted to the train via coded track circuits. After a speed reduction order (a yellow signal with possible speed indication) the driver has to apply the brakes until the permitted speed is reached. In the case of a yellow signal the driver has to reduce speed to a restricted speed of 40 km/h and then may drive on-sight for the remainder of the block. If a driver does not brake sufficiently after a speed reduction order then ATB warns the driver and if the driver still does not react ATB will intervene with an emergency brake to standstill. Speeds below

40 km/h are not supervised assuming that the driver runs on-sight and stops in time before a red signal. The supervised speeds correspond to the nearest ATB speed step above the permitted speed. For instance, if the permitted speed is 100 km/h then the ATB-code corresponds to 130 km/h which is shown to the driver as the supervised speed in the cabin. It is the responsibility of the driver to stay below 100 km/h. So the supervised speeds in the cabin are not equal to the actual permitted speeds. The driver can therefore not rely on the ATB cabin information but needs to pay attention to the wayside speed signs and signals. If the signal ahead improves to an aspect with an improved ATB-code (e.g. from yellow to green) then this is directly communicated to the train via the coded track circuits and the driver may increase speed again promptly.

## ETCS Level 2

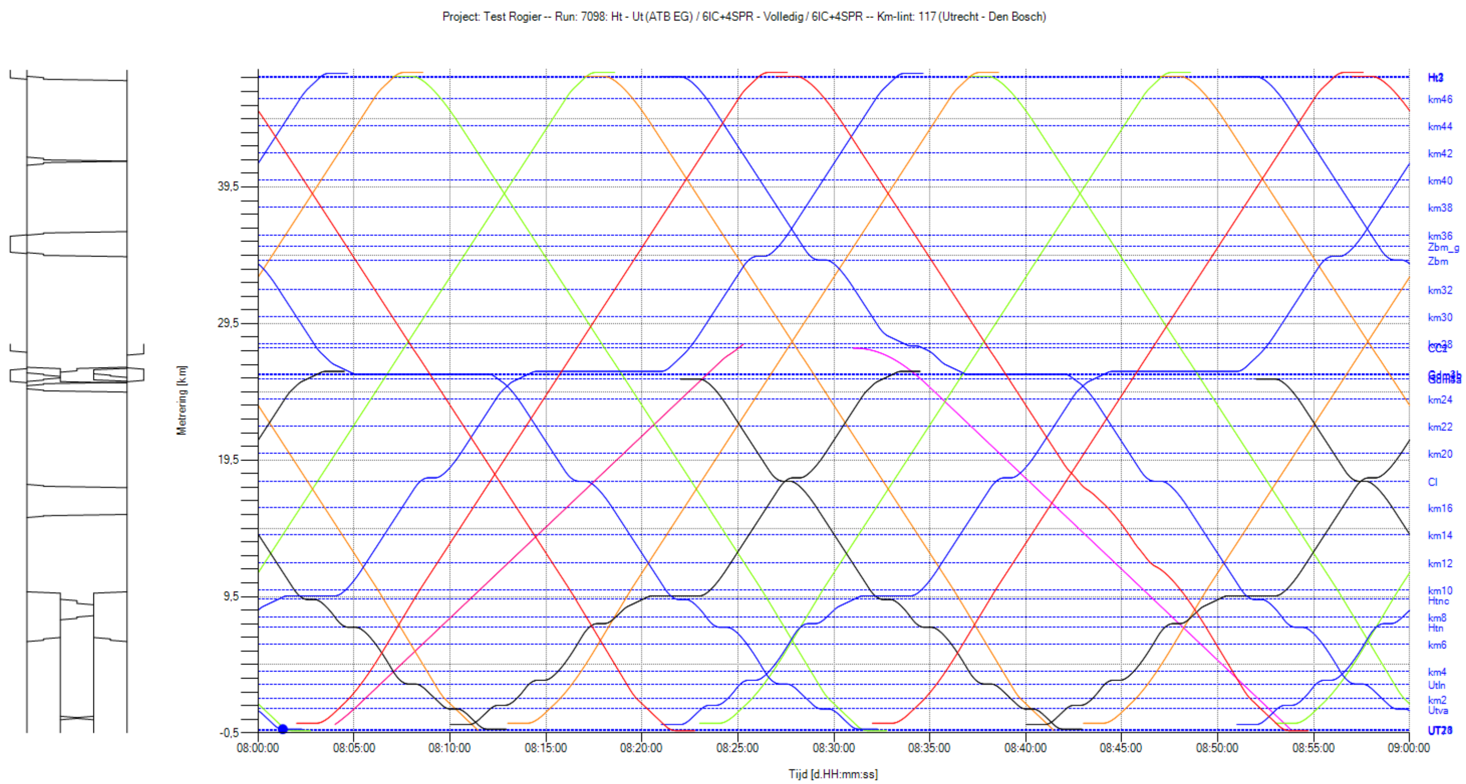
ETCS Level 2 is an integrated cab signalling and train protection system which will follow up the NS'54/ATB system in the future. Also ETCS L2 is a fixed block system with track-free detection that the interlocking system uses to set routes. However, now the interlocking communicates the set route(s) to a Radio Block Centre (RBC). The RBC translates the set route(s) into a Movement Authority (MA) and sends this together with a track description to the ETCS L2 on-board computer. The ETCS on-board computer computes a dynamic speed profile based on this incoming message, the current speed, and the train characteristics (maximum speed, braking curve), taking into account all track speed restrictions until the End of Authority (EoA). The MA and dynamic speed profile are displayed to the driver in the cabin, and are supervised continuously, i.e., the on-board computer supervises both the permitted speed at each location as well as the braking curves at each speed reduction. The driver must start braking only at the braking distance before an intermediate speed restriction or the EoA. This indication point depends on the track and train characteristics and the current speed. Note that an ETCS L2 train does not have to brake at the entry of a block but just before the computed braking curve. Trackside signals are no longer used. If a driver does not brake sufficiently and exceeds the (warning) braking curve a warning is given; if he then still does not brake enough an emergency brake intervention to standstill is carried out. Figure 6 illustrates the different braking and acceleration behaviour of ATB versus ETCS trains near a local speed restriction due to a crossover. In NS'54/ATB the local speed restriction is signalled to the driver by the three subsequent signal aspects Yellow 4 – Flashing Green 4 – Green. The ATB train has to brake from the approach signal to 40 km/h and maintain this restricted speed in the remainder of this block. The maximum speed of 40 km/h also must be maintained over the entire next block with the Flashing Green 4 aspect that includes the crossover section. Only when passing the next signal with the Green aspect the train can accelerate again to the line speed. In contrast, the ETCS train only has to brake just in rear of the crossover section according to the braking curve from line speed to restricted speed computed on-board. It then can reaccelerate again as soon as the back of the train has passed the crossover section. As a result of these local higher speeds of the ETCS train the running time of this train is shorter than the ATB train. Moreover, ETCS L2 has a two-way safe communication with the RBC via the GSM-R radio connection. The train sends (semi-)continuously its position and speed to the RBC by which the RBC can also request new routes to the interlocking which are then again translated in new MA's that are send back to the train. As soon as a new

MA has been received the current MA is extended with the new MA. So if for instance a train was braking before an EoA and the MA is extended then the driver can immediately accelerate again.

## ETCS Level 3

ETCS Level 3 works in may ways the same as Level 2, but is not fully developed yet. It is also an integrated cab signalling and train protection system, in which the RBC translates the set routes to a Movement Authority, after which the train itself computes its braking curve to the EnA. The difference is that in Level 3, track detection is done away with. Trains themselves have to communicate their integrity status with the RBC. Each train has to be equipped with a Train Integrity Monitoring System (TIMS) for this. The blocks are then no longer dependent on the track-side detection, and can be either virtual blocks to separate trains or moving blocks. Moving blocks allow trains to follow one another a braking distance, providing even more capacity than ETCS Level 2. Virtual blocks can be between 50 and 1000+ meter, dependent on the need for capacity. MA are provided then up to the beginning of the next occupied or restricted virtual block.

## 7.3. Appendix B - Basic Hour Pattern



## 7.4. Interviews

### 7.4.1. Interview with H. Thomassen - director ERTMS program - 27-08-2019 Utrecht

*Employed at the Ministry of Infrastructure and Water, Mr. Thomassen is the director of the ERTMS program in the Netherlands. Mr. Thomassen has had several management functions inside ProRail, Keyrail and NS, and therefore has a broad, nuanced knowledge of the rail sector as a whole.*

Conclusions of the interview:

- Reliability (or resilience) has not been a main criterion in the decision making process in 2015, as it is a criterion with a relatively small effect compared to the other criteria, such as costs, need for replacement of current infrastructure, and capacity.

- The ERTMS program is and needs to be an adaptive program, to cope with the impact of new technologies or new findings. An example of this is the development of Radio Block Centres, which in the beginning of the program could only manage 20 trains at a time, whereas current RBCs can already manage up to 400 trains, thus requiring only one or two for the whole Netherlands. However, adaptations to the initial program will always cost extra and this has to be weighed off against the benefits. The later in the program, the more the adaptations will cost.

- He states that there currently are too little connections between the university and the rail sector, although the necessity has been indicated and several programs have been set up. The length and complexity of the decision making process weakens the role of the university. A role he sees for the university is research firstly on how traffic control measures will or must change with the implementation of ERTMS; secondly what the capacity impact would be of ETCS L3; and thirdly research to back up or change the norms for headways when ERTMS is combined with Automatic Train Operation. Physical expansion of the railway network is a difficult and expansive task, so the norms need to be sharpened to use the the current capacity better.

- A stakeholder-based approach can be used to see which stakeholder has most interest in an increased robustness of the train traffic, and how much power this stakeholder has.

- In the transition period from ATB to ERTMS, the punctuality of operators is likely to decrease, due to technical or procedural failures. However, agreements have been made that operators are not to be blamed for this.