Optimising Railway Traffic Management under Radio-Based Distance-to-Go Signalling

Nina D. Versluis

Optimising Railway Traffic Management under Radio-Based Distance-to-Go Signalling

Nina D. Versluis

Cover image by Stefan Verkerk

Optimising Railway Traffic Management under Radio-Based Distance-to-Go Signalling

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus Prof. dr. ir. H. Bijl,
chair of the Board for Doctorates,
to be defended publicly on
Friday 30 January 2026 at 12:30 o'clock

by

Nina Dorien VERSLUIS
Master of Science in Applied Mathematics
Delft University of Technology
born in Amstelveen, the Netherlands

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus	Chairperson
Prof. dr. R. M. P. Goverde	Delft University of Technology, promotor
Dr. ing. E. Quaglietta	Delft University of Technology, copromotor

Independent members:

Prof. dr. D. C. Gijswijt	Delft University of Technology
Prof. dr. M. M. de Weerdt	Delft University of Technology
Dr. B. Atasoy	Delft University of Technology
Dr. ing. J. Beugin	Gustave Eiffel University, France
Dr. Y. Zhu	Delft University of Technology, reserve member

Other member:

Dr. P. Pellegrini	Gustave Eiffel University, France

The research leading to this dissertation has received funding from the Shift2Rail Joint Undertaking (JU) under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101015416 (PERFORMINGRAIL).

TRAIL Thesis Series no. T2026/2, the Netherlands Research School TRAIL
TRAIL
P.O. BOX 5017
2600 GA Delft
The Netherlands
E-mail: info@rsTRAIL.nl

ISBN: 978-90-5584-376-3

Printed in the Netherlands

Sometimes we stop at the green lights
If only we'd realize that it's time to go

- Hollow Coves

Acknowledgements

With the realisation that my defence date marks and concludes five years of this PhD journey, it is time to acknowledge those who have supported me throughout. I would like to start by thanking my supervisors - for favouring mathematical over railway knowledge, for continuing to believe in me throughout the journey, and for acknowledging that regular rescheduling is required not only in railway operations, but also in PhD trajectories.

Rob and Egidio, it was a pleasure to conduct my research in the Digital Rail Traffic Lab under your supervision. Rob, as my promotor, you were more approachable and present than I had expected. Our shared mathematical background, combined with your railway knowledge, made working together both instructive and motivating. Thank you for all the time and effort you have put in, both in guiding the content of my research and in supporting me throughout the process. Egidio, as my daily supervisor, you guided me into the project and provided support along the way, particularly moral support. Despite our personal differences, you were there for me when I really needed guidance.

Paola and Joaquin, it was a pleasure collaborating with you as my external supervisors. Thank you for having me join the COSYS-ESTAS team for ten months; I genuinely enjoyed my stay, with smooth automatic metro rides most of the time. Paola, thank you for the regular (online) meetings where I could ask questions freely, and for your support with the experiments (I think I still owe you a beer...). Joaquin, thank you for sharing your knowledge and experience whenever asked. Unfortunately, you could not be part of my PhD committee, but I am all the more grateful that you accepted the invitation to join my defence.

Besides my supervisors, I would like to thank the remaining members of my PhD committee. First, for accepting the invitation, and second, for accepting my dissertation. Dion, thank you for being the red thread through my academic journey, as part of my BSc, MSc and now PhD committee. Mathijs, thank you for also connecting with me in Delft, having first met during my internship at NS, and later at ODS conferences in Italy. Julie, thank you for the collaboration having worked together on the same project and hallway. Bilge, thank you for joining, even though we only met when I joined the interfaculty Freight and Logistics Lab for my post-doc. Yongqiu, thank you for returning to Delft and agreeing to join as a reserve member.

I would like to thank my colleagues from the Department of Transport and Planning, especially those in the old and new management team and the wellbeing taskforce, for having welcomed me into their midst and for the opportunities and support provided. A special thanks goes to my (former) Digital Rail Traffic Lab colleagues. I am especially grateful to my roommates at 4.17 for warmly welcoming me back after my visit and for creating a supportive environment.

I would also like to recognise my COSYS-ESTAS colleagues for their hospitality during my visit, as well as my current colleagues from the Freight and Logistics Lab and/or room 4.30.

I owe a special thanks to two colleagues in particular, my paranymphs Renate and Federico, for being my closest support in Delft and Lille, respectively. Having attended both your weddings, it is a true pleasure to have you as my academic maid of honour and best man.

Een warm dankjewel aan mijn vrienden en familie voor jullie interesse en begrip. Een speciale gedachte gaat uit naar mijn grootouders, die helaas niets van mijn academische weg hebben kunnen meemaken.

In het bijzonder wil ik degenen bedanken die het dichtst bij mij staan. Mam, ik waardeer je onvoorwaardelijke liefde en morele steun. Ook al lijkt het soms een ver-van-je-bed-show, ik weet dat je altijd het beste voor me wilt en trots op me bent - en daar ben ik je ontzettend dankbaar voor. Pap, ik wil je bedanken voor je voortdurende en ook praktische steun en aanmoediging door de jaren heen. Zonder jouw suggestie om ook eens bij de TU Delft te gaan kijken, was ik waarschijnlijk niet op dit punt gekomen. Judith, mijn grote zus wiens stappen ik steeds op mijn eigen manier volg. Zo ook promoveren, al hadden we dat allebei nooit van tevoren gedacht. Ik hoop dat je hebt genoten van je 14 maanden als enige dr. Versluis in de familie.

Colin, wat een geluk dat ik jou gedurende dit hele traject naast me heb gehad. Ik heb zo veel gehad aan je voortdurende liefde en steun. Samen zijn we ver gekomen, ondanks - of misschien juist wel dankzij - mijn tijd in Lille. Het afgelopen jaar heb ik nog meer van jouw liefdevol zorgende kant mogen zien, na het adopteren van Milo en Roos. Ook jullie, onze schattieflappies, bedankt voor de gezelligheid, de fijne afleiding en de wijze les dat zorg en aandacht voor elkaar het verschil maken.

Nina Versluis
Rijswijk, November 2025

Contents

Chapter 1

Introduction

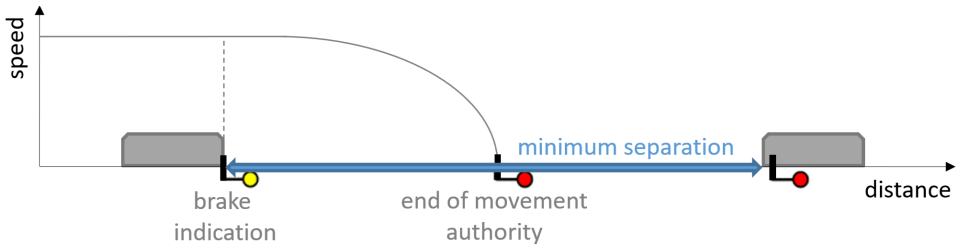
1.1 Context and background

Railways are widely acknowledged as a sustainable, cost-effective, safe and efficient transport mode, for both passenger and freight transportation (European Rail Supply Industry Association, 2022). While sustainability and cost-effectiveness are mostly relevant at strategic and tactical level, safety and efficiency are the most critical aspects in the context of real-time operations. Safe railway operations are primarily ensured by railway signalling systems, while railway traffic management is responsible for the efficient execution of scheduled railway operations. As this thesis contributes to the intersection of railway signalling and railway traffic management, these two topics are further introduced in this section.

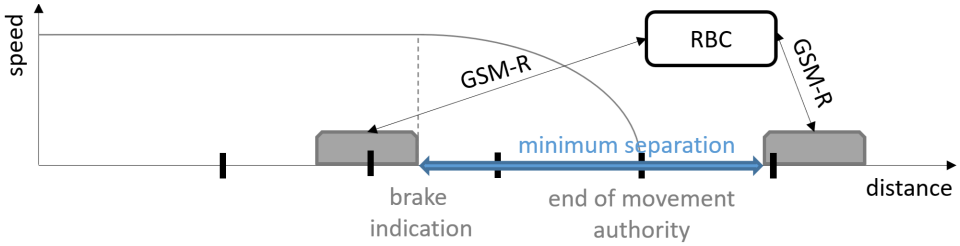
1.1.1 Railway signalling systems

Railway operations are characterised by long train braking distances and the reliance on safe train routes. The long braking distances, well beyond sight distance, are due to the low wheel-rail friction. While this low friction is advantageous for energy efficiency, it necessitates meticulous safety measures. Safe routes are guaranteed by infrastructure-based control mechanisms such as interlocking systems that manage routes over switches to divert trains to alternative tracks. Besides ensuring safe train routes, maintaining train separation and providing train protection are two key responsibilities regarding safe railway operations, which are all the responsibility of railway signalling. The safe separation of trains involves providing and supervising train movement authorities. That is, the permission granted to a train to move to a specific location. The operational rules regarding route protection and, more so, train separation highly depend on the implemented signalling system. Figure 1.1 illustrates train separation under various signalling systems, which, along with others, are introduced in the following paragraphs.

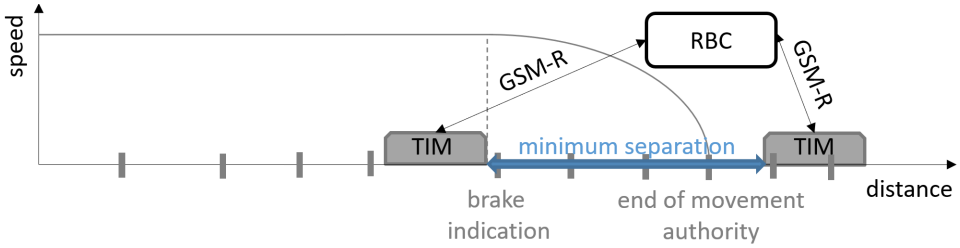
The conventional signalling systems are fixed-block multi-aspect systems with automatic train protection (ATP). As in any fixed-block system, the railway track is divided into blocks that are fixed both in space and length, and can only be occupied by one train at a time. The block entries are protected by trackside multi-aspect signals which indicate whether an approaching train can proceed, needs to start braking (either to a target speed or to a standstill), or is required to stop. The ATP system supervises that the trains follow the signal aspect and intervenes when



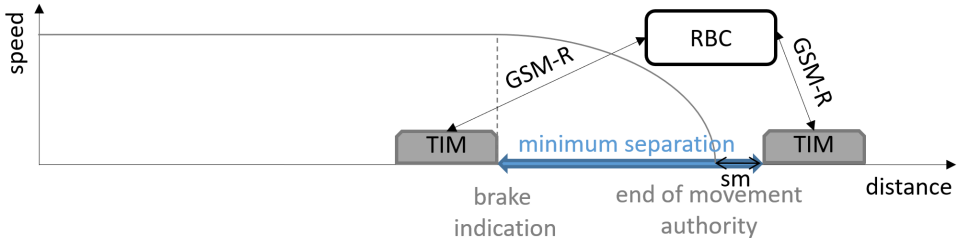
(a) Conventional fixed-block three-aspect signalling system.



(b) Fixed-block distance-to-go signalling system ETCS Level 2 with Trackside Train Detection.



(c) Virtual-block signalling system ETCS Level 2 Virtual Block.



(d) Moving-block signalling system ETCS Level 2 Moving Block.

Figure 1.1: Illustration of minimum train separation under different signalling systems with indicated features: Radio Block Centre (RBC), radio communication (GMS-R), on-board train integrity monitoring (TIM) and safety margin (sm).

a train does not start braking after a brake indication. Hence, minimum safe train separation under conventional fixed-block signalling is based on a number of blocks with lengths based on worst-case braking distances, i.e., ca. 1500 metres (for 140 km/h), see Figure 1.1a.

To reduce safe train separation and increase railway capacity, advanced signalling systems have been proposed. A first example is fixed-block systems featuring distance-to-go ATP with continuous braking curve supervision from the train front to the block exit corresponding to the end of the train's movement authority. With this, brake indications can be provided at any point along the track, and train separation is based on the train's absolute braking distance. That is, the distance the train needs to decelerate from its current speed to a standstill.

In Europe, railway signalling systems are developed in the context of the European Train Control System (ETCS) within the European Railway Traffic Management System (ERTMS)

specifications with interoperability throughout Europe as overall objective (European Rail Supply Industry Association, 2022). The ERTMS/ETCS framework supports multiple application levels and variants, which are defined by how information is communicated between trackside and train. ETCS Level 2 features radio-based cab signalling, with bidirectional radio communication between a Radio Block Centre (RBC) on the trackside and the trains facilitated by the Global System for Mobile Communication - Railways (GSM-R). The most advanced state-of-practice is ETCS Level 2 with Trackside Train Detection (TTD). ETCS Level 2 with TTD is a radio-based fixed-block distance-to-go system in which the fixed blocks correspond to TTD sections. These are the sections of the track in which the absence of a train is automatically detected, for example monitored by axle counters or track circuits, with typical lengths of 500 metres. As shown in Figure 1.1b, the minimum train separation is significantly shorter under ETCS Level 2 with TTD than under conventional fixed-block signalling, due to a combination of factors such as the train-specific braking distance and the brake indication point independent of section boundaries.

Despite the gain in capacity obtained from implementing fixed-block distance-to-go systems such as ETCS Level 2 with TTD, it remains a challenge to fulfil future railway demand. In Europe, the passenger demand is forecast to have increased by 30% in 2050 compared to 2000 (European Environment Agency, 2021). Moreover, the build and maintenance costs of trackside equipment are high. As a solution to the high costs and the limited capacity of existing railway networks, it is being considered to remove TTD and replace it with onboard train position and train integrity monitoring (TIM). With onboard TIM, movement authorities are no longer restricted to end at a block boundary. Signalling systems featuring both radio-based distance-to-go ATP and onboard TIM are referred to as next-generation signalling systems in this thesis.

Within ERTMS/ETCS, next-generation signalling is captured in ETCS Level 2 with Onboard TIM, previously known as ETCS Level 3 (European Rail Supply Industry Association, 2024). ETCS Level 2 with Onboard TIM includes two variants: Level 2 Virtual Block and Level 2 Moving Block. In Level 2 Virtual Block, the discretisation of the track into TTD sections is exchanged for a digital discretisation into shorter virtual blocks, with typical lengths between 25 to 500 metres. In Level 2 Moving Block, the track is represented continuously, with a total dismissal of track discretisation. This allows the train movement authority to end at the end of the last free virtual block (in case of Virtual Block) or at a safety margin in rear of the train ahead (in case of Moving Block), as shown in Figures 1.1c and 1.1d.

Table 1.1 provides an overview of the introduced signalling systems and their relevant characteristics in terms of track discretisation, distance-to-go ATP, and train integrity. Additionally, ETCS Hybrid Train Detection (HTD) is included. ETCS HTD, previously referred to as ETCS Hybrid Level 3, is currently developed by the railway industry as intermediate signalling solution between ETCS Level 2 with TTD and ETCS Level 2 with Onboard TIM (EEIG ERTMS Users Group, 2024), mixing TTD sections and virtual blocks to allow a mixture of trains equipped with and without TIM. In this thesis, the focus is on the three ETCS Level 2 variants: with TTD, Virtual Block and Moving Block.

As also illustrated in Figure 1.1, signalling features such as radio communication, distance-to-go ATP, finer track discretisation, and onboard TIM allow for shorter safe separation, thereby increasing the capacity on the existing railway network. For the efficient use of this increased capacity during operations, railway traffic management is relied upon, as explained in the following section.

Table 1.1: Overview of railway signalling systems and their characteristics in terms of track discretisation, radio-based distance-to-go ATP (DTG) and train integrity (TTD or onboard TIM).

Signalling system	**Track discretisation**	**DTG**	**Train integrity**
Conventional fixed-block	Fixed blocks	$\times$	TTD
Fixed-block distance-to-go	Fixed blocks	$\vee$	TTD
ETCS Level 2 with TTD	TTD sections	$\vee$	TTD
ETCS Level 2 Virtual Block	Virtual blocks	$\vee$	TIM
ETCS Level 2 Moving Block	$\times$	$\vee$	TIM
ETCS Hybrid Train Detection	TTD sections / Virtual blocks	$\vee$	TTD / TIM

1.1.2 Railway traffic management

Railway operations are scheduled as train services in a timetable. A timetable contains the routes and departure and arrival times of the scheduled services. In practice, scheduled services may deviate from the timetable due to, for example, variations in rolling stock, dwell times and driver behaviour. These deviations can lead to conflicting operations if they exceed the buffer times in the timetables. Maintaining conflict-free timetables during operations is the responsibility of railway traffic management. To ensure conflict-free operations, railway traffic management continuously monitors and predicts the traffic state containing train positions and speeds. The predicted traffic state is checked for any conflicting operations. Specifically considered are track conflicts, which occur whenever more than one train requires the same part of the track during overlapping time periods.

In case of track conflicts due to disturbances, traffic management is responsible for taking rescheduling measures to resolve the conflicts while minimising the impact on the railway network. These measures include retiming, reordering and local rerouting of trains. Retiming involves shortening or extending train running and/or dwell times. Reordering refers to changing the passing sequence of trains. Local rerouting corresponds to changing train routes through stations and/or platform assignment. Global rerouting, short-turning and cancelling are more involved measures that are typically only applied in case of disruptions, such as those caused by train or trackside equipment failure.

To support dispatchers in taking optimised rescheduling decisions, mathematical optimisation models have been developed. Of these conflict detection and resolution models, the vast majority proposed in the literature refers to conventional fixed-block signalling, with modelling approaches relying on the system's inherent infrastructure discretisation and neglecting the speed-dependency of train separation. With the literature on conflict detection and resolution under next-generation signalling limited to some preliminary works (Büker et al. (2019), Janssens (2022), Meunier et al. (2021) and Pochet et al. (2016, 2017)), the impact of next-generation signalling on railway traffic management is unclear. With that, it is also unknown how railway traffic management can be supported in maintaining safe and efficient railway operations under next-generation signalling.

1.2 Research aim and questions

The aim of the research presented in this thesis is to develop models to support the effective management of real-time railway operations under radio-based distance-to-go signalling. The focus is on the proactive prevention of track conflicts for such railway signalling systems. To achieve this aim, the following five research questions are addressed:

1. What are the current research gaps and challenges in modelling conflict detection and resolution under moving-block as next-generation radio-based distance-to-go railway signalling system?
2. How can conflict detection and resolution under radio-based distance-to-go railway signalling be modelled?
3. What are the effects of radio-based fixed-block distance-to-go signalling over conventional fixed-block signalling in conflict detection and resolution?
4. What is the impact of track discretisation granularity on conflict detection and resolution under virtual-block signalling?
5. How can railway traffic management be supported in mitigating hazardous traffic conditions under radio-based distance-to-go signalling with onboard train integrity monitoring?

1.3 Thesis contributions

In addressing the research questions, this thesis provides contributions for both the scientific community and society, as set out in Sections 1.3.1 and 1.3.2.

1.3.1 Scientific contributions

The research presented in this thesis contributes to the scientific field of railway operations, specifically by adding to the knowledge regarding the interaction of railway traffic management and railway signalling. The five main contributions to this research field are as follows:

- **Research gaps and challenges.** Current research gaps and challenges in the modelling of conflict detection and resolution under moving-block signalling are identified based on a literature review. With most research focusing on conventional fixed-block signalling, radio-based distance-to-go signalling remains underrepresented. Consequently, the impact of radio-based distance-to-go signalling on conflict detection and resolution models is not yet understood.

- **Crucial modelling steps.** Necessary steps for modelling conflict detection and resolution under radio-based distance-to-go signalling are established to address the identified research gaps. Implementation of these steps can enable the detection and resolution of conflicts under radio-based distance-to-go signalling by explicitly accounting for speed-dependent train separation.

- **Conflict detection and resolution model for distance-to-go signalling.** A state-of-the-art conflict detection and resolution model for conventional fixed-block signalling is

enhanced to incorporate speed-dependent train separation, representing distance-to-go signalling. Application of the obtained model demonstrates how radio-based distance-to-go signalling affects the performance of railway traffic management.

- **Track discretisation impact.** The impact of track discretisation on conflict detection and resolution for next-generation signalling is assessed. Sensitivity analyses of the developed model show how track section lengths affect minimum safe train separation and potential rescheduling decisions under virtual-block signalling. These analyses extend existing knowledge of the influence of the virtual block section configuration on operational performance.

- **Early-warning hazard prediction framework.** An early-warning framework for hazardous railway traffic conditions under next-generation signalling is proposed. Extending the general railway traffic management framework, it provides a model-based approach to proactively predict and mitigate performance hazards. Incorporating the developed conflict detection and resolution model for radio-based distance-to-go signalling, it specifically addresses hazards that may arise from the reliance on satellite-based train integrity monitoring. Application of the framework shows how these hazards affect rescheduling, and conversely how rescheduling can support operations under hazards, adding to the knowledge on the interplay between hazards and operational performance in next-generation signalling systems.

1.3.2 Societal contributions

The relevance of the research in this thesis to society can be considered through several perspectives, including passenger, dispatcher, railway industry and society in general. Given these perspectives, the main societal contributions are as follows:

- **Sustainable transport.** Mitigating climate change requires a modal shift toward lower-emission transport options such as railways. To meet future demand and keep railways an attractive mode, radio-based distance-to-go signalling systems are being implemented. The research in this thesis supports their implementation by providing insights into the effective management of railway operations under these signalling systems.

- **Safe and efficient railway operations.** To increase railway capacity without expanding networks or compromising safety, radio-based distance-to-go signalling systems can be implemented. To ensure efficient operations under these systems, railway traffic management must consider potential effects of distance-to-go characteristics on operational decisions. This thesis provides insights into how radio-based distance-to-go signalling influences railway traffic management.

- **Decision support for human dispatchers.** Decision support systems can aid human dispatchers in taking optimised decisions in case of disturbed operations. These support systems should be fitted to the signalling system in place. This thesis contributes to the development of decision support systems for radio-based distance-to-go signalling by developing traffic management models that can be incorporated in such systems.

- **Signalling system design considering traffic management.** A crucial aspect of implementing virtual-block signalling is the granularity of the track discretisation. Understanding how section lengths influence railway operations is essential, also during disturbances that require rescheduling. This thesis provides insights into the effects of section lengths on rescheduling decisions, supporting the design of radio-based distance-to-go signalling systems.

1.4 Thesis structure

This thesis consists of six chapters, structured as visually presented in Figure 1.2. As main chapters, Chapters 2 to 5 are self-contained and collectively address the research questions posed in Section 1.2. Chapters 3 and 4 together form the modelling part of the thesis, grounded in Chapter 2 and integrated into Chapter 5.

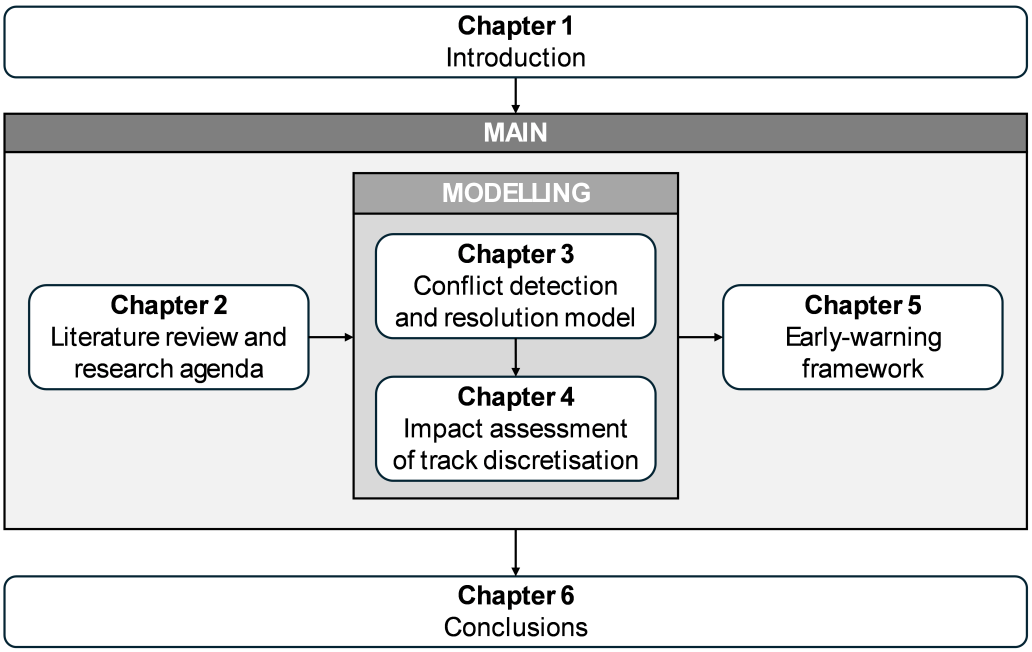


Figure 1.2: Thesis structure.

In Chapter 2, the literature on conflict detection and resolution models is reviewed with the aim to identify research gaps and challenges in the modelling of conflict detection and resolution under moving-block signalling. In line with the identified gaps, a research agenda with future steps in the development of conflict detection and resolution models for moving-block signalling is proposed, which guides the developments in the subsequent modelling chapters.

Chapter 3 addresses the research gap regarding conflict detection and resolution models for signalling systems with distance-to-go ATP. Speed profile options and train separation based on speed-dependent braking distances are proposed and embedded into a state-of-the-art conflict detection and resolution model to obtain a model describing fixed-block distance-to-go railway operations. Specifically for ETCS Level 2 with TTD, the obtained model is compared with the original one to investigate the effects of fixed-block distance-to-go over conventional fixed-block signalling in conflict detection and resolution.

Subsequently, Chapter 4 further extends the fixed-block distance-to-go model developed in Chapter 3 to describe ETCS Level 2 with Onboard TIM. For this model, the impact of discretising the railway track into virtual blocks of varying lengths is assessed.

Chapter 5 proposes an early-warning framework to support dispatchers in mitigating hazardous traffic conditions under next-generation signalling. The core railway traffic management framework is enhanced with an early-warning module, in which the state of onboard TIM parameters is monitored and predicted. In cases of increased parameter values, the potentially hazardous effects are estimated and presented to the dispatcher. In the latter part, the conflict detection and resolution model obtained in Chapters 3 and 4 is considered.

Finally, Chapter 6 concludes the thesis, presenting the main findings by answering the research questions. Additionally, recommendations for practice and future research are provided towards the effective traffic management of next-generation railway operations.

Chapter 2

Conflict detection and resolution under moving-block signalling: A literature review and research agenda

---

The concepts of conflict detection and resolution and railway signalling are introduced in Chapter 1. Specifically, mathematical conflict detection and resolution models are presented as support tool optimising the rescheduling of disturbed railway operations. The modelling of conflict detection and resolution under radio-based distance-to-go signalling systems such as moving block is, however, underrepresented in literature.

This chapter presents future steps in the modelling of conflict detection and resolution under moving-block as next-generation radio-based distance-to-go railway signalling system. First, the available literature on conflict detection and resolution models is reviewed with the aim to identify current gaps and challenges in modelling conflict detection and resolution for moving-block railway operations. Subsequently, research steps to address the identified gaps and challenges are proposed based on a comparative analysis of applicable modelling approaches.

Apart from minor changes, this chapter has been published as:

Versluis, N. D., Quaglietta, E., Pellegrini, P., Goverde, R. M. P., & Rodriguez, J. (2024). Real-time railway traffic management under moving-block signalling: A literature review and research agenda. *Transportation Research Part C: Emerging Technologies*, 158, 104438.

2.1 Introduction

Moving-block signalling is an innovative railway signalling solution, offering increased efficiency and capacity compared to fixed-block signalling (European Rail Supply Industry Association, 2022). In conventional fixed-block signalling systems, the track is divided into fixed-length block sections, which are protected by trackside signals. These systems rely on trackside train detection (TTD) for train position and integrity monitoring, with train integrity referring to whether a train has not accidentally split. In these fixed-block signalling systems, train headways, i.e., the minimum safe head-to-head separation distance between two trains, are based on a preset number of block sections considering worst-case braking distances and number of signal aspects. In moving-block signalling systems, the fixed block sections are eliminated. Trackside signalling and train detection are replaced by radio-based cab signalling and onboard train positioning and train integrity monitoring (TIM), respectively. With this, moving-block train headways are continuously based on absolute braking distances to the tail of the preceding train, i.e., the distance a train needs to decelerate to a standstill from its current speed. The continuous braking curve calculation is characteristic to distance-to-go signalling, which is not only featured in moving-block systems but also in advanced fixed-block systems. In fixed-block distance-to-go signalling systems, the fixed-block track division is combined with radio-based cab signalling to enable train headways derived from absolute braking distances to the end of a block.

In moving-block systems, the radio communication system, the train positioning technology and the onboard TIM device are safety-critical components (Martinez & Martin, 2020). So far, TIM is only available for homogeneous, fixed-composition trains and closed networks with low complexity. As a result, moving-block technology is typically deployed on urban railways which have these characteristics (Martinez & Martin, 2020). There, moving-block technology is implemented within Communications-Based Train Control (CBTC) systems. For mainline railways with heterogeneous traffic, variable train compositions and more complex networks, fixed-block signalling is still the standard. In this area, current research focuses on the further development of the safety-critical aspects, rather than on the operational efficiency of the moving-block system (e.g., Himrane et al. (2023) and Lazarescu & Poolad (2021)).

Crucial for the operational efficiency is real-time railway traffic management. It is responsible for the detection and resolution of conflicting operations arising from disturbances, i.e., relatively small delays originating from, for example, variations in rolling stock, dwell times and driver behaviour. In current practice, human dispatchers take rescheduling measures mostly based on experience and preset rules. In the literature, conflict detection and resolution (CDR) models exist that can support dispatchers in minimising the impact of disturbances on the network. CDR models generally consider measures such as retiming, i.e., shortening or extending running and/or dwelling times, reordering, i.e., changing passing sequences, and local rerouting, i.e., changing station route and/or platform assignment (Cacchiani et al., 2014). In urban railways, typically only retiming is considered (Pochet et al., 2017).

The literature on the modelling of CDR under moving-block signalling is limited to the preliminary works of Büker et al. (2019), Janssens (2022), Meunier et al. (2021) and Pochet et al. (2016, 2017). The vast majority of CDR models proposed in the literature refers to fixed-block signalling, with modelling approaches relying on the system's inherent infrastructure dis-

cretisation and a limited dependence of train headways on speed (e.g., D'Ariano et al. (2007a), Reynolds et al. (2020) and Törnquist & Persson (2007)). A knowledge gap exists regarding the modelling of CDR for moving-block operations in terms of infrastructure representation and speed-headway relation, as well as with respect to the impact of moving-block CDR models on the management of heterogeneous traffic under disturbed conditions.

In this chapter, the available literature on CDR models is reviewed with the aim to identify existing gaps and to propose future steps in the research on CDR under moving-block signalling. Specifically, a comparison of modelling approaches considering infrastructure and speed modelling is included. The main contributions of the chapter are:

- The identification of gaps and challenges in the research on CDR under moving-block signalling based on a comprehensive review of the existing literature.
- A comparative analysis of modelling approaches for the application to moving-block CDR.
- A research agenda proposing future steps in the development of CDR models for moving-block signalling.

This chapter is structured as follows. Section 2.2 provides background on railway signalling systems. Section 2.3 presents a literature review on CDR models. In Section 2.4, gaps in the research on CDR under moving-block signalling are identified and research challenges are formulated. Section 2.5 presents a research agenda on CDR under moving-block signalling. The chapter finalises with concluding remarks in Section 2.6.

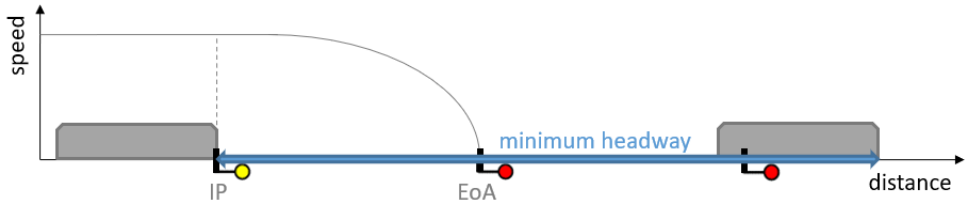
2.2 Railway signalling systems and blocking time theory

In Section 2.1, the fixed-block and the moving-block signalling systems have been shortly introduced. Here, the systems are described in more detail and compared to one another in terms of minimum headways. Additionally, blocking time theory is introduced as one of the most frequently used models for minimum headways.

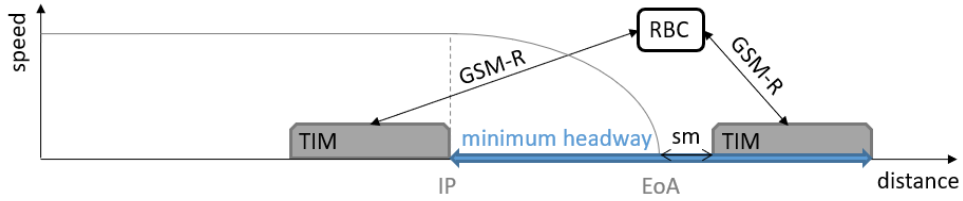
Conventional fixed-block signalling systems are characterised by the division of the track into fixed-length block sections, of which the entry points are protected by trackside signals. These signals use multiple aspects to communicate train movement authorities (MAs), i.e., the permission for a train to move to a specific location, and corresponding speed commands. The colour aspects indicate different instructions to an approaching train. For instance, in the case of three-aspect signalling, they indicate whether the train needs to stop (red), needs to start braking and prepare to stop (yellow) or can proceed without restrictions into the following block section (green). A train needs to stop if the block section ahead is assigned to another train or otherwise not available. It needs to start braking if it must slow down to be able to halt before the next stop indication. If the system has more than three aspects, additional restrictive aspects exist before a red signal, and hence, increase the number of block sections available to come to a halt.

Figure 2.1a illustrates the fixed-block multi-aspect system, providing a schematic representation of its main features. These include the brake indication point (IP) at the block entry corresponding to a trackside signal showing yellow, and the end of authority (EoA) at the end of the block right before the first red signal. For the monitoring of train position and integrity, this signalling system relies on trackside train detection devices, such as track circuits or axle

counters. To protect against driver errors, multi-aspect block systems are complemented by an automatic train protection (ATP) system, which supervises adherence to the signal aspects. The ATP system can intervene when a train fails to respect restrictive indications.



(a) Fixed-block multi-aspect signalling system with block sections and trackside signals.



(b) Moving-block signalling system with onboard train integrity monitoring (TIM), Radio Block Centre (RBC), radio communication (GSM-R) and safety margin (sm).

Figure 2.1: Schematic layout of the minimum headway between two trains in a speed-distance diagram under different signalling systems. The headway is related to the brake indication point and the end of authority depending on the braking curve.

Moving-block signalling systems feature radio-based cab signalling with a distance-to-go ATP system and onboard train positioning and TIM, while dismissing the concept of fixed block sections (European Union Agency for Railways, 2016). The elimination of the fixed block sections, in combination with cab signalling, onboard train positioning and TIM, allows trains to maintain headways based on absolute braking distances. Cab signalling is enabled by bidirectional radio communication, facilitated by, e.g., the Global System for Mobile Communication - Railways (GSM-R), between a Radio Block Centre (RBC) on the trackside and the trains. This bidirectional communication allows a train's MA to be continuously provided up to the tail of the preceding train, or up to a movable track element such as a switch, respecting a safety margin. Distance-to-go refers to the onboard computation and supervision of dynamic speed profiles including continuous braking curves to an EoA, enabling the IP to lie anywhere along the track.

Figure 2.1b illustrates the moving-block signalling system on an open line, providing a schematic representation of its main features. These include the EoA for the following train at a safety margin behind the tail of the leading train, and the corresponding IP upstream along the track, depending on the train's braking distance.

Moving-block system specifications are defined for urban (Institute of Electrical and Electronics Engineers, 2025) and mainline railways (European Union Agency for Railways, 2016). For the urban railways, the specifications are in the context of CBTC. For the mainline railways, the moving-block concept is developed within the framework of the European Rail Traffic Management System (ERTMS), consisting of the European Train Control System (ETCS) and GSM-R (European Rail Supply Industry Association, 2022). In this context, moving-block signalling is incorporated within ERTMS/ETCS Level 2 Moving Block.

Fixed-block distance-to-go signalling is a signalling solution defined and developed within ERTMS/ETCS as Level 2 with TTD. In fixed-block distance-to-go signalling systems, the track division into fixed blocks is combined with bidirectional trackside-to-train radio communication. This enables distance-to-go signalling. Hence, the IP can lie anywhere along the track, while the EoA is consistently set at the end of a block section (European Union Agency for Railways, 2016). Another distance-to-go solution is ERTMS/ETCS Level 2 Virtual Block. Under virtual-block signalling, the infrastructure is virtually divided into short blocks with fixed lengths. The system relies on onboard train positioning and TIM for train-tail detection to clear a virtual block, so that the EoA lies at the end of the last released virtual block (European Union Agency for Railways, 2016; Furness et al., 2017).

The signalling system deployed determines the minimum headway between two trains. Indeed, the minimum headway depends on the EoA and the IP of the following train. In fixed-block (multi-aspect and distance-to-go) systems, the EoA of the following train is located at the end of the block section last released by the leading train. In moving-block systems, the EoA of the following train is consistently maintained at a safety margin behind the tail of the leading train, irrespective of the leading train's position. In fixed-block multi-aspect signalling systems, the IP corresponds to the first block entry upstream ensuring at least the braking distance before reaching the EoA. In signalling systems with distance-to-go ATP, i.e., fixed-block distance-to-go and moving-block systems, the IP is situated precisely the train's safe braking distance further upstream of its EoA. Figure 2.1 illustrates the minimum headway in terms of EoA and IP under fixed-block multi-aspect (Figure 2.1a) and moving-block (Figure 2.1b) signalling.

In the literature, blocking time theory is used to describe minimum train headways. It it is a well-known concept for the detection of track conflicts (Hansen & Pachl, 2014) and it will be used in the following to compare various modelling approaches for CDR. Blocking time is the time a track part, e.g., a block section, is assigned to a train and hence blocked for other trains; it starts when a train requests the track part for its route and ends after the train has traversed and cleared it. The total blocking time of a track part constitutes of the following components:

- Setup time - to request, set and lock the route over the track part;
- Reaction time - to perceive and respond to speed/braking indication;
- Approach time - to run from the IP to the track part;
- Running time - to run over the track part, including possible dwell time;
- Clearing time - to run until the train has left the track part with its full length; and
- Release time - to release the route.

The blocking time and (most of) its components depend on the signalling system (Büker et al., 2019). The difference in setup time between fixed-block multi-aspect and distance-to-go systems relates to the shift from trackside to cab signalling. In the former, the MA is communicated via the trackside signals, while in the latter, the MA is communicated by the RBC through radio communication. Similarly, the reaction time is affected by the shift from trackside to cab signalling. The time to perceive a signal depends on whether the signal is continuously visible (cab signalling) or only from a certain sight distance (trackside signalling). The time to respond to a signal is similar in all systems as the MA information, however received, needs to be translated into an action. The approach time is significantly different in the respective systems due to its dependency on the IP, which is determined differently as explained before. The running time component is the other main change due to the track parts considered: fixed block sections, virtual block sections, or no sections at all in moving block, where the running time component

is negligible. The clearing time can be considered the same for the different systems; the time it takes to run over a train length is independent of the signalling system. The release time is influenced by the system's reliance on either onboard TIM or TTD for the track-clear detection because of the different communication systems. Also, whereas in fixed-block multi-aspect systems the signals need to be released, i.e., set to the default value, this subcomponent can be omitted in the systems with distance-to-go signalling. Under moving-block signalling, the blocking times around movable track elements are defined in a similar manner as under fixed-block distance-to-go signalling. For more details on railway signalling, see Theeg & Vlasenko (2020).

2.3 Literature review of conflict detection and resolution models

In this section, the existing literature on CDR models is reviewed as a springboard for the identification of gaps in the literature regarding CDR under moving-block signalling. Figure 2.2 illustrates that CDR under moving-block signalling is staying behind on the trend of an increasing number of publications on moving-block signalling. A similar trend holds for the number of publications on CDR in general.

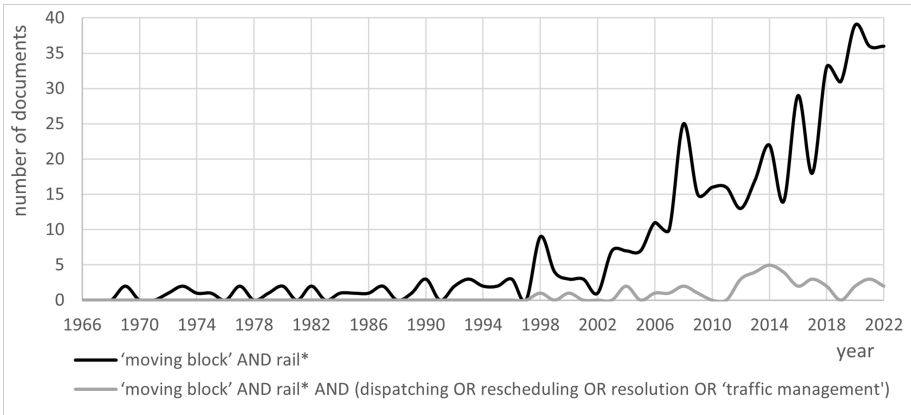


Figure 2.2: Available documents in Scopus (2023) by search queries 'moving block' AND rail and 'moving block' AND rail* AND (dispatching OR rescheduling OR resolution OR 'traffic management').*

To provide a comprehensive analysis, the review extends its focus beyond moving-block signalling, also including the other types of signalling systems described in Section 2.2. First, the review methodology is described in Section 2.3.1. Then, the review of CDR models under fixed-block multi-aspect, fixed-block distance-to-go and moving-block signalling are presented in Sections 2.3.2 and 2.3.3 and 2.3.4, respectively. Finally, a summary is given in Section 2.3.5.

2.3.1 Literature search, selection and classification

Two prominent review papers on CDR for conventional fixed-block operations, i.e., Cacchiani et al. (2014) and Corman & Meng (2015), serve as guideline for the literature to consider until

2015. For the period after that, the literature selection is conducted using forward snowballing, i.e., by identifying follow-up papers or other papers referencing the ones reviewed by Cacchiani et al. (2014) and Corman & Meng (2015), cross-referenced with the search query "rail* AND (dispatching OR rescheduling OR resolution OR 'traffic AND management')" in Scopus (2023). Also, the authors' expertise is used to identify the relevant papers within and beyond the selection procedure.

Foremost, the papers are separated based on the modelled signalling system: fixed-block multi-aspect, fixed-block distance-to-go or moving-block. Within this classification, several (CDR) modelling aspects are considered that relate to the critical importance of train headways in both the modelling of (possibly conflicting) railway operations and moving-block signalling: modelling approach, infrastructure modelling, speed modelling, headway modelling, consideration of rerouting, solution method and objective function.

For CDR, many different modelling approaches have been proposed and applied by many different researchers (Cacchiani et al., 2014).

In the modelling of railway infrastructure, there are two main levels of representation: macroscopic and microscopic. Macroscopic models depict the infrastructure at a lower detail level, often merging parallel tracks into one line and omitting block sections. Microscopic models provide a more detailed representation, including block sections and parallel tracks in junctions and station areas. In the context of CDR, the infrastructure is typically modelled at the microscopic level (Cacchiani et al., 2014). This allows the use of blocking time theory in the modelling of minimum train headways.

While speed modelling is not strictly considered as part of CDR, it plays a crucial role in translating (updated) schedules to train operations (Quaglietta et al., 2016). CDR models can be classified into two types: fixed-speed or variable-speed models. In fixed-speed models, speed variation dynamics resulting from traffic conditions are neglected, and unplanned stops are assumed to lead to immediate stand-stills, without accounting for acceleration and deceleration. In contrast, variable-speed models do incorporate speed dynamics, also in case of unplanned braking.

Headway modelling is heavily influenced by the other aspects. For instance, whether the modelling approach and infrastructure representation level allow for the application of blocking time theory (Cacchiani et al., 2014; Corman & Meng, 2015).

In the context of disturbed railway operations, relevant rescheduling measures are retiming, reordering and local rerouting of trains (Cacchiani et al., 2014). By definition, retiming and reordering measures are considered in CDR models (for mainline railways). The inclusion of rerouting is often explicitly mentioned, e.g., as 'train routing and scheduling' (Pellegrini et al., 2014).

The real-time aspect of CDR introduces the need to find a balance between solution quality and computation time. Both commercial solvers and customised heuristics are used for solving CDR problems (Corman & Meng, 2015).

The primary objective of CDR is to minimise the impact of disturbances on the network. Various objective functions can be considered. Typically, they refer to the minimisation of delays (Cacchiani et al., 2014; Corman & Meng, 2015).

2.3.2 Conflict detection and resolution under fixed-block multi-aspect signalling

Three main classes of CDR models for fixed-block multi-aspect signalling are: alternative graph (AG) models (e.g., D'Ariano et al. (2007a) and Mazzarello & Ottaviani (2007)); disjunctive mixed integer linear programming (MILP) models (e.g., Luan et al. (2018), Pellegrini et al. (2015) and Törnquist & Persson (2007)); and time-indexed MILP models (e.g., Bettinelli et al. (2017), Lusby et al. (2013) and Reynolds et al. (2020)).

Other modelling approaches explored in the literature are model predictive control (Caimi et al., 2012), Monte Carlo tree search (Lövétei et al., 2021), stochastic programming (Meng & Zhou, 2011) and constraint programming (Rodriguez, 2007; Marlière et al., 2023). Despite some promising results of these approaches, they have not been further picked up in the research field yet.

Alternative graph models

In AG-based models, the (re)scheduling of railway operations is considered as a no-wait job shop scheduling problem. For this well-known scheduling problem, an AG formulation is developed by Mascis & Pacciarelli (2002). The approach was first applied to railway operations by D'Ariano et al. (2007a) and Mazzarello & Ottaviani (2007). The railway network is represented as an alternative graph consisting of nodes, fixed arcs and alternative arcs, as illustrated in Figure 2.3. The nodes correspond to the entry times of trains into block sections, while the weight of the fixed arcs represent the train running times within the block sections and the dwell times at station platform tracks along a route. In D'Ariano et al. (2007a), the running time reflects the time it takes to traverse the block section at the scheduled speed, whereas Mazzarello & Ottaviani (2007) consider the minimum running time based on the maximum line speed. In both cases, the running time is considered fixed for each block section.

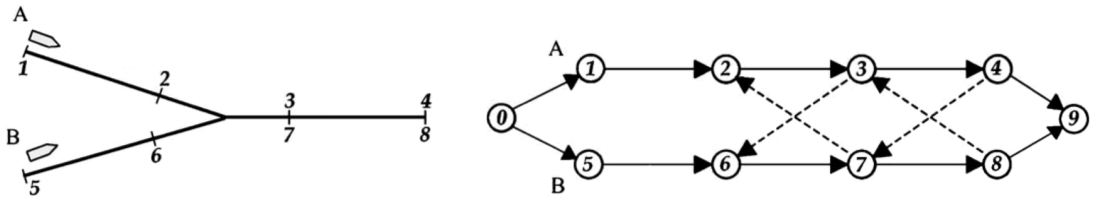


Figure 2.3: A simple network (left) and the associated alternative graph (right). The network features trains A and B, and their block section entries 1 to 8. The alternative graph features a source and sink node, 0 and 9 respectively, nodes corresponding to the block entries of trains A and B, fixed arcs (solid) between block entries on a train route, and pairs of alternative arcs (dashed) between conflicting operations. Adapted from Mazzarello & Ottaviani (2007).

The alternative arcs represent conflicting pairs of operations, where the determination of a train passing order and a minimum train headway is required. This is done through selecting one arc of each pair of alternative arcs. The selection process is modelled by disjunctive constraints, formulated as follows:

$$entry_j\text{–}entry_i \geq a_{ij} \ \vee \ entry_k\text{–}entry_h \geq a_{hk}, \ \forall \text{ pairs of alternative arcs } ((i,j),(h,k)),$$

with *entry* the block entry time of a train and a the weight of the alternative arc between the nodes corresponding to the entries. In the models presented by D'Ariano et al. (2007a) and Mazzarello & Ottaviani (2007), the alternative arcs connect subsequent block entries of pairs of running trains backwards, as shown in Figure 2.3. As a result, the running time over the block is directly included into the headway time. The rest of the headway is determined by the weight of the alternative arcs. D'Ariano et al. (2007a) include the time between the block exit of the head and of the tail of the train, i.e., the clearing time, supplemented with a default value representing the route setup and release time. In Mazzarello & Ottaviani (2007), the weight consists of a fixed and a variable term corresponding to blocking time components. The latter term depends on the train length and speed, so it includes the clearing time. Note that the time required to traverse the number of sections corresponding to the number of signalling aspects, i.e., the approach time, is not considered in the headway calculation in these models. In a subsequent version of the AG model, Corman et al. (2009) address the omission by incorporating the approach time in the headway calculation by connecting alternative arcs between two nodes corresponding to block entries that are the number of signal aspects apart. Figure 2.4 illustrates the difference in the construction of the alternative graph.

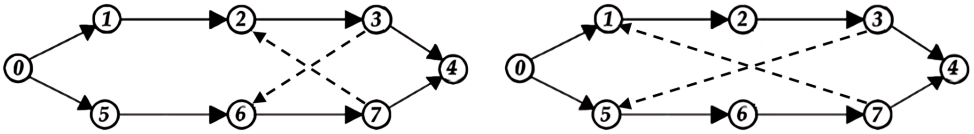


Figure 2.4: An alternative graph that does not include approach time in the headway (left), and the same alternative graph that does (right) for three-aspect signalling. Adapted from Mazzarello & Ottaviani (2007).

AG-based CDR models require the timetable and the current delays as input, next to the AG representation of the infrastructure. The model objective is to minimise the maximum secondary delay, i.e., the delay trains face due to interaction with trains suffering initial delays. Compliant with this objective, a graph-based solution algorithm is employed to find a selection of arcs that contains source-to-sink paths for all trains, while avoiding the presence of any positive directed cycle and minimising the weight of the critical path. An example of a heuristic solution algorithm is the truncated branch-and-bound algorithm proposed by D'Ariano et al. (2007a). To extend the model to consider objective functions beyond maximum secondary delay, it is necessary to transform the model into a MILP formulation. In that case, the AG serves as an intermediate model to specify the problem and to derive the disjunctive MILP formulation.

The AG approach is particularly suitable for retiming and reordering, though efforts are made to include rerouting (D'Ariano et al., 2008; Mazzarello & Ottaviani, 2007). Other extensions of D'Ariano et al. (2007a) consider the inclusion of speed control. D'Ariano et al. (2007b) propose an iterative approach similar to the method used by Mazzarello & Ottaviani (2007). This iterative approach involves an overarching traffic control algorithm that iterates between the fixed-speed CDR model and an external speed profile optimisation model. Corman et al. (2009) apply the so-called green wave approach, which assumes trains to only stop and wait at stations.

Disjunctive mixed integer linear programming models

In a disjunctive MILP CDR model, the railway operations are described by decision variables and mixed integer linear constraints, complemented with a linear objective function (e.g., Luan et al. (2018), Pellegrini et al. (2014) and Törnquist & Persson (2007)).

The decision variables indicate which train passes which track section at what time. The (re)scheduling decision variables relate to: timing, i.e., at what time does a train pass; ordering, i.e., which one of a pair of trains enters first; and routing, i.e., which route does a train take (Luan et al., 2018; Pellegrini et al., 2014; Törnquist & Persson, 2007). Time variables are typically continuous, indicating entry and exit times of trains on track sections. Binary variables decide on the order in which trains enter sections, as well as on the route trains are assigned.

To obtain linear constraints, disjunctive MILP models rely on the general big-M linearisation method (Bazaraa et al., 2008). The drawback of this method is that it results in a relatively weak linearisation, which makes it hard to solve the model to optimality (Reynolds et al., 2020). For this reason, improved solution methods are proposed in follow-up works of both Törnquist & Persson (2007) and Pellegrini et al. (2014) (Törnquist, 2012; Pellegrini et al., 2015, 2019).

The big-M constraints are used to describe the capacity. More specifically, they are the disjunctive constraints that ensure that possibly conflicting operations (or events) are separated in time. A generic pair of big-M constraints for the CDR problem is represented by the following equations:

$$start_{t',s} - end_{t,s} \geq \Delta_{t,t',s} + M\,(1 - order_{t,t',s}),\ \forall t,t',s,$$

$$start_{t,s} - end_{t',s} \geq \Delta_{t',t,s} + M\,order_{t,t',s},\ \forall t,t',s,$$

with t,t' trains, s a track section, *start* and *end* the start and end time, respectively, Δ the minimum separation time, M a sufficiently large number (the big M), and *order* a binary ordering variable. Depending on whether train t goes before train t' on section s (if $order_{t,t',s} = 1$), a certain minimum separation time ($\Delta_{t,t',s}$ or $\Delta_{t',t,s}$) is to be respected between the two relevant operations, e.g., the occupation of section s by trains t and t'.

In Pellegrini et al. (2014), operations are separated by determining the order in which a track-clear detection section is blocked by a train according to blocking time theory. All blocking time components are included. Based on the number of aspects in the considered multi-aspect signalling system, a reference section is determined for each track-clear detection section. The blocking starting time is linked to the entry time of the reference section. The running time from the reference section to the considered track detection section is based on minimum running times, possibly with an additional component in case of a decision to run slower or to perform an unplanned stop. Only in absence of conflicts and at planned stops the train times are strictly aligned with the scheduled times.

Luan et al. (2018) also consider the full blocking times, but on the level of block sections. The blocking times dynamically depend on train speeds. A mixed integer non-linear programming model is converted into a MILP by approximating the non-linear terms with piece-wise affine functions. In this model, speed is included through predetermined speed profile alternatives per train-block section pair.

In Törnquist & Persson (2007), both station areas and open lines between them are considered in terms of sets of parallel block sections. Being a fixed-speed model, the minimum

running time over a block section is considered as the duration of the activity associated with a pair of train and block section. For the separation of events associated with different trains and one block section, however, a fixed time is used, independent of the specific (parallel) block section or train.

Lamorgese & Mannino (2015) address the limitations of big-M formulations by introducing an exact macro/micro decomposition of the CDR problem, inspired by the Benders' decomposition approach. In the master problem, stations are macroscopically represented by nodes. On this macroscopic level, the CDR model is solved exactly, proposing tentative arrival and departure times. In each subproblem, a station is represented microscopically. For each train, a route through the station is searched, to fit the tentative timings. If the subproblem is infeasible, the violations are added to the master problem in an iterative process. The decomposition approach significantly reduces the number of big-M constraints and it has been applied in pilot tests in practice. In follow-up works, the approach is enhanced to speed up the solution process (Lamorgese et al., 2016) and the decomposition is further strengthened by removing the big-M constraints from the master problem (Lamorgese & Mannino, 2019).

The reason to translate an AG model into a MILP, is the flexibility in terms of objective functions (Samà et al., 2015). As long as the objective is formulated as a linear function, all types of objectives can be included into the optimisation model such as a weighted sum of delays, see, e.g., Luan et al. (2018), Pellegrini et al. (2014, 2015) and Törnquist & Persson (2007).

Time-indexed mixed integer linear programming models

Time-indexed MILP models are considered as an alternative to disjunctive MILP formulations as they are known for their strong linearisation and good approximations for scheduling problems (Van den Akker et al., 2000). The CDR problem is formulated as a time-indexed MILP by Bettinelli et al. (2017), Lusby et al. (2013) and Reynolds et al. (2020).

Time-indexed MILP models consider a uniform time discretisation on top of the fixed-block space discretisation. The model formulation builds on resources which correspond to pairs of a time unit and a track part (Reynolds et al., 2020; Lusby et al., 2013; Bettinelli et al., 2017). In Figure 2.5, a time-space graph is constructed in which nodes correspond to time-space resources and the arcs to parts of possible train routes indicated by source-sink paths. In general, the time-indexed approach is well-suited for the consideration of rerouting (Reynolds et al., 2020; Lusby et al., 2013; Zwaneveld et al., 2001; Bettinelli et al., 2017).

Typically, binary variables indicate the use of a time-space resource, and capacity is modelled using set packing constraints (Reynolds et al., 2020; Lusby et al., 2013; Bettinelli et al., 2017) such as the following:

$$\sum_{t \in T} x_t^r \leq 1, \ \forall r,$$

with T the set of trains and r a resource. In this generic set packing constraint, the number of trains t assigned to time-space resource r ($x_t^r = 1$) is restricted to at most one. With this, the occurrence of track conflicts is excluded. Resources can also be defined in such a way that they have a capacity greater than 1. For example, when the infrastructure is represented at a macroscopic level in which resources correspond to, e.g., parallel tracks or platforms (Bettinelli et al., 2017).

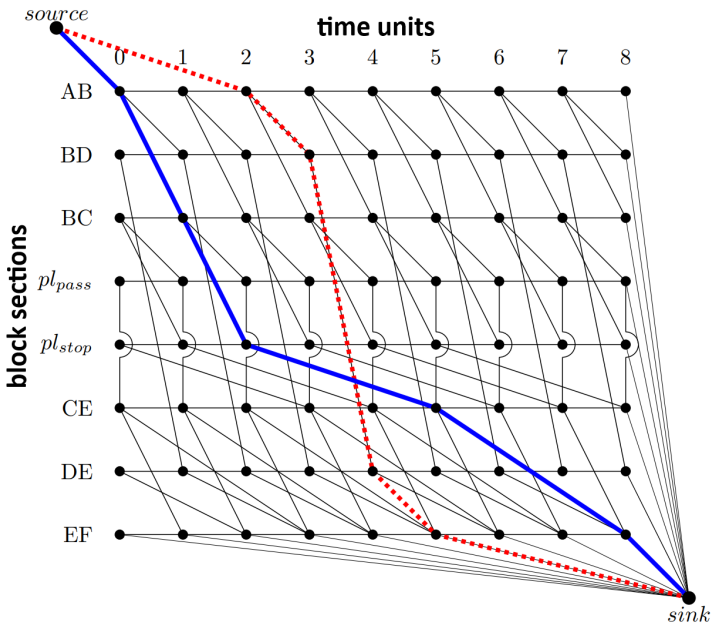


Figure 2.5: Example time-space graph with two train routes, i.e., source-sink paths. Adapted from Reynolds et al. (2020).

Time-indexed modelling requires to express the running and headway times with a precision of the time unit. Reynolds et al. (2020) estimate a 'near minimum' running time based on historical real-life data and consider the minimum headway as this running time plus an additional separation time, e.g., 30 seconds, based on blocking time theory.

The main drawback of the time-indexed approach is the model size (Van den Akker et al., 2000; Reynolds et al., 2020). This can be a restriction in the real-time application of the approach (Reynolds et al., 2020). Several solution methods are proposed to enable the use of the approach for CDR. A validated combination is the use of a heuristic method and a model decomposition (Lusby et al., 2013; Reynolds et al., 2020), while Bettinelli et al. (2017) propose a parallel algorithm based on an iterative greedy approach. The time-indexed formulation allows for various objective functions.

Time-indexed models can consider speed beyond fixed-speed assumptions. Lusby et al. (2013) and Reynolds & Maher (2022) approximate speed profiles by considering the alternatives of running at constant speed and accelerating/decelerating.

2.3.3 Conflict detection and resolution under fixed-block distance-to-go signalling

The wors of Gonzalez et al. (2010) and Mera et al. (2016) on fixed-block distance-to-go signalling nicely illustrates the role of speed on train headways and line capacity at a microscopic level. For the fixed-block distance-to-go signalling system implemented on the Madrid metro line, Gonzalez et al. (2010) describe a line capacity optimisation algorithm. Speed profiles are calculated for trains to come from the scheduled speed (40-90 km/h) to a standstill at a stopping point, i.e., at a station or at the entry of an occupied section. These speed profiles are input for an investigation of track section lengths (typically 100-200 m) and the effects on the

capacity parameters of train headway and running times. In Mera et al. (2016), the metro line capacity algorithm is enhanced with the introduction of speed signalling. Based on the number of free sections ahead, a target speed (code) is communicated to the train. This is done per section, obtaining a discrete braking curve.

Literature on CDR under fixed-block distance-to-go signalling is available on the Chinese signalling system, i.e., Level 3 of the Chinese Train Control System (CTCS Level 3), also called the quasi-moving block signalling system in Chinese literature. Note that the CTCS Level 3 system is designed for the Chinese high-speed (up to 300 km/h) railway lines. In accordance with the distance-to-go system characteristics (as described in Section 2.2), CDR models for CTCS Level 3 consider speed-dependent running times (Liu et al., 2021; Xu et al., 2017). The literature proposes different speed modelling approaches to allow a direct modelling relation between speed and train headway. Figure 2.6 illustrates this relation, showing that the speed (level) is translated one-on-one to the number of free block sections needed as minimum headway. The concept of a discrete set of speed levels is included in Xu et al. (2017) and its extension (Xu et al., 2021) to allow speed selection alongside the rescheduling decisions of retiming and reordering (and rerouting). The speed selection and retiming decisions are directly linked via running time constraints. Liu et al. (2021) also consider speed levels, but only in the initial solving phase. In the next step, continuous speed profiles matching the solution are selected from a predefined set. If no speed profile can be found, another iteration of the CDR model is required. Also in Xu et al. (2017), the model is first simplified. To obtain an initial solution for the whole problem, some speed-related variables are fixed according to the scheduled situation, so assuming no effect from the considered disturbances.

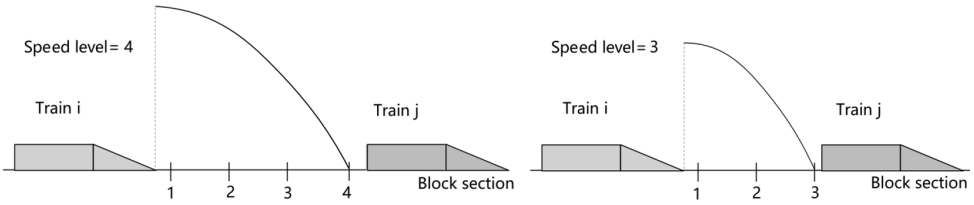


Figure 2.6: Example of relation between speed level and train headway in terms of block sections. Adapted from Liu et al. (2021).

The presented CDR models for CTCS Level 3 rely on a MILP formulation (Liu et al., 2021; Xu et al., 2017). Xu et al. (2017) build upon the AG model introduced in D'Ariano et al. (2007a), resulting in a AG-based MILP. The original model is extended by introducing pairs of alternative arcs related to a possible track conflict for each speed level. For a feasible solution, still only one alternative arc related to this conflict can be selected. As objective function, the total final secondary delay is considered in Xu et al. (2017), in its extension (Xu et al., 2021) and in Liu et al. (2021).

2.3.4 Conflict detection and resolution under moving-block signalling

The detection and resolution of conflicts under moving-block signalling is barely addressed in the literature. Pochet et al. (2016, 2017) consider CDR for mixed CBTC and non-CBTC operations on suburban railways. A model predictive control approach is presented in a rolling horizon framework, with a genetic algorithm optimising punctuality and regularity indicators.

An example of the latter is the deviation from the scheduled headway intervals. Overall, there is no notion of speed. In case of delays up to three minutes, the model can adjust running times and/or dwell times, considering fixed arrival and departure headways. The approach is incorporated in a microscopic simulation tool of the French train operator SNCF. In the practical experiments described, few scenarios in terms of disturbances and networks are considered, and the change of train orders is allowed. Meunier et al. (2021) follow up on the work of Pochet et al. (2017), introducing inter-station running times dependent on traffic conditions, considering disturbances. In their disjunctive MILP model, the running times on single-track lines are related to departure headways based on empirical simulation data.

Janssens (2022) presents a model concept based on the fixed-block CDR model described in D'Ariano et al. (2007a). In the development of a moving-block CDR model, the possibility is explored to model station areas and the open line, i.e., the track between stations, distinctively. The fixed block sections in the station areas are modelled as in the AG model described in D'Ariano et al. (2007a). For the moving block sections, the model is extended with virtual nodes and corresponding fixed arcs and pairs of alternative arcs. These model components are visualised in Figure 2.7, showing an alternative graph representation of two trains on a single track under moving-block signalling. Note that, in the figure, the alternative arcs are already selected according to train ordering. The virtual nodes correspond to grid points resulting from a (fixed) discretisation of the line based on train lengths. The virtual nodes are connected by fixed arcs whose weights correspond to minimum train clearing times, i.e., the time it takes the train to traverse its length when running at maximum speed. The alternative arcs connect subsequent nodes of two trains, indicating the order of the trains and the minimum headway between them. The moving-block minimum headway is derived from the fixed-block blocking times, leaving out the (fixed) block traversing time. For a small case study, the approach provides a valid result within reasonable time using a commercial solver, with the objective to minimise the maximum secondary delay. However, the model size significantly increases because of the additional components. Moreover, the model does not consider rerouting and adopts the fixed-speed assumption from D'Ariano et al. (2007a).

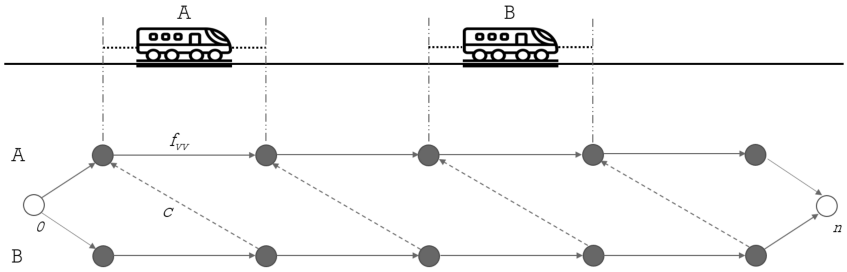


Figure 2.7: Alternative graph for two trains on a single track under moving-block signalling. The virtual nodes (dark) represent the discretisation of the line into train lengths. Adapted from Janssens (2022).

The general moving-block literature showcases the relevance of including speed modelling beyond fixed-speed assumptions, with the speed-dependency of the moving-block headway as a main point of focus (Gao et al., 2020; Liu, 2016; Wang et al., 2014; Xu et al., 2014). Gao et al. (2020) point out that under moving-block signalling, the minimum headway is negatively correlated with the running time, which corresponds to the positive correlation between braking distance and speed: the lower the speed, the shorter the braking distance and therefore the

headway. The balance between the capacity, as result of the minimum headway, and effective operations, related to the running time or speed, is addressed by, e.g., Liu (2016) and Xu et al. (2014). Liu (2016) proposes a dynamic system approach to address the dual control problem of finding the optimal speed and a safe following distance. Xu et al. (2014) dynamically optimise the balance between running times and headways, e.g., by introducing speed limits to reduce braking distances and, hence, headways.

Dynamic system approaches are well-represented in moving-block literature (e.g., Gao et al. (2020), Liu (2016), Wang et al. (2014) and Xu et al. (2014)). Dynamic systems are characterised by the use of differential equations, and they are widely applied in optimal train control and train trajectory models (Gao et al., 2020; Liu, 2016; Wang & Goverde, 2019). In the optimal control models of Gao et al. (2020) and Liu (2016), the focus is on speed profile optimisation depending on train dynamics and constrained by track characteristics. This train trajectory optimisation is done for individual trains. The modelling of moving-block headways would require cross-terms between the state trajectories of successive trains resulting in complex multi-train trajectory optimisation models (Wang & Goverde, 2019). Wang & Goverde (2019) apply multi-train trajectory optimisation in the context of fixed-block timetabling, considering timing, ordering and routing options.

Wang et al. (2014) implement the problem of train separation within the context of trajectory optimisation using the pseudospectral method and by formulating a MILP, both solved with commercial solvers. The two methods show similar performance with respect to the balance between solution quality and computation time, even though the linearisation necessary for the MILP results in a discrete-space model. (Disjunctive) MILP modelling is also used in timetable optimisation for moving-block operations (Schlechte et al., 2022). Train orders and routes are scheduled with headways based on emergency braking curves consistent with a predefined set of possible speeds.

Besides the speed-dependency and the infrastructure modelling, also moving-block as a system needs to be considered for the modelling of moving-block headways. In the context of ERTMS/ETCS Level 2 Moving Block, Büker et al. (2019) incorporate the system characteristics such as the bidirectional radio communication and the onboard train positioning and integrity into the blocking time components.

2.3.5 Summary

Table 2.1 provides a summary of the reviewed papers specifically on CDR modelling. The table classifies a representative selection of papers for fixed-block multi-aspect, fixed-block distance-to-go and moving-block signalling in terms of the modelling aspects introduced in Section 2.3.1. Publications are deemed representative if appointed as such in the general literature, e.g., in the review papers of Cacchiani et al. (2014) and Corman & Meng (2015), and/or if they are unique in terms of the considered modelling aspects. In this way, Table 2.1 provides a concise overview of the type of models existing for CDR. In case of multiple publications by the same authors that do not differ much in terms of aspects, the first paper is included.

It emerges that the most commonly applied modelling approaches are AG, disjunctive MILP (D-MILP) and time-indexed MILP (TI-MILP), as also found in the literature (Reynolds et al., 2020). Note that AG-based MILP (AG-MILP) is an example of a disjunctive MILP approach.

Table 2.1: Overview of literature on conflict detection and resolution under fixed-block multi-aspect, fixed-block distance-to-go and moving-block signalling.

Publication	Modelling approach	Infra modelling	Speed modelling	Headway modelling	Re-routing	Solution method	Objective function
Fixed-block multi-aspect signalling							
Bettinelli et al. (2017)	TI-MILP	BS	Fixed	Blocking	Yes	Iterated greedy	Total arrival and departure
Corman et al. (2009)	AG	BS	Green wave	Blocking	No	Heuristic	Max secondary
D'Ariano et al. (2007a)	AG	BS	Fixed	Occupation	No	B&B	Max secondary
D'Ariano et al. (2007b)	AG	BS	Iterative	Occupation	No	B&B	Max secondary
D'Ariano et al. (2008)	AG	BS	Fixed	Occupation	Yes	B&B	Max secondary
Lamorgese & Mannino (2015)	D-MILP	BS	Fixed	Default	Yes	Macro/micro decomposition	Total arrival delay cost
Luan et al. (2018)	D-MILP	BS	Levels	Blocking	No	Two-level	Total mean absolute
Lusby et al. (2013)	TI-MILP	TDS	Variable	Occupation	Yes	B&P	Weighted total
Mazzarello & Ottaviani (2007)	AG	BS	Iterative	Occupation	Yes	Two-level	Max secondary
Pellegrini et al. (2014)	D-MILP	TDS	Fixed	Blocking	Yes	Solver	Max/total secondary
Pellegrini et al. (2015)	D-MILP	TDS	Fixed	Blocking	Yes	Heuristic	(Weighted) total final
Reynolds et al. (2020)	TI-MILP	TDS	Fixed	Blocking	Yes	B&P	Custom utility
Reynolds & Maher (2022)	TI-MILP	TDS	Levels	Blocking	Yes	B&P	Custom utility
Samà et al. (2015)	AG-MILP	BS	Fixed	Occupation	No	Solver	Multi-criteria
Törnquist & Persson (2007)	D-MILP	BS	Fixed	Default	Yes	Solver	Total/weighted final
Törnquist (2012)	D-MILP	BS	Fixed	Default	Yes	Greedy depth-first	Total final
Fixed-block distance-to-go signalling							
Liu et al. (2021)	MILP	BS	Iterative	Default	No	Bi-level	Total final secondary
Xu et al. (2017)	AG-MILP	BS	Levels	Default	No	Two-step	Total final secondary
Xu et al. (2021)	AG-MILP	BS	Levels	Default	Yes	Two-step	Total final secondary
Moving-block signalling							
Janssens (2022)	AG	(M)BS	Fixed	Occupation	No	Solver	Max secondary
Meunier et al. (2021)	AG-MILP	Micro	Variable	Blocking	No	Solver	Punctuality regularity
Pochet et al. (2016)	MPC	Micro	Fixed	Default	No	Genetic algorithm	Punctuality regularity

AG: alternative graph, D-MILP: disjunctive mixed integer linear program, TI-MILP: time-indexed mixed integer linear program, AG-MILP: AG-based mixed integer linear program, MPC: model predictive control, (M)BS: (moving) block section, TDS: track-clear detection section, B&B: branch-and-bound, B&P: branch-and-price, Max: maximum.

An exception is the model predictive control (MPC) model applied by Pochet et al. (2016) in a suburban context. Each approach has its strengths and weaknesses in terms of, e.g., inclusion of speed modelling, suitability for rerouting, computational performance and flexibility in objective functions. For example, AG models are restricted to path-based objectives, such as minimising the maximum (secondary) delay.

Practically all CDR models describe the infrastructure at a microscopic level. Within the microscopic level, a distinction can be made based on whether or not the block sections (BSs) are represented as a sequence of track-clear detection sections (TDSs). The main advantage

of considering the infrastructure in terms of TDSs is the possibility to model sectional route release in interlocking areas, which can provide capacity benefits (Hansen & Pachl, 2014). Nevertheless, the TDSs are not often modelled.

Conventionally, CDR models are fixed-speed models. While fixed-speed models capture most essential characteristics and provide highly practical solutions in terms of rescheduling decisions (Pellegrini et al., 2014), they may lead to infeasible solutions in terms of exact timings due to their reliance on minimum or scheduled running times in case of disturbances (Reynolds & Maher, 2022). As alternative to the fixed-speed assumptions, the modelling of speed levels is introduced. This allows for a better capture of the speed-dependency of the headway and for the usage of the speed variability to influence the headway. Specifically for the modelling of fixed-block distance-to-go operations, there is the tendency to move away from fixed-speed assumptions.

The modelling of train headways is clearly interwoven with other modelling aspects such as the modelling approach and the speed modelling. Independently of other aspects, the options are to determine minimum train headways based on the physical occupation, the longer blocking times of trains, or to fix them at default values. This generally follows from the foreseen model application. For instance, the default (arrival) headway between homogeneous suburban trains (Pochet et al., 2016).

The rescheduling measures of retiming and reordering are always considered, even for the CBTC application on suburban lines (Pochet et al., 2016). Rerouting decisions, however, are only included in part of the models. If rerouting is considered, it is regularly in follow-up work. This is, for example, the case with the rerouting model for fixed-block distance-to-go, i.e., Xu et al. (2021). Notable is the lack of rerouting in moving-block applications. The general reason is twofold. First, in practice, disturbances resulting in delays of a couple of minutes are typically resolved with retiming and reordering alone. Second, incorporating rerouting presents modelling challenges, primarily related to computational performance (Pellegrini et al., 2015).

The literature on CDR is dominated by heuristic solution methods (Cacchiani et al., 2014; Corman & Meng, 2015). Indeed, various heuristics method to solve CDR problems are proposed and applied. Even the commercial solvers typically rely on heuristics when taking into account a limited computation time. The two favoured streams of heuristic methods are tailored branch-and-bound (B&B), among which truncated branch-and-price (B&P) algorithms, and multi-step optimisation methods.

The objective of minimising the impact of disturbances on the network is generally considered by minimising the propagation of train delays. In line with that, practically all model objectives are a type of delay minimisation. While the AG models are limited to the minimisation of the maximum (secondary) delay, the MILP models include various (linear) objective functions, e.g., total mean, weighted total or final secondary delay. The exception to delay minimisation is the maximisation of punctuality and regularity in a suburban context by Meunier et al. (2021) and Pochet et al. (2016). Note that not only are different objectives not compatible with all modelling approaches, but also that different objectives can lead to different (optimal) solutions for the same model.

2.4 Research gaps and challenges

This section is dedicated to the identification of research gaps and the related challenges regarding the modelling of CDR under moving-block signalling. The literature review in Section 2.3 shows that the majority of existing CDR models refers to fixed-block signalling systems whereas only little research addresses CDR under moving-block signalling, leaving a clear gap to fill in future research.

The main characteristic for the modelling of railway traffic under moving-block signalling is the minimum train headway based on absolute braking distance. For minimum headways, default values can be determined per section, per train and/or per speed level. The alternative is to determine headways using blocking time theory such that they rely on the time a track part is assigned to a train. The blocking time includes the occupation time, during which a train is physically present on the considered block section, plus the other blocking time components including in particular the approach time corresponding to the braking behaviour before the block section. Hence, to model signalling conflicts, and not only physical conflicts, minimum headways should include all blocking time components.

Overall, the gap to bridge in the research on CDR under moving-block signalling is the development of a model that describes railway traffic operating at moving-block characteristic headways. This gap gives rise to two main challenges to be faced in future research.

The first challenge is to adapt CDR models such that the EoA can be assigned to any point on the open line instead of only at block entries as customary under fixed-block signalling. This yields the need of a continuous modelling of the infrastructure rather than the conventional discretisation into track sections. A continuous representation of the infrastructure would also accommodate IPs anywhere on the track, including in station areas where fixed sections remain in place to protect switches in interlocking areas.

The second challenge relates to the modelling of the headways as a continuous function of train speed. While in fixed-block multi-aspect signalling system the role of speed is limited, in signalling systems with distance-to-go ATP the impact of speed on rescheduling decisions increases. However, the extent of the impact of modelling a fully continuous speed-headway relation is unclear. The effect on rescheduling decisions, the main output of the model, may be limited; these decisions are generally taken in interlocking areas where speed limits and protective blocks are in place, possibly diminishing the effects of the continuous speed-headway relation.

2.5 Research agenda

A research agenda is drafted based on the gaps and challenges identified in Section 2.4. Four main research steps towards CDR under moving-block signalling are: 1) an analysis of infrastructure modelling options, 2) an analysis of speed modelling options, 3) the identification of modelling options for moving-block CDR, and 4) an assessment of modelling approaches for the identified modelling options. Figure 2.8 presents the research steps and their connections. In the following, the research steps are addressed to provide stepping stones in the research on CDR under moving-block signalling.

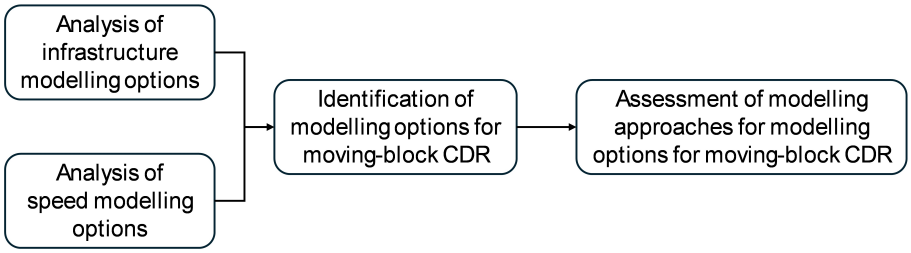
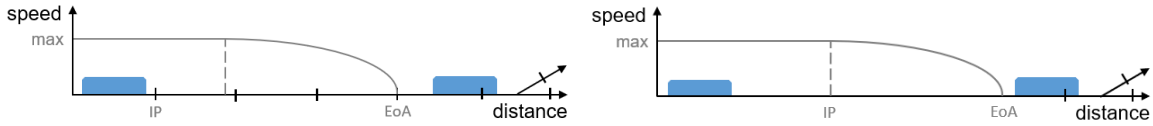


Figure 2.8: Flow chart of the proposed research agenda.

2.5.1 Infrastructure modelling options

An analysis of infrastructure modelling options can help to assess the impact of infrastructure modelling on CDR models in terms of, e.g., delay propagation minimisation, rescheduling decisions and computational performance. Various modelling options, discrete and continuous, can be considered and compared.

In case of a discrete representation of the infrastructure, the open line is divided into sections by a discretisation grid. Because of the infrastructure discretisation, an MA up to the position of the tail of the preceding train cannot be modelled. Instead, the EoA lies at the grid point that is released last by the preceding train. Similar to the EoA, the IP, i.e., where an approaching train receives the indication to start braking to stop at the EoA, has to be related to a discrete grid point. The effect of a discrete representation of the infrastructure on the EoA and the IP is illustrated in Figure 2.9a. The impact of discretising the infrastructure is expected to strongly depend on the type and fineness of the discretisation.



(a) Discrete infrastructure. *(b) Continuous infrastructure.*

Figure 2.9: Illustrative examples of infrastructure modelling options: (a) discrete and (b) continuous. The end of authority (EoA) and the brake indication point (IP) for maximum speed are indicated.

There are roughly two types of discretisation that can be applied: equidistant discretisation or variable discretisation. With equidistant discretisation, the interval length between grid points is set equal to an overall fixed value. This value can be, e.g., derived from train lengths or from the inherent moving-block system discretisation due to discontinuous communication. Based on this, 200 metres can be considered as starting point in the search for an appropriate discretisation interval (Furness et al., 2017). With variable discretisation, the interval lengths can vary over the infrastructure. For example, shorter intervals can be considered closer to critical points such as in station areas, or interval lengths can be related to the (maximum) track speed (the lower, the shorter).

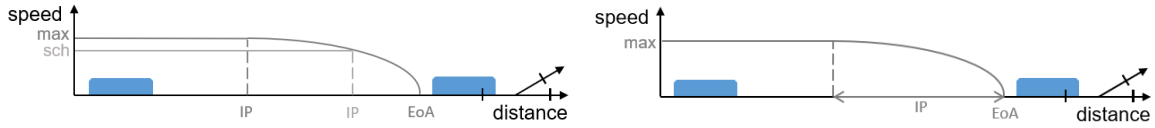
The consideration of the open line in continuous space directly affects the EoA and the IP. In continuous space, the EoA can be set at a safety margin behind the tail of the preceding train. The approaching train can then run freely up to the ‘actual’ braking point, i.e., where the

train should start braking. Figure 2.9b illustrates the effect of a continuous representation of the infrastructure on the EoA and the IP.

2.5.2 Speed modelling options

As for infrastructure modelling, an analysis of the speed modelling options can help to assess their impact on CDR models in term of, e.g., delay propagation minimisation, rescheduling decisions and computational performance. Various options to incorporate speed dynamics, discrete and continuous, can be considered and compared. Key is to keep in mind the possibly limited effect due to speed limits and protection zones/blocks around decision points at switches.

With speed modelled discretely, braking distances can be (pre)calculated for each of the discrete speed alternatives. The speed alternatives can correspond to, e.g., speed levels or speed profiles. Speed levels can be defined to align with different track speed (limits), while speed profiles can be derived from the scheduled and/or minimum running times. In both cases, the impact of discretising speed strongly depends on the fineness of the discretisation, i.e., the number of alternatives. Figure 2.10a relates the discrete modelling of speed to the EoA and the IP; the EoA is independent of speed, but the IP can be different for each speed alternative.



(a) Discrete speed. *(b) Continuous speed.*

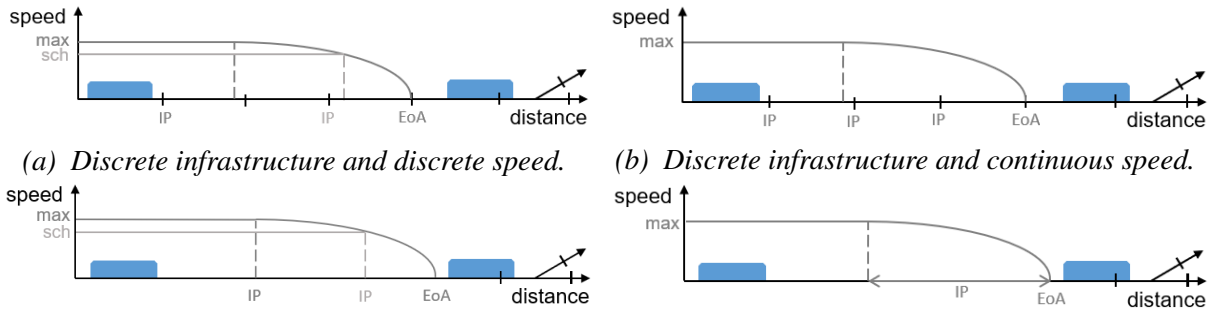
Figure 2.10: Illustrative examples of speed modelling options: (a) discrete and (b) continuous. The end of authority (EoA) and the possible brake indication points (IP) are indicated based on maximum (max) and scheduled (sch) speed.

With continuous modelling of speed, the minimum headway can be better approximated by computing the braking distance based on the actual speed. Hence, the IP cannot be precomputed. As illustrated in Figure 2.10b, the IP can lie anywhere between the start of the braking curve when running at maximum speed and the EoA.

2.5.3 Modelling options for moving-block conflict detection and resolution

Following from the infrastructure and speed modelling options, the modelling options to consider in the development of a CDR model for moving-block signalling are: 1) discrete infrastructure and discrete speed, 2) discrete infrastructure and continuous speed, 3) continuous infrastructure and discrete speed, and 4) continuous infrastructure and continuous speed. The four options are shortly described in the following and illustrated in Figure 2.11.

Figure 2.11a features an equidistant infrastructure discretisation and two speed profile alternatives (maximum and scheduled), to exemplify the modelling option of discrete infrastructure and discrete speed. Indeed, the EoA lies at the grid point that is last released by the preceding train and the speed-dependent IP at the grid point before the start of the relevant braking curve.



(a) Discrete infrastructure and discrete speed. *(b) Discrete infrastructure and continuous speed.*

(c) Continuous infrastructure and discrete speed. *(d) Continuous infrastructure and continuous speed.*

Figure 2.11: Illustrative examples of moving-block modelling options in terms of the modelling of the infrastructure (discrete on the top versus continuous on the bottom) and speed (discrete on the left versus continuous on the right). The end of authority (EoA) and the possible brake indication points (IP) are indicated.

Figure 2.11b features an equidistant infrastructure discretisation and continuous speed, to exemplify the modelling option of discrete infrastructure and continuous speed. Indeed, the EoA lies at the grid point that is last released by the preceding train and the speed-dependent IP at the grid point before the start of the braking curve of the actual train speed.

Figure 2.11c features a continuous infrastructure and two speed profile alternatives (maximum and scheduled), to exemplify the modelling option of continuous infrastructure and discrete speed. Indeed, the EoA lies at a safety margin behind the tail of the preceding train and the speed-dependent IP at the start of the relevant braking curve.

Figure 2.11d exemplifies the fourth modelling option by featuring a continuous infrastructure and continuous speed. Indeed, the EoA lies at a safety margin behind the tail of the preceding train and the speed-dependent IP at the start of the braking curve of the actual train speed, which can be anywhere on the track between the IP related to the maximum speed profile and the EoA.

The modelling options can be further specified based on the first two research steps. Within the underlying infrastructure and speed modelling options, the possibilities may be narrowed down and/or a candidate possibility may be put forward. Crucial for any option is the balance between computational effort and solution quality, both in terms of the approximation of the moving-block headway and the optimality of the rescheduling decisions. This balance does not only depend on the modelling options but also on the modelling approach. Hence, modelling approaches are analysed with respect to the presented modelling options in Section 2.5.4.

2.5.4 Modelling approaches for moving-block conflict detection and resolution

For the four modelling options, the CDR modelling approaches of AG, disjunctive MILP and time-indexed MILP, as well as the moving-block dynamic system approach are analysed. The comparative analysis is based on the advantages and disadvantages of the approaches, which are summarised in Table 2.2.

Table 2.2: Advantages, disadvantages and proposed moving-block modelling options of modelling approaches.

Approach	**Advantages (+) / Disadvantages (–)**	**Proposed modelling option(s)**
Alternative graph	(+) Continuous time (+) Model extension for moving block (–) Iterative speed modelling (–) Rerouting requires meta-heuristic (–) Decision points from infra disretisation (–) Principally headway based on occupation (–) Single objective function	Discrete infrastructure and discrete speed Discrete infrastructure and continuous speed
Disjunctive MILP	(+) Continuous time (+) Speed level applications (+) Suitable for rerouting (+) Headway based on blocking time (+) General objective function (–) Decision points from infra discretisation (–) Weak linearisation due to big-M (–) Linear objective and constraints	Discrete infrastructure and discrete speed Discrete infrastructure and continuous speed
Time-indexed MILP	(+) Strong linearisation (+) Potential for speed modelling (+) Suitable for rerouting (+) General objective function (–) Imposed time discretisation (–) Model size (–) Set packing capacity constraint	Discrete infrastructure and discrete speed
Dynamic system	(+) Applied to moving-block operations (+) Continuous infrastructure (+) Dynamic speed-headway relation (+) Applied to timetabling (–) Not used in rescheduling (–) Complex multi-train models	Continuous infrastructure and discrete speed Continuous infrastructure and continuous speed

The three CDR approaches inherently rely on a discretisation of the infrastructure for the modelling of decisions. In all three cases, a finer discretisation grid will increase the model size significantly, whether in number of nodes and arcs (AG) or variables (MILP). For time-indexed MILP, the model size is already a restriction as it considers time-space resources as a consequence of its inherent time discretisation. In this also lies its strength, because it facilitates the strong model linearisation based on set packing constraints for the modelling of capacity. The set packing constraints require resources with a fixed capacity, which can restrict the model in its space and time discretisation. As the headway is expressed in space and/or time units, the headway is also limited in its flexibility.

The time-continuous disjunctive MILP does not have this restriction as its linearisation is based on the big-M method. The big-M constraints allow for a (more) flexible headway definition. However, it comes at a price: a weak linearisation that makes it harder to solve the model to optimality. The necessary linearisation for MILP models does not only apply to the model

constraints, but also to the objective function. Besides that, MILP objective functions are very general, in contrary to the single possible objective in AG models.

Two other aspects in which MILP models seem to outperform AG models are the suitability for rerouting and the potential to include speed dynamics. Both disjunctive MILP and time-indexed MILP models have considered speed alternatives and initially include rerouting, while the inclusion of speed modelling and rerouting in AG models requires external speed profile optimisation models and meta-heuristics. On a different note, only the AG approach has so far been applied to moving-block CDR.

Despite the difference between the approaches, they all build on the important aspect of headway modelling. The MILP models derive headways from train blocking times, while initial AG models only consider physical occupation times. This is not necessarily a restriction, because the missing approach times can be included by a slight reformulation. Anyway, modelling moving-block headways will require a reconfiguration, certainly in an attempt to include speed beyond fixed-speed assumptions or speed levels. For the AG approach, continuous speed would require to rebuild the graph after every changed speed decision because of the graph structure in which alternative arcs define headways by connecting infrastructure points. In theory, MILP models can include continuous speed variables. However, in time-indexed MILP the effects would be downscaled because of the expression of time in fixed units. Moreover, continuous speed would mean that running and clearing times will become variable as they can no longer be precalculated. With this, the requirement of linear constraints complicates the model formulation. For example in combination with route decisions because whether a route is used or not co-determines the running and clearing times.

In the end, only the dynamic system approach has proven to be able to describe the dynamic relation between speed and headway. The approach has been applied to general moving-block models, but not to CDR. Retiming decisions are included through the continuous speed modelling, but reordering and rerouting decisions have no clear connection to the dynamic system. Also of influence is the need to consider switches as discrete sections. In general, the consideration of multiple trains result in complex models which are hardly considered in literature.

To conclude, proposed modelling approaches per modelling option are (also see Table 2.2):

1. Discrete infrastructure and discrete speed: AG, disjunctive MILP or time-indexed MILP.
2. Discrete infrastructure and continuous speed: AG or disjunctive MILP.
3. Continuous infrastructure and discrete speed: dynamic system.
4. Continuous infrastructure and continuous speed: dynamic system.

2.6 Conclusions

Conflict detection and resolution (CDR) under moving-block signalling remains underexposed in the literature, in comparison to moving-block signalling and CDR in general. This chapter reviewed the literature on CDR models for fixed-block multi-aspect, fixed-block distance-to-go and moving-block signalling with the aim to identify research gaps and to propose future steps towards the development of moving-block CDR models.

The main gap is identified to be the lack of CDR models that accurately describe railway operations considering moving-block characteristic headways, based on absolute braking dis-

tances. Therefore, two main research challenges are the modelling of the infrastructure in continuous space and the inclusion of speed dynamics.

It is proposed to address the identified research gaps by an analysis of various options for infrastructure and speed modelling, followed by an assessment of modelling approaches for the different options. The CDR approaches of alternative graph (AG) and (disjunctive) mixed integer linear programming (MILP) are proposed for modelling options including a discrete infrastructure representation, while a shift to dynamic system approaches is suggested for the modelling of the infrastructure in continuous space.

Consistently with the proposal, future research should aim at (further) investigating the possibilities of the existing approaches of AG and disjunctive MILP in terms of infrastructure discretisation and speed alternatives. This has been taken up in the work of Versluis et al. (2023); a CDR model approximating moving-block operations is obtained by enhancing the disjunctive MILP model of Pellegrini et al. (2015). Furthermore, a comparison of the different modelling options is needed to determine the extent to which continuous modelling of infrastructure and speed brings advantages over discretised models in terms of solution quality and computational efficiency. Such a comparison will also improve the understanding of the context in which accurate modelling of moving-block operations does imply different rescheduling decisions compared to existing fixed-block CDR models. These insights are invaluable for the development of effective moving-block CDR models.

Chapter 3

Conflict detection and resolution under distance-to-go signalling

---

Chapter 2 proposed research steps for the modelling of conflict detection and resolution under moving-block as next-generation radio-based distance-to-go railway signalling system. A key characteristic shared by moving-block and other radio-based distance-to-go railway signalling systems is the implementation of continuous braking curve supervision. Moving-block railway operations can be approximated in discrete space. Therefore, modelling conflict detection and resolution for distance-to-go operations in discrete space can be considered as a step towards conflict detection and resolution under moving-block signalling.

In this chapter, a conflict detection and resolution model describing fixed-block distance-to-go railway operations is obtained by enhancing a state-of-the-art rescheduling model originally developed for conventional fixed-block signalling. In two case studies, the enhanced model is compared with the original one to investigate the effects of fixed-block distance-to-go over conventional fixed-block signalling in conflict detection and resolution.

Apart from minor changes, this chapter has been accepted for publication as:

Versluis, N. D., Pellegrini, P., Quaglietta, E., Goverde, R. M. P., & Rodriguez, J. (in press). Conflict detection and resolution for distance-to-go railway signalling. *Transportmetrica A: Transport Science.*

3.1 Introduction

Worldwide, railways are experiencing a continuously increasing travel demand. The existing railway networks have limited capacity and fewer and fewer extension possibilities due to the costly and land-consuming infrastructure. Especially at capacity bottlenecks such as in and around stations, extensions are often not available. To fulfil future railway demand in a different way, advanced signalling systems such as distance-to-go (DTG) systems are developed as capacity-increasing alternatives to conventional fixed-block multi-aspect signalling with automatic train protection systems.

In conventional fixed-block systems, the track is partitioned into blocks of fixed lengths which can be occupied by at most one train at a time. The block entries are protected by trackside multi-aspect signals which indicate whether an approaching train can proceed, needs to start braking or is required to stop. The automatic train protection system supervises that the trains follow the signal aspects and intervenes when a train does not start braking after a brake indication. Hence, train separation distances are determined based on a number of blocks of fixed lengths.

Radio-based fixed-block DTG systems feature radio-based cab signalling and continuous braking curve supervision from the train front position to the block (or section) entry corresponding to the train's end of movement authority (MA), i.e., the permission to move to a specific location under supervision. With this, brake indications no longer need to be provided at the entry of a block. Moreover, train separation distances become speed-dependent due to the approach distance to the first occupied block being based on the train's absolute braking distance. Note that with DTG referring to the dynamic speed profile supervision to a target distance, DTG is not synonymous to a fixed-block signalling system, which can also be implemented with stepwise braking at trackside signals (Schön et al., 2013).

For effective railway operations, real-time traffic management is crucial. In case of small delays originating from variations in rolling stock, dwell times and driver behaviour, traffic management can apply rescheduling measures such as retiming, reordering and local rerouting to resolve track conflicts while minimising the delay propagation in the network. To support human dispatchers in taking mathematically optimised rescheduling decisions, conflict detection and resolution (CDR) models describing variants of the real-time railway traffic management or rescheduling problem are developed (Cacchiani et al., 2014; Pellegrini et al., 2014). In short, the problem can be formally described as follows: given a railway network and timetable together with delayed events, find a new timetable by rescheduling trains such that all track conflicts are resolved and total delay (or other relevant objective) is minimised.

The existing CDR models mostly refer to conventional fixed-block signalling systems. As a result, they are typically microscopic models that inherently rely on the discretisation of the track into fixed blocks. Moreover, they generally neglect changes in the speed profile to fit the adjusted timetable (Reynolds & Maher, 2022). Hence, these models cannot accurately capture the speed-dependent separation distances typical for DTG signalling due to the implementation of braking curve supervision (Versluis et al., 2024).

In this chapter, we address the research gap regarding CDR under DTG signalling by developing a CDR model for DTG and applying it to assess the operational relevance of modelling CDR for DTG. To obtain a model that represents and identifies optimised rescheduling deci-

sions for DTG railway operations, we propose enhancements for existing CDR models. The enhancements include the introduction of speed profile options and train blocking times considering train- and speed-dependent braking distances. These two enhancements are significant as they are crucial in enabling the modelling of train- and speed-dependent brake indication points in CDR models. As a proof of concept, we incorporate the enhancements into the state-of-the-art CDR model of RECIFE-MILP (Pellegrini et al., 2015), together with a distinction between switches and the rest of the track. The resulting model is assessed through a comparative analysis with the original one in two case studies, specifically in terms of total train delay and rescheduling decisions.

With this, the main contributions of this chapter are:

- A modelling approach for enhancing conflict detection and resolution models to apply to distance-to-go signalling, introducing train- and speed-dependent train separation and brake indication points.
- A mixed-integer linear programming model for conflict detection and resolution under distance-to-go signalling, obtained by integrating distance-to-go principles into the state-of-the-art RECIFE-MILP model.
- A comparative analysis illustrating the operational relevance of incorporating distance-to-go principles into conflict detection and resolution modelling, as different rescheduling decisions achieving reductions in total train delay are proposed.

In line with the contributions, the chapter is organised as follows. In Section 3.2, the state-of-the-art related to the modelling of CDR for DTG signalling is provided. Section 3.3 describes the enhancement approach for CDR models to describe DTG signalling. The mathematical formulation of the fixed-block CDR model RECIFE-MILP enhanced for DTG signalling is presented in Section 3.4. Section 3.5 presents the results of the comparative analysis. Finally, Section 3.6 concludes the chapter.

3.2 State of the art in conflict detection and resolution for distance-to-go signalling

In this section, we provide an overview of the state-of-the-art relevant for the topic of modelling CDR under DTG signalling. Section 3.2.1 focuses on DTG signalling as the de-facto standard in the development of next-generation signalling systems. Section 3.2.2 provides an overview on modelling approaches for CDR existing in the literature.

3.2.1 Distance-to-go signalling systems

In the development of next-generation signalling systems, DTG is the de-facto standard. For the mainline railways, DTG signalling is included within the European Train Control System (ETCS) and the Chinese Train Control System (CTCS). For metro lines, Communications-Based Train Control (CBTC) applies DTG signalling principles. Examples of deployed DTG systems are the fixed-block ETCS Level 2 with Trackside Train Detection (TTD) and the similar CTCS Level 3. The latter is designed for the Chinese high-speed (up to 300 km/h) railway lines (Xu et al., 2017). In this chapter, we focus on ETCS Level 2 with TTD, hereafter referred to

as ETCS L2, which is implemented on freight and passenger railways throughout the world (European Rail Supply Industry Association, 2022).

To enable cab signalling, ETCS L2 features radio communication between a Radio Block Centre (RBC) on the trackside and the trains as follows. The RBC receives train position information from a train. This information supports the request for a route extension to the interlocking system. Based on safety logic and input from TTD, the interlocking system sets and locks a new route for the train. As a next step, the RBC updates the MA before sending it to the train. With the received MA, the train recomputes and supervises a dynamic speed profile including continuous braking curves.

Corresponding to the braking curves, the brake indication point for a train can lie anywhere along the track. The end of authority, however, is consistently set at the end of a block or TTD section (TDS). The TDSs correspond to the parts of the track in which the absence of trains is automatically detected. Whether the end of authority corresponds to a block or TDS depends on the implemented ETCS L2 variant. It can also depend on whether it is on the open line, in a station or around switches. A schematic representation of the ETCS L2 system is provided in Figure 3.1. In the figure, two trains at the speed-dependent minimum separation distance are included.

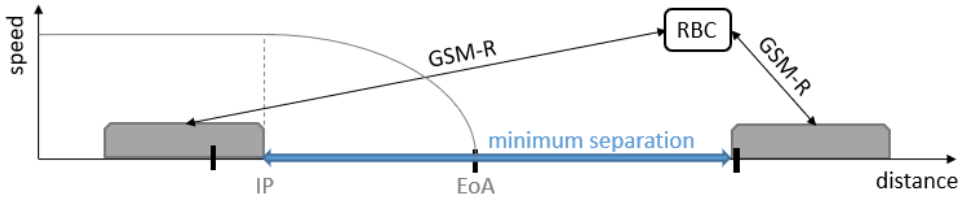


Figure 3.1: Schematic layout of the minimum separation between two trains in a speed-distance diagram under ETCS L2 with (block) section boundaries, Radio Block Centre (RBC) and radio communication (GSM-R). The separation is related to the brake indication point (IP) and the end of authority (EoA) depending on the braking curve.

Within ETCS and CTCS, more advanced application levels are being developed, but not yet deployed. In particular, ETCS Level 2 with Onboard Train Integrity Monitoring (TIM), previously known as ETCS Level 3. This allows the end of authority to be located at the end of the last free virtual (non-physical) block, e.g., in ETCS Level 2 Virtual Block, or even at a safety margin behind the rear position of a moving train, e.g., in ETCS Level 2 Moving Block (European Union Agency for Railways, 2016).

3.2.2 Conflict detection and resolution modelling

Based on the DTG signalling principles as described in Section 3.2.1, three relevant aspects for CDR models are the infrastructure representation, the assumptions related to speed and the modelling of train separations (Versluis et al., 2024). In the context of CDR, the infrastructure is typically modelled at the microscopic level (Cacchiani et al., 2014), which allows the use of the well-known blocking time theory (see, e.g, Hansen & Pachl (2014)) in the modelling of train separation. The alternatives of mesoscopic and macroscopic approaches are more often considered in case of disruptions rather than disturbances (Cacchiani et al., 2014; Zhan et al., 2022).

In this section, we provide an overview of the existing models and methods, to further specify the gap in the modelling of conflict detection and resolution for distance-to-go railway operations. First, the main types of model formulations in the literature are introduced. Then, solution methods are discussed in the light of practical applicability. Table 3.1 provides a summary of the reviewed literature aimed at comparing the work presented in this chapter with the key existing works. We specifically consider the following features: whether the infrastructure is modelled in terms of blocks or TDSs, how speed is modelled, whether and how speed is considered in the train separation distance, which rescheduling decisions are considered, which solution method is applied, what model objective is considered, and to what kind of scenario is it applied.

For a more extensive review of the existing literature on CDR models in the context of DTG signalling, we refer to Versluis et al. (2024).

Table 3.1: Comparison of this work with related works on key features in conflict detection and resolution modelling.

Work	Infra modelling	Speed modelling	Speed in sep. dist.	Rescheduling decisions	Solution method	Model objective	Application scenario
D'Ariano et al. (2007b)	Blocks	Iterative	No	RT,RO	B&B	Max secondary	20-km area
Lamorgese & Mannino (2015)	Blocks	Fixed	Fixed	RT,RO,RR	Macro/micro decomposition	Mean station delay cost	100-km lines
Luan et al. (2018)	Blocks	Multiple profiles	Yes	RT,RO,RR	Two-level	Total mean	50-km corridor station delay
Mazzarello & Ottaviani (2007)	Blocks	Iterative	No	RT,RO,RR	Two-level	Max secondary	44-km area
Pellegrini et al. (2015)	TDSs	Fixed	No	RT,RO,RR	Heuristic	Total final delay	Various control areas
Reynolds & Maher (2022)	TDSs	Two profiles	No	RT,RO,RR	B& P	Custom utility	Ext. station area
Törnquist & Persson (2007)	Blocks	Fixed	No	RT,RO	Solver	Total delay	400-km network
Xu et al. (2017)	Blocks	Multiple levels	Yes	RT,RO	Two-step	Total secondary delay	high-speed corridor like network
This work	TDSs	Two profiles	Yes	RT,RO,RR	Two-step solver	Total delay, max speed penalty	17-km junction, 68-km corridor

sep. dist.: train separation distance, TDSs: trackside train detection section, B&B: branch-and-bound, B&P: branch-and-price, RT: retiming, RO: reordering, RR: rerouting, Max: maximum.

Models

Three main model classes can be defined: alternative graph, disjunctive (or big-M) mixed integer linear programming (MILP) and time-indexed MILP models.

The main benefit of formulating the CDR problem as an alternative graph is the relation with general scheduling problems for which proven solution algorithms exist (D'Ariano et al., 2007a; Mascis & Pacciarelli, 2002; Mazzarello & Ottaviani, 2007). Alternative graph models typically rely on a microscopic representation of the infrastructure, assuming an infrastructure discretisation corresponding to the fixed-block signalling system. However, the specific formulation also limits the flexibility of the resulting models. For example, as an objective function only the minimisation of the maximum secondary delay can be considered. Also, the initial formulation does consider train occupation times in the separation modelling, rather than the full blocking times. This is however shown to be surmountable by explicitly adding the approach

time component (Corman et al., 2009). Other downsides are the serious difficulties in the inclusion of rerouting and speed beyond the fixed-speed assumption. Some efforts have been made concerning both (D'Ariano et al., 2008; Mazzarello & Ottaviani, 2007). The fixed-speed CDR model can be used in an iterative scheme with a speed profile generation model to find a valid speed profile for a computed rescheduling solution (D'Ariano et al., 2007b; Mazzarello & Ottaviani, 2007). The model of D'Ariano et al. (2007a) is considered in the application to moving-block signalling by Janssens (2022). Unsurprisingly, model restrictions in terms of objective, rerouting, speed, and train separation resonate throughout this early-stage moving-block research (Versluis et al., 2024).

In disjunctive MILP CDR models, by which we mean MILP models with disjunctive constraints representing 'or' conditions, railway operations are described by decision variables and mixed integer linear constraints, complemented with a linear objective function (Luan et al., 2018; Pellegrini et al., 2015; Törnquist & Persson, 2007; Xu et al., 2017). Though linearity of a model is generally desirable, there is also a downside to it. To obtain linear constraints, the approach relies on the big-M linearisation method (Bazaraa et al., 2008). The drawback of that method is that it is a weak linearisation, which makes the resulting model hard to solve to optimality (Reynolds et al., 2020). The formulation is flexible in terms of model objectives and rescheduling, including rerouting. It is also fairly flexible in terms of infrastructure, speed, and separation modelling, though some remarks are needed here (Versluis et al., 2024). First, some form of discretisation is required for the infrastructure. Second, fixed-speed assumptions based on one speed profile are most common. However, Luan et al. (2018) and Xu et al. (2017) considered speed levels in their applications to fixed-block and distance-to-go signalling, respectively. Immediately, train separation times become speed-dependent through the consideration of full blocking times. The latter is typical in disjunctive MILP models, though Törnquist & Persson (2007) instead assume fixed train separation times.

Known for their strong linearisation and good approximation for scheduling problems (Van den Akker et al., 2000), time-indexed MILP models are considered as an alternative to disjunctive MILP formulations for the CDR problem (Bettinelli et al., 2017; Lusby et al., 2013; Reynolds et al., 2020). The models feature resources that correspond to pairs of a time unit and a track part. Possible train routes can be visualised as source-sink paths in a time-space graph corresponding to the problem instance. This makes the time-indexed approach generally well-suited for the consideration of rerouting (Bettinelli et al., 2017; Lusby et al., 2013; Reynolds et al., 2020). These resources, however, seriously limit the model's application possibilities. The necessary discretisation of time and space restricts the flexibility in the modelling of running and separation times, and capacity, respectively. It is also the cause of the main drawback of the approach, i.e., the model size (Reynolds et al., 2020; Van den Akker et al., 2000). Lusby et al. (2013) and Reynolds & Maher (2022) model speed beyond fixed-speed assumptions. They consider alternative speed profiles corresponding either to running at a constant speed or to accelerate/decelerate around a stop.

From this overview, we conclude that DTG signalling is underrepresented within the variety of CDR models, especially in the context of mainline railways. Besides model formulation, important limitations of existing models are their dependence on infrastructure discretisation and the incorporation of speed in train separation modelling. For application of CDR models in the context of DTG signalling, it is necessary to increase the general modelling flexibility. In this chapter, we contribute to the challenge of making CDR models more applicable for DTG

signalling by addressing both infrastructure and speed modelling. Our goal is to model train separation based on full blocking times taking into account absolute braking distances, with a focus on mainline railways, to assess the operational relevance of DTG signalling in CDR.

Solution methods and practical applications

Given the presented overview of the main types of models, the next step is to look at the solution methods considered in literature. The solution approach is of great importance for the practical applicability of CDR models due to the real-time nature of the modelled problem. Hence, the focus here is on approaches that have an impact on computation times.

Traditionally, solving MILP models involves branch-and-bound algorithms, which are also implemented in commercial solvers such as CPLEX. Branch-and-bound is an exact method but can be enhanced by heuristic techniques to speed up the solution process. Examples applied to CDR are the truncated branch-and-bound algorithm proposed by D'Ariano et al. (2007a) or the tailored branch-and-price algorithm of Reynolds et al. (2020). Also, iterative (meta)heuristics are considered in literature to obtain rescheduling solutions within effective computation times (Zhang et al., 2024).

As alternative solution approaches to tackle the challenge of finding good solutions quickly, exact decomposition methods are considered, e.g., by Lamorgese & Mannino (2015) and Leutwiler & Corman (2022). Lamorgese & Mannino (2015) introduce an exact macro-/microscopic decomposition of the CDR problem, inspired by the Bender's decomposition approach. The resulting iterative process of solving the problem at a macroscopic level and searching for feasible train routes through stations, has been applied in pilot tests in practice. Leutwiler & Corman (2022) apply a logic-based Benders decomposition, in which the subproblem is the feasibility counterpart of the master optimisation problem. Applied on a case study from the Swiss Federal Railways (SBB), the presented approach is up to 40 times faster than the centralised approach combined with a commercial solver.

Several other decomposition methods are considered in literature, which can, e.g., be found in the review papers of Leutwiler & Corman (2023) and Marcelli & Pellegrini (2021). Here, we mention two works that are based on the earlier mentioned MILP formulations of Törnquist & Persson (2007) and Pellegrini et al. (2015), respectively. Lippes (2024) and Yi et al. (2023) both propose an iterative approach based on a geographical decomposition of the network and a coordination algorithm to combine the local rescheduling solutions into an overall feasible solution. Per iteration, Lippes (2024) obtains a (moving-block) rescheduling solution for one subarea, while Yi et al. (2023) obtains (fixed-block) rescheduling solutions for all subareas separately. Both non-exact methods lead to significant computation time reduction compared to the centralised counterpart in most scenarios.

On a different note, model predictive control (MPC) is proposed as alternative modelling approach for the CDR problem by, e.g., Caimi et al. (2012) and Pochet et al. (2016). Caimi et al. (2012) propose an MPC approach for rescheduling within complex central station areas, incorporating the retiming and rerouting of trains as well as partial speed profile coordination. With an SBB case study, it was demonstrated that this approach is viable for practical applications. Pochet et al. (2016) apply an MPC-based approach in the context of suburban railway

lines with mixed CBTC traffic, considering retiming and reordering measures. The approach is incorporated in a microscopic simulation tool of the French train operator SNCF.

In general, actual real-life implementation of CDR in traffic management systems are rare. Lamorgese et al. (2018) mention a few, of which here we report the ones with supporting documentation. Well-represented here is Bombardier Transportation (acquired by Alstom in 2021), with an optimisation-based support system (temporary) embedded in the traffic management system of some terminal stations of the Milano Undergound (Mannino & Mascis, 2009), ans on various main lines in Italy (Mannino, 2011) and Latvia. In the Milano system, an exact branch-and-bound was used as solution algorithm. In the latter two, the stricter business rules required a heuristic approach, with which computational speed was gained with minimal loss in the 'optimal' performance. As already shortly referred to, the exact decomposition method of Lamorgese & Mannino (2015) has been applied on a Norway railway line with the involvement of the local infrastructure manager and train operating companies.

Overall it is clear that for practical applications, the straightforward MILP solving methods are unlikely to be sufficient. We note that practical application is not a focus point in this chapter, but we acknowledge the importance of it and the related real-time performance. Specifically, when evaluating the model's performance and in following steps.

3.3 Distance-to-go signalling in conflict detection and resolution models

In this section, we present our approach for enhancing microscopic fixed-block CDR models to describe DTG signalling. The proposed enhancements focus on capturing the characteristics of train separation under DTG signalling. As observed, the dependence of the train separation on speed and infrastructure discretisation touches upon gaps in the CDR literature. For an accurate modelling of the continuous relation between speed and train separation under braking curve supervision, we need to relax the (fixed-block) discretisation of the infrastructure and the fixed-speed assumptions. This allows for the incorporation of (train- and) speed-dependent brake indication points. Along that line, we address the following three aspects: 1) switch and track areas, 2) speed profile options, and 3) blocking times for DTG signalling.

3.3.1 Switch and track areas

The discretisation of the track can be different around switches and on the (rest of the) track. To accommodate the resulting differences in DTG blocking times, we divide the infrastructure into switch areas and tracks, as commonly considered in the field. A switch area contains a collection of sections, e.g., TDSs or (virtual) blocks, such that it has more than two entry and/or exit points. Depending on the track layout, i.e., the relative closeness of switches, a switch area can contain one or more switches. In practice, two switch areas should lie more than the maximum train length apart. This allows all trains to come to a standstill between the switches without blocking traffic over either of them. Two switch areas are connected by one or more tracks. We distinguish two types of tracks: station tracks and open line stretches. If a track lies at a station, it is a station track. If not, the track is an open line stretch. Parallel tracks

connecting two switch areas are considered as separate station tracks and/or open line stretches. Figure 3.2 illustrates the division of the infrastructure into switch areas, station tracks and open line stretches.

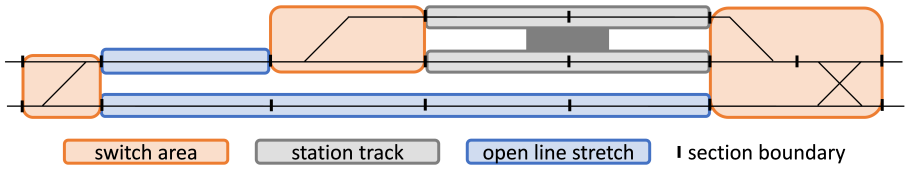


Figure 3.2: Infrastructure divided into switch areas, station tracks and open line stretches.

We propose a discretisation of the track beyond the fixed-block structure. The discretisation can correspond to physical blocks, TDSs or virtual blocks. We introduce the concept of locations to indicate the discretisation points corresponding to section entry points. In the rest of the chapter, when suitable, locations in switch areas are distinguished from locations on the tracks, by referring to them as *switch locations* and *track locations*, respectively. This distinction enables the modelling of possibly different blocking times in switch areas and on the track. For example, by reserving switch locations within a switch area together and track locations independently from one another.

3.3.2 Speed profile options

Consistent with the speed dependency of train separation under DTG signalling, the notion of speed modelling is extended to two speed profile options. That is, the option to run according to the *maximum speed profile* and the option to run according to a *scheduled speed profile*. In general, maximum speed profiles are characterised by maximum acceleration, a target cruising speed equal to the maximum speed and maximum deceleration, and scheduled speed profiles by a lower target cruising speed and, possibly, coasting phases. Figure 3.3 shows examples of a maximum and a scheduled speed profile between two stops, considering an intermediate track speed restriction.

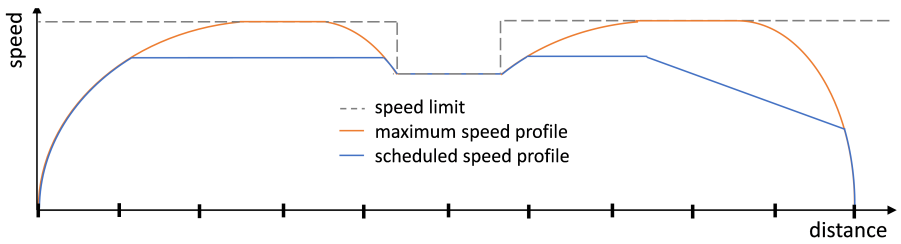


Figure 3.3: Maximum and scheduled speed profile between two stops with a speed restriction.

We propose to consider speed profiles only on the track, i.e., on open lines and station tracks. Within switch areas, speed is typically restricted and track occupation times should be kept to a minimum. Hence, we assume unique train running times in switch areas, independently of the speed profile assigned before or after. We consider the speed profile to not change on a track, but different profiles can be used on different tracks by a train. This may significantly reduce the size of the model compared to allowing changes at every location separately. For every

track, the first location is selected as *speed assignment location*. Only at these locations, speed profiles are assigned (one for each train along each route).

The introduction of speed profile options affects the train running and clearing times. The minimum running and clearing times corresponding to the maximum speed profile are taken as reference. For cases in which trains are operating according to the scheduled speed profile, the longer running and clearing times with respect to the minimum times can be predetermined and included as input parameters.

Typically, scheduled speed profiles are assigned. Trains should only run according to the maximum speed profile if the situation requires it. For example, maximum speed profiles can be followed if the train is delayed or when it can reduce the impact of a delayed train. To encourage the assignment of scheduled speed profiles, the minimisation of the number of maximum speed profile assignments can be included in the model objective.

We note that despite the introduction of a second speed profile option, the enhanced model will still be a fixed-speed model. Also in case of considering more than two speed profile options, this notion will remain to a certain extent. In this work, we have opted for two speed profile options as a compromise between the computational efficiency, which is negatively affected by more speed profiles options through the increasing model size, and realism, which is positively affected by more speed profiles that increases the model accuracy. This trade-off is based on the matter that CDR models are not actually representing railway operations itself, as opposed to, for example, simulation models.

3.3.3 Blocking times for distance-to-go signalling

With their direct relation to minimum train separation, blocking times are the core of microscopic CDR models. In the blocking time components, DTG principles are incorporated to describe the minimum train separation based on speed-dependent braking distances.

The minimum separation between a pair of trains at a specific location can be expressed in terms of their blocking time components. Related to the leading train, we need to take into account the train's clearing and release time of the section starting at the location. Concerning the following train, we need to consider its setup, reaction and approach time related to that location. Due to locations referring to fixed discrete points, we should also include the leading train's running time from one location to the next.

The train- and speed-dependency of the separation distance is clearly presented in the approach time, i.e., the time it takes to cross the approach distance. The DTG approach distance is defined as the braking distance supplemented with a safety margin. The inclusion of a safety margin accounts for the uncertainties in train position and speed due to, for example, measurements errors. To be able to describe this in the model, we introduce the concept of *reference (brake) location*. For track locations, reference locations are set to lie minimally the approach distance before the location's position on the track. For switch locations, reference locations are obtained similarly, but always considering the previous location in the switch area along a specific route.

In the following, we formalise the DTG approach distance based on its definition in terms of braking distance and safety margin. For a given combination of train, location and (approaching)

speed, the DTG approach distance can be determined as follows (Brünger & Dahlhaus, 2014):

$$\text{approach distance} = \frac{\text{speed}^2}{2\,(\text{braking rate} + \text{resistance deceleration})} + \text{safety margin}, \tag{3.1}$$

with the approach distance and safety margin in metres, the speed in metres per second, and the braking rate and resistance deceleration in metres per square second. For a correctly defined approach distance, we assume the braking rate to be constant. That is, for a certain train and speed, we assume one braking rate that best approximates the train's braking curve up to the given speed. The deceleration due to the resistance of the train on the track depends on the train, the speed of the train at the location and the track gradient at the location, and can be, e.g., expressed according to a quadratic equation of speed (Davis, 1926).

We note that our definition of braking distance differs from the ones that are based on the ETCS braking curve computations of the European Union Agency for Railways (2020). The main difference is that we consider the distance traversed by the train from the moment the brakes start to work, i.e., the braking point, while the European Union Agency for Railways (2020) considers the distance traversed by the train from the moment a brake indication, i.e., at the indication point, is issued by the system. The (travelling) time between the indication and braking points can be described as the human and system reaction times, which is modelled in the blocking time component 'reaction time'.

Given a location, train and speed, the approach distance formula provides the braking point before the location. As illustrated in Figure 3.4, this point may not coincide with a modelled discrete track location: the last track location before the braking point is defined to be the reference (brake) location. Hence, the reference location lies some lag distance earlier along the track than the braking point. Note that the lag distance is strictly shorter than the length of the track between the reference location and its succeeding location, i.e., the entry point of the next section along the train's route. This length is also referred to as the 'reference section length'. Moreover, reference locations and the corresponding lag distances depend on the speed profile assigned.

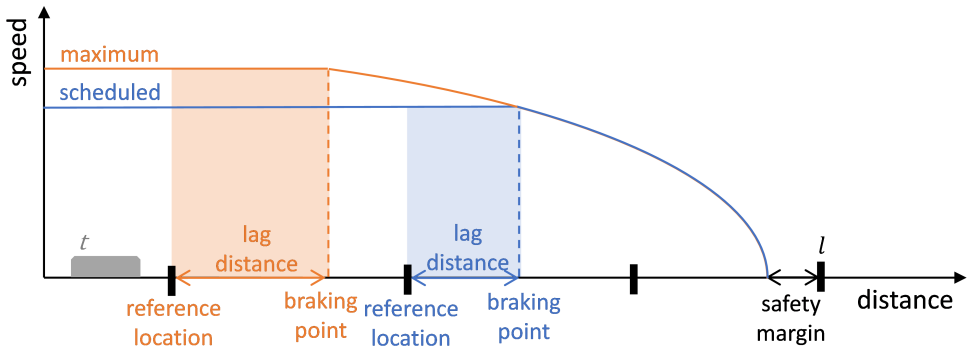
In addition to the definition of the reference brake location, we enable the modelling of continuous braking curve supervision. Indeed, the blocking of a location by an approaching train starts at the moment the train passes the braking point, minus the setup and reaction time, rather than the associated reference brake location. This implies that only if a train's MA reaches beyond a certain location, it can pass the braking point associated with that location. For the approximation of the passing times at braking points, we introduce the concept of *reservation lag*. For a specific speed profile, the reservation lag indicates the time interval by which the reservation can be postponed, with respect to the passing time of the corresponding reference location. Hence, it is the time it takes the train to traverse the earlier defined lag distance, which is approximated by the following formula:

$$\text{reservation lag} = \frac{\text{lag distance}}{\text{reference section length}}\ \text{running time over reference section},$$

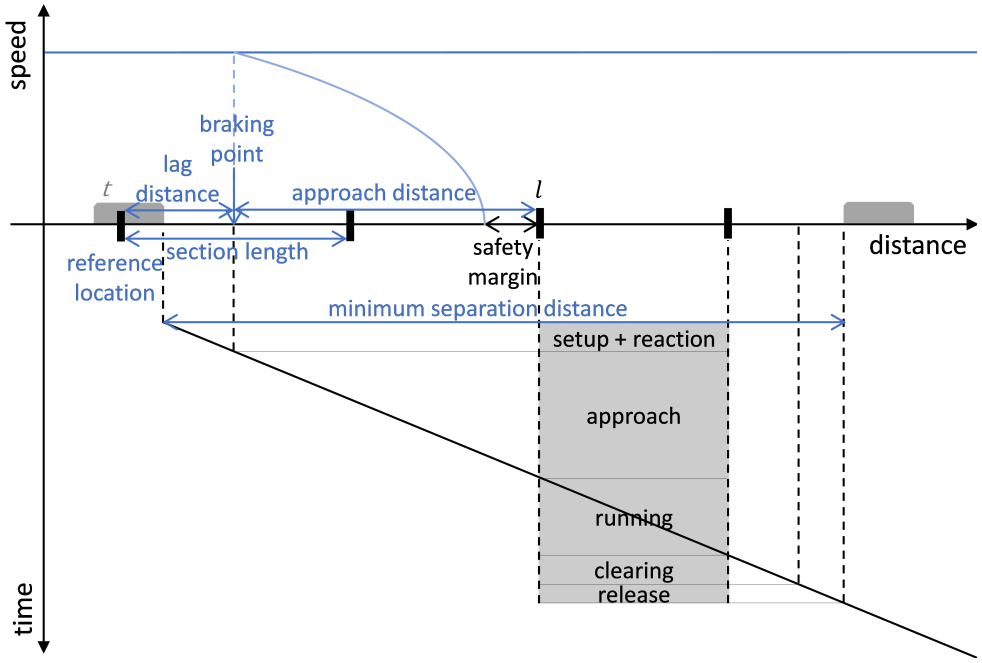
with the reservation lag and running time in seconds and the lag distance and section length in metres. Note that the values of all terms but the reference section length depend on the speed profile.

The modelling concept for fixed-block DTG signalling is illustrated in Figure 3.4. Figure 3.4a indicates the speed-dependent reference (brake) locations and associated lag distances for train t approaching location l according to its maximum or scheduled speed profile. It shows that the *maximum-speed reference location* lies before the *scheduled-speed reference location* since the higher the speed, the longer the braking distance. It also shows that the lag distances, or the reservation lags, for the maximum and scheduled speed profiles are not equal as they depend on the braking point in relation to the discretisation grid. In Figure 3.4b, the reference location and lag distance for the scheduled speed profile are related to the approach distance and the reference section length. Also, the blocking time components are related to the train separation. We note that the trains in the figure are separated by the minimum separation distance under worst-case conditions, i.e., the largest value that the minimum separation can take.

We note that this enhancement does not modify the blocking time theory but only allows to provide a more accurate computation of train blocking times to capture the DTG signalling feature of braking curve supervision.



(a) For maximum and scheduled speed profiles.



(b) Train blocking time components related to minimum train separation for scheduled speed profile.

Figure 3.4: Modelling concept including braking points, reference (brake) locations and (reservation) lag distances for train t approaching location l.

3.4 Enhancement of RECIFE-MILP to distance-to-go signalling

In this section, we apply the proposed enhancements to the state-of-the-art fixed-block RECIFE-MILP model (Pellegrini et al., 2015), obtaining a MILP formulation of a CDR model for DTG signalling. First, we shortly describe the RECIFE-MILP model from the literature (Section 3.4.1). Second, we provide the mathematical formulation resulting from enhancing RECIFE-MILP (Section 3.4.2).

3.4.1 The RECIFE-MILP model

RECIFE-MILP is a CDR model for conventional fixed-block signalling developed by Pellegrini et al. (2014, 2015). The model features a microscopic representation of the infrastructure, in terms of TDSs grouped into blocks. A train route is considered as a sequence of TDSs from its entry to its exit of the considered infrastructure.

Due to the implementation of the route-locking sectional-release principle, the blocking time of a TDS for a train includes the setup and reaction time (together: the formation time), the time to run with the train's head from the entry of the reference block a fixed number of blocks in advance to the entry of the TDS (the approach time), the time to run with the train's head from the entry to the exit of the TDS (the running time), the time to run with the train's length over the TDS exit (the clearing time), and the release time (Hansen & Pachl, 2014).

RECIFE-MILP is a fixed-speed model assuming minimum train running and clearing times, and dismissing speed dynamics in case of unplanned declarations or stops. Station dwell times and delays are included in the model as extended occupation times. The objective of RECIFE-MILP is to minimise the weighted cumulative total delay, i.e., the delays of all trains upon entering the infrastructure, upon arriving at scheduled stops and upon exiting the considered infrastructure (or upon reaching its final destination). To this end, the model allows for retiming, reordering and rerouting of trains.

RECIFE-MILP has been applied to various case studies and has been proven to perform well in real-time applications (Pellegrini et al., 2015). Also, the model has been evaluated in open- and closed-loop frameworks with a simulation environment (Pellegrini et al., 2016; Quaglietta et al., 2016), enabling a realistic assessment of the implications of the rescheduling decisions optimised under the fixed-speed assumption. For the mathematical formulation of RECIFE-MILP and further details, we refer to Pellegrini et al. (2014, 2015, 2019).

3.4.2 MILP formulation for distance-to-go conflict detection and resolution

Here, we present the new mathematical model of RECIFE-MILP for DTG signalling. The sets, parameters and variables, as well as the objective function and the constraints of the MILP formulation for DTG CDR are included. Note that not all details of the RECIFE-MILP model are given here. For that, we refer the reader to Pellegrini et al. (2014, 2015, 2019).

Sets, parameters and variables

The sets, parameters and variables of the model are listed in Table 3.2. The sets represent collections of elements that are used for the model notation. The four main sets are the set of trains T, the set of routes R, the set of locations L and the set of stations S. Subsets $R_t \subset R$ and $S_t \subset S$ are defined to represent routes and stations relevant to train $t \in T$.

For the set of locations, multiple subsets are defined. First, the subsets indicating the relevant locations per train $L_t \subset L$ with $t \in T$ and per route $L^r \subset L$ with $r \in R$. Next, the sets of *occupied locations* $OL_{t,r,l} \subset L^r$, containing the locations along route $r \in R_t$ for which it holds that if train $t \in T$ starts occupying it, the train has not yet cleared location $l \in L^r$. $L_{t,s} \subset L_t$ is the set of locations that can be used by train $t \in T$ to stop at station $s \in S_t$. $\hat{L}_{t,t',l} \subset L_t \cap L_{t'}$ is the set of locations that trains t, $t' \in T$ need to traverse in the same order, i.e., if train t precedes t' on $l \in L$, then t precedes t' on $l' \in \hat{L}_{t,t',l}$. Finally, the set of speed assignment locations $\mathrm{P} \subset L$ containing the track locations to which a speed profile is assigned for the whole track. Subset $\mathrm{P}^r \subset \mathrm{P}$ contains the speed assignment locations along route $r \in R$.

The parameters are provided as input to the model. First, we define the parameters related to the infrastructure. $\pi_{r,l}$ and $\sigma_{r,l}$ are defined to represent the preceding and succeeding location of location $l \in L^r$ along route $r \in R$, respectively. $\rho_{r,l} \in \mathrm{P}^r$ indicates the speed assignment location associated with location $l \in L^r$ along route $r \in R$. As mentioned in Section 3.3.2, it is the first location of the track to which l belongs. Additionally, s_l indicates whether location $l \in L$ is a switch location ($s_l = 1$), or a track location ($s_l = 0$). Also, dummy location $l_\infty \in L$ represents the dummy destination of all trains, added at the end of each route.

Next, we define the parameters related to the timetable. The initial entry time and the scheduled destination arrival time of train $t \in T$ are given by $init_t$ and $sched_t$, respectively. The minimum dwell time of train $t \in T$ at station $s \in S_t$ is given by $dw_{t,s}$. The arrival and departure time of train $t \in T$ at station $s \in S_t$ are given by $a_{t,s}$ and $d_{t,s}$, respectively.

Then, we define the parameters related to the blocking times. The formation time, which includes the setup and reaction time, and the release time of location $l \in L^r$ along route $r \in R$ are given by $for_{r,l}$ and $rel_{r,l}$, respectively. The running time of train $t \in T$ from location $l \in L_t$ to its succeeding location $\sigma_{r,l} \in L^r$ along route $r \in R_t$ is described by two parameters. $rt_{t,r,l}$ gives the minimum running time corresponding to running according to the maximum speed profile. $\Delta rt_{t,r,l}$ gives the additional running time in case the train follows a scheduled speed profile. Similarly, the clearing time of train $t \in T$ of location $l \in L^r$ along route $r \in R_t$ is given by the minimum clearing time $cl_{t,r,l}$ and the possible additional clearing time $\Delta ct_{t,r,l}$.

Another group of parameters is related to the approach time: the reference (brake) locations and the reservation lags. They are defined in line with their introduction in Section 3.3.3. Locations $ref^{\mathrm{s}}_{t,r,l}$, $ref^{\mathrm{m}}_{t,r,l} \in L^r$ represent the reference brake locations for location $l \in L^r$ along route $r \in R_t$ for train $t \in T$ approaching according to the scheduled or maximum speed profile, respectively. Reservation lag parameters $lag^{\mathrm{s}}_{t,r,l}$ and $lag^{\mathrm{m}}_{t,r,l}$ are defined to indicate the time by which the reservation of location $l \in L^r$ along route $r \in R_t$ for train $t \in T$ can be postponed after passing the corresponding reference brake location. We refer back to Figure 3.4a for the illustration of the reference brake location and reservation lags.

Table 3.2: Sets, parameters and variables in the MILP model formulation.

Sets	
T	set of trains
R	set of routes
$R_t \subset R$	set of routes available to train $t \in T$
L	set of locations
$L_t \subset L$	set of locations which can be used by train $t \in T$
$L^r \subset L$	set of locations along route $r \in R$
$OL_{t,r,l} \subset L^r$	set of locations along route $r \in R_t$ such that if train $t \in T$ starts occupying it, the train has not yet cleared location $l \in L^r$, $l \notin OL_{t,r,l}$
S	set of stations
$S_t \subset S$	set of stations where train $t \in T$ has a scheduled stop
$L_{t,s} \subset L_t$	set of locations that can be used by train $t \in T$ for stopping at station $s \in S_t$
$\hat{L}_{t,t',l} \subset L$	set of locations $l' \in L_t \cap L_{t'}$ which may be used by trains $t, t' \in T$ such that if t precedes t' on l, then t precedes t' on l'
$\mathrm{P} \subset L$	set of speed assignment locations
$\mathrm{P}^r \subset \mathrm{P}$	set of speed assignment locations along route $r \in R$
Parameters	
$\pi_{r,l}, \sigma_{r,l} \in L^r$	preceding/succeeding location of location $l \in L^r$ along route $r \in R$
$\rho_{r,l} \in \mathrm{P}^r$	speed assignment location associated with location $l \in L^r$ along route $r \in R$
$s_l \in \{0,1\}$	$= 1$ if location $l \in L$ lies in a switch area
$l_\infty \in L$	dummy location considered as destination for all trains
$init_t \in \mathbb{R}_+$	earliest time at which train $t \in T$ can be operated
$sched_t \in \mathbb{R}_+$	scheduled arrival time of train $t \in T$ at dummy destination location $l_\infty \in L$
$dw_{t,s} \in \mathbb{R}_+$	minimum dwell time for train $t \in T$ at station $s \in S_t$
$a_{t,s}, d_{t,s} \in \mathbb{R}_+$	scheduled arrival/departure time for train $t \in T$ at station $s \in S_t$
$for_{r,l} \in \mathbb{R}_+$	formation time, i.e., setup and reaction time, of location $l \in L^r$ along route $r \in R$
$rel_{r,l} \in \mathbb{R}_+$	release time of location $l \in L^r$ along route $r \in R$
$rt_{t,r,l} \in \mathbb{R}_+$	minimum running time for train $t \in T$ from location $l \in L_t$ to $\sigma_{r,l}$ along route $r \in R_t$
$\Delta rt_{t,r,l} \in \mathbb{R}_+$	additional running time for train $t \in T$ from location $l \in L_t$ to $\sigma_{r,l}$ along route $r \in R_t$ in case of scheduled speed profile
$ct_{t,r,l} \in \mathbb{R}_+$	minimum clearing time for train $t \in T$ of location $l \in L_t$ along route $r \in R_t$
$\Delta ct_{t,r,l} \in \mathbb{R}_+$	additional clearing time for train $t \in T$ of location $l \in L_t$ along route $r \in R_t$ in case of scheduled speed profile
$ref^{\mathrm{s}}_{t,r,l} \in L^r$	reference brake location for location $l \in L^r$ along route $r \in R_t$ for train $t \in T$ approaching according to scheduled speed profile
$ref^{\mathrm{m}}_{t,r,l} \in L^r$	reference brake location for location $l \in L^r$ along route $r \in R_t$ for train $t \in T$ approaching according to maximum speed profile
$lag^{\mathrm{s}}_{t,r,l} \in \mathbb{R}_+$	time by which blocking of location $l \in L^r$ by train $t \in T$ running according to scheduled speed profile along route $r \in R_t$ can be postponed after passing $ref^s_{t,r,l}$
$lag^{\mathrm{m}}_{t,r,l} \in \mathbb{R}_+$	time by which blocking of location $l \in L^r$ by train $t \in T$ running according to maximum speed profile along route $r \in R_t$ can be postponed after passing $ref^m_{t,r,l}$
$M \in \mathbb{R}_+$	a large constant
$w_t \in \mathbb{R}_+$	priority weight of train $t \in T$ for objective function
$w \in \mathbb{R}_+$	maximum speed profile penalty
Variables	
$y_{t,t',l} \in \{0,1\}$	$= 1$ if train $t \in T$ blocks location $l \in L_t \cap L_{t'}$ before train $t' \in T$
$x_{t,r} \in \{0,1\}$	$= 1$ if train $t \in T$ uses route $r \in R_t$
$v^{\mathrm{s}}_{t,r,l} \in \{0,1\}$	$= 1$ if train $t \in T$ passes speed assignment location $l \in \mathrm{P}^r$ along route $r \in R_t$ according to scheduled speed profile
$v^{\mathrm{m}}_{t,r,l} \in \{0,1\}$	$= 1$ if train $t \in T$ passes speed assignment location $l \in \mathrm{P}^r$ along route $r \in R_t$ according to maximum speed profile
$o_{t,r,l} \in \mathbb{R}_+$	occupation start time of train $t \in T$ on location $l \in L^r$ along route $r \in R_t$
$o^+_{t,r,l} \in \mathbb{R}_+$	extended occupation time of train $t \in T$ between locations $l \in L^r$ and $\sigma_{r,l} \in L^r$ along route $r \in R_t$
$b^{\mathrm{s}}_{t,l}, b^{\mathrm{e}}_{t,l} \in \mathbb{R}_+$	time at which train $t \in T$ starts/ends blocking location $l \in L_t$
$z_t \in \mathbb{R}_+$	delay suffered by train $t \in T$ when exiting the infrastructure and/or arriving at destination
$z_{t,s} \in \mathbb{R}_+$	delay suffered by train $t \in T$ when stopping at station $s \in S_t$

Finally, we define M to be a large constant and the weights w_t, $t \in T$ and w, for the objective function. w_t is a priority weight for train $t \in T$ and its (total) delay, and w is a penalty weight for the assignment of maximum speed profiles.

The model variables are either binary decision variables or continuous timing variables. The binary decision variables capture the scheduling and speed profile decisions. The passing order of two trains $t, t' \in T$ at common location $l \in L_t \cap L_{t'}$ is determined by variable $y_{t,t',l}$. The order variables are only defined for one representative location per $\hat{L}_{t,t',l}$ set. Recall that within $\hat{L}_{t,t',l}$, train orders do not change. The route assignment of train $t \in T$ is captured by variable $x_{t,r}$, indicating for route $r \in R_t$ whether or not the route is used by train $t \in T$. The assignment of speed profiles is described by two different sets of variables. $v^{\mathrm{s}}_{t,r,l}$ indicates whether or not train $t \in T$ runs according to the scheduled speed profile over speed assignment location $l \in \mathrm{P}^r$ along route $r \in R_t$, while $v^{\mathrm{m}}_{t,r,l}$ indicates whether or not train $t \in T$ runs according to the maximum speed profile over speed assignment location $l \in \mathrm{P}^r$ along route $r \in R_t$.

The timing variables include the decision variables that indicate the physical occupation start time: $o_{t,r,l}$ with $t \in T$, $r \in R_t$ and $l \in L^r$. Additionally, we have timing variables that depend on the occupation start time and the binary decision variables. These auxiliary variables represent the extended physical occupation times due to dwell time, delay and/or scheduling decisions ($o^{+}_{t,r,l}$ with $t \in T$, $r \in R_t$ and $l \in L^r$), the blocking start and end times ($b^{\mathrm{s}}_{t,l}$ and $b^{\mathrm{e}}_{t,l}$ with $t \in T$ and $l \in L_t$), the final delays (z_t with $t \in T$), and the delays at scheduled stops ($z_{t,s}$ with $t \in T$ and $s \in S_t$).

Objective function

In line with the general objective of rescheduling in case of disturbances (Cacchiani et al., 2014), the main objective of the model is to minimise total train delay. As secondary objective, we want to enforce the assignment of scheduled speed profiles where possible. This results in the following objective function:

$$\text{minimise} \sum_{t \in T} \Big(w_t \big(z_t + \sum_{s \in S_t} z_{t,s} \big) + w \sum_{\substack{r \in R_t: \\ l \in \mathrm{P}^r}} v^{\mathrm{m}}_{t,r,l} \Big). \tag{3.2}$$

The first term includes the weighted cumulative delay, i.e., the weighted sum of train delays upon arriving at scheduled stops ($z_{t,s}$ with $t \in T$ and $s \in S_t$) or upon exiting the infrastructure, either by leaving the area or by reaching its terminus (z_t with $t \in T$). The weights w_t with $t \in T$ can, for example, be interpreted as priority factors.

The second term counts the number of maximum speed profiles assigned. With maximum speed profiles only assigned when it reduces total delay by more than w seconds, it can be ensured that typically more energy-efficient scheduled speed profiles are assigned by default. To align with the secondary nature of this objective, weight w should be at most 1. From here on, this second term is also referred to as *speed penalty*.

Constraints

The constraints describing the DTG version of RECIFE-MILP are given by Equations (3.3) to (3.18). Compared to the original MILP, Constraints (3.6), (3.13), (3.14) and (3.15) are new, Constraints (3.7), (3.10), (3.12) and (3.16) are changed, and Constraints (3.3), (3.4), (3.5), (3.8), (3.9), (3.11), (3.17) and (3.18) are the same. The presented constraints are explained in the following. For the changed constraints, we add a short description of how they differ from the original version.

$$o_{t,r,l} \geq init_t\, x_{t,r} \qquad \forall t \in T,\ r \in R_t,\ l \in L^r, \quad (3.3)$$

$$o_{t,r,l} \leq M\, x_{t,r} \qquad \forall t \in T,\ r \in R_t,\ l \in L^r, \quad (3.4)$$

$$\sum_{r \in R_t} x_{t,r} = 1 \qquad \forall t \in T, \quad (3.5)$$

$$v^{\mathrm{m}}_{t,r,l} + v^{\mathrm{s}}_{t,r,l} = x_{t,r} \qquad \forall t \in T,\ r \in R_t,\ l \in \mathrm{P}^r, \quad (3.6)$$

$$o_{t,r,\sigma_{r,l}} = o_{t,r,l} + o^{+}_{t,r,l} + rt_{t,r,l}\, x_{t,r} + \Delta rt_{t,r,l} v^{\mathrm{s}}_{t,r,\rho_{r,l}} \qquad \forall t \in T,\ r \in R_t,\ l \in L^r, \quad (3.7)$$

$$o^{+}_{t,r,l} \geq \sum_{\substack{s \in S_t: \\ l \in S_{t,s} \cap L^r}} dw_{t,s}\, x_{t,r} \qquad \forall t \in T,\ r \in R_t,\ l \in \bigcup_{s \in S_t} L_{t,s}, \quad (3.8)$$

$$o_{t,r,\sigma_{r,l}} \geq \sum_{\substack{s \in S_t: \\ l \in L_{t,s} \cap L^r}} d_{t,s}\, x_{t,r} \qquad \forall t \in T,\ r \in R_t,\ l \in \bigcup_{s \in S_t} L_{t,s}, \quad (3.9)$$

$$z_{t,s} \geq \sum_{r \in R_t} \sum_{l \in L^r \cap L_{t,s}} \left(o_{t,r,l} + rt_{r,t,l} x_{t,r} + \Delta rt_{t,r,l} v^{\mathrm{s}}_{t,r,\rho_{r,l}}\right) - a_{t,s} \qquad \forall t \in T, s \in S_t, \quad (3.10)$$

$$z_t \geq \sum_{r \in R_t} o_{t,r,l_\infty} - sched_t \qquad \forall t \in T, \quad (3.11)$$

$$b^{\mathrm{s}}_{t,l} \leq \sum_{\substack{r \in R_t: \\ l \in L^r}} \left(o_{t,r,ref^{\mathrm{s}}_{t,r,l}} + \left(lag^{\mathrm{s}}_{t,r,l} - for_{r,l}\right) x_{t,r} \right)$$

$$\forall t \in T,\ l \in L_t : s_l = 0 \vee \nexists r \in R_t : s_{\pi_{r,l}} = 1, \quad (3.12)$$

$$b^{\mathrm{s}}_{t,l} \leq \sum_{\substack{r \in R_t: \\ l \in L^r}} \left(o_{t,r,ref^{\mathrm{m}}_{t,r,l}} + \left(lag^{\mathrm{m}}_{t,r,l} - for_{r,l}\right) x_{t,r} + M\, v^{\mathrm{s}}_{t,r,\rho_{r,ref^{s}_{t,r,l}}} \right)$$

$$\forall t \in T,\ l \in L_t : s_l = 0 \vee \nexists r \in R_t : s_{\pi_{r,l}} = 1, \quad (3.13)$$

$$b^{\mathrm{s}}_{t,l} \leq \sum_{\substack{r \in R_t: \\ l \in L^r}} \left(o_{t,r,ref^{\mathrm{s}}_{t,r,\pi_{r,l}}} + \left(lag^{\mathrm{s}}_{t,r,\pi_{r,l}} - for_{r,\pi_{r,l}}\right) x_{t,r} \right.$$

$$\forall t \in T,\ l \in L_t : s_l = 1 \wedge \exists r \in R_t : s_{\pi_{r,l}} = 1, \quad (3.14)$$

$$b^{\mathrm{s}}_{t,l} \leq \sum_{\substack{r \in R_t: \\ l \in L^r}} \left(o_{t,r,ref^{\mathrm{m}}_{t,r,\pi_{r,l}}} + \left(lag^{\mathrm{m}}_{t,r,\pi_{r,l}} - for_{r,\pi_{r,l}}\right) x_{t,r} + M\, v^{\mathrm{s}}_{t,r,\rho_{r,ref^{s}_{t,r,\pi_{r,l}}}} \right)$$

$$\forall t \in T,\ l \in L_t : s_l = 1 \wedge \exists r \in R_t : s_{\pi_{r,l}} = 1, \quad (3.15)$$

$$b^{\mathrm{e}}_{t,l} = \sum_{\substack{r \in R_t: \\ l \in L^r}} \left(o_{t,r,\sigma_{r,l}} + \left(ct_{t,r,l} + rel_{r,l}\right) x_{t,r} + \Delta ct_{t,r,l} v^{\mathrm{s}}_{t,r,\rho_{r,l}} + \sum_{\substack{l' \in L^r: \\ l' \in OL_{t,r,l}}} o^{+}_{t,r,l'} \right)$$

$$\forall t \in T,\ l \in L_t, \quad (3.16)$$

$$b^{\mathrm{e}}_{t,l} - M\left(1 - y_{t,t',\hat{l}}\right) \leq b^{\mathrm{s}}_{t',l} \qquad \forall t,t' \in T, \text{index } t < \text{index } t',\ l,\hat{l} \in L_t \cap L_{t'} : l \in \hat{L}_{t,t',\hat{l}}, \quad (3.17)$$

$$b^{\mathrm{e}}_{t',l} - M\, y_{t,t',\hat{l}} \leq b^{\mathrm{s}}_{t,l} \qquad \forall t,t' \in T, \text{index } t < \text{index } t',\ l,\hat{l} \in L_t \cap L_{t'} : l \in \hat{L}_{t,t',\hat{l}}. \quad (3.18)$$

Constraints (3.3) force train t to start operating no earlier than its initial entry time $init_t$ on its assigned route, while Constraints (3.4) set the occupation start time of all locations along the train's unassigned routes to 0. Constraints (3.5) ensure that a single route is assigned to each train.

Constraints (3.6) ensure that either a maximum or a scheduled speed profile is assigned to speed assignment locations along the route of a train. No speed profile is assigned to locations along routes that are not assigned to the train ($v^{\mathrm{m}}_{t,r,l} = v^{\mathrm{s}}_{t,r,l} = 0$).

Constraints (3.7) describe the difference in occupation start times between succeeding locations in terms of extended occupation time and running time. By Constraints (3.8), the extended occupation time $o^{+}_{t,r,l}$ includes the dwell time at scheduled stops. Additionally, it includes the difference in running time in case the assigned speed profile cannot be followed due to a delayed train in front. The running time of a train from a location to the succeeding one depends on the assigned speed profile. If train t runs over location l along route r according to the maximum speed profile, i.e., $v^{\mathrm{m}}_{t,r,\rho_{r,l}} = 1$ and hence $v^{\mathrm{s}}_{t,r,\rho_{r,l}} = 0$, then only the minimum running time $rt_{t,r,l}$ is considered. If train t runs over location l along route r according to the scheduled speed profile, i.e., $v^{\mathrm{s}}_{t,r,\rho_{r,l}} = 1$, then the additional running time $\Delta rt_{t,r,l}$ is also included. This additional running time term is the DTG addition to the original constraints.

Constraints (3.8) ensure that the station dwell times are included in the extended occupation time, while Constraints (3.9) ensure that train t does not leave its stopping location $l \in L_{t,s}$ before its scheduled departure time from station s.

Constraints (3.10) and (3.11) quantify non-negative delay at each station where train t has a scheduled stop ($z_{t,s}$) and at its exit from the infrastructure and/or when reaching its final destination (z_t). Note that t is assumed to stop at the end of the TDS where the stop occurs. In Constraints (3.10), the occupation start time of train t on the stopping location along route r at station s is compared with its scheduled arrival time at station s ($a_{t,s}$). For DTG, the speed-dependency of the running time over the stopping TDS is included. In Constraints (3.11), the occupation start time of train t on the dummy destination location is compared with the scheduled exit time ($sched_t$).

Constraints (3.12) to (3.15) set the blocking start times. Constraints (3.12) and (3.13) describe the speed-dependent blocking start times of track locations, i.e., location l such that $s_l = 0$, and of locations that can be the first of a switch area for a specific train, i.e., location l such that $\nexists r \in R_t : s_{\pi_{r,l}} = 1$.

Constraints (3.12) ensure that the blocking of track location l by train t starts at the latest the formation time before the train passes the braking point corresponding to the scheduled speed profile, that is the moment the train starts occupying the scheduled-speed reference location

$ref^{\mathrm{s}}_{t,r,l}$ along the assigned route, postponed with the scheduled-speed reservation lag $lag^{\mathrm{s}}_{t,r,l}$. Indeed, the blocking of a track location starts earlier when a train is approaching according to the maximum speed profile than to the scheduled speed profile because of the longer braking curve (see Figure 3.4a). These constraints differ from the fixed-block version in two main ways. First, it has become an inequality rather than an equality constraint due to the introduction of a speed profile alternative. Second, the right term is redefined to correspond to the braking point which can lie anywhere on the track rather than the entry of the train- and speed-independent reference block in the fixed-block model.

Constraints (3.13) ensure that in case the approaching train is running according to the maximum speed profile, the blocking starts earlier. Namely, at the moment the train passes the maximum-speed braking point ($o_{t,r,ref^{\mathrm{m}}_{t,r,l}} + lag^{\mathrm{m}}_{t,r,l}$) along the assigned route, minus the formation time. However, the blocking starti time in case of a train approaching according to the scheduled speed profile must not be restricted. For that purpose, a big-M term is added. The value of M should include the running time from the maximum-speed braking point to the scheduled-speed brake start point and the possible longer stay at the locations in between.

Constraints (3.14) and (3.15) deal with the locations in switch areas that are not the first switch location for a specific train, i.e., with l such that $s_l = 1$ while $\exists r \in R_t : s_{\pi_{r,l}} = 1$. The blocking of such a switch location starts as soon as the preceding location would start being blocked (if not also the exit of a switch section). Hence, Constraints (3.14) and (3.15) are the same as Constraints (3.12) and (3.13), except that they refer to (the braking point of) the preceding location.

Constraints (3.16) set the blocking end times. The blocking of a location lasts until the train has fully passed the succeeding location along its route, plus the release time. The clearing time is included in the blocking time. In case of a maximum speed profile, only the minimum clearing time is considered, while in case of a scheduled speed profile, an additional clearing time component is included. This speed-dependent term in the clearing time is the DTG alteration to the original constraint. Additionally, if the train is long enough to keep occupying a location when its head is at the following locations (included in set $OL_{t,r,l}$), also the extended occupation times of the train for these locations has to be accounted for.

Finally, disjunctive Constraints (3.17) and (3.18) ensure that the location blocking times of two trains do not overlap, depending on train orderings. The passing order of a pair of trains is defined per set of locations that the two trains need to pass in the same order.

3.5 Results

In this section, we present the results of the assessment of the fixed-block DTG CDR model obtained by enhancing RECIFE-MILP. The assessment is carried out through model experiments in two distinct case studies, which are introduced in Section 3.5.1. To evaluate the applicability to DTG and the operational performance of the model, we conduct a comparative analysis between the enhanced DTG model and the original model for conventional fixed-block signalling. The model solutions are compared in terms of the objective function and the underlying rescheduling decisions, and the effects of the model enhancements, specifically the shorter train separation due to the introduction of train- and speed-dependent blocking times

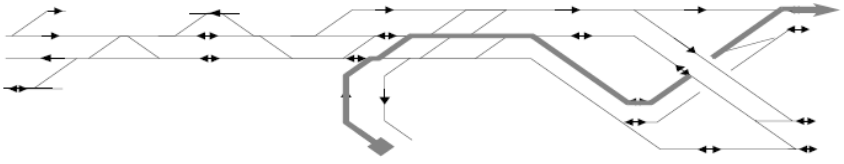
and brake indication points, are discussed. The results of the comparative analysis under the setup presented in Section 3.5.2 are discussed in Section 3.5.3. To further illustrate the model for fixed-block DTG operations, we zoom in on a specific case study instance in Section 3.5.4.

We run the experiments on an Intel(R) Xeon(R) CPU Gold 6226R CPU @ 2.90GHz, 16 cores, 256 GB RAM. The implementation uses IBM ILOG CPLEX Concert Technology for C++, version 20.1. The optimisation models are run to find the rescheduling decisions that give the minimum objective value. The computation time limit is set to 3600 seconds to seek optimality; we obtain a maximum optimality gap of 0.68%. With this, the time limit does not correspond to a real-time application, but it allows for a thorough assessment of the enhanced model: we aim to understand whether the model enhancements indeed result in a better description of DTG operations and, if so, whether they lead to different rescheduling decisions to be optimal. Given the foreseen real-time implementation, a discussion of the real-time applicability of the model is added in Section 3.5.5.

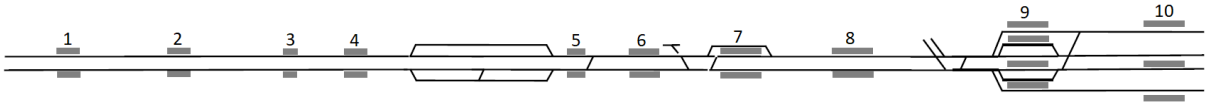
3.5.1 Case studies: Gonesse junction and Rosny-StEtienne corridor

The model experiments are performed in two case studies representing traffic control areas in France: the Gonesse junction and the Rosny-StEtienne corridor. The Gonesse area is a 17-kilometre long complex junction north of Paris with dense mixed traffic. Figure 3.5a provides a schematic representation of the Gonesse junction. The junction includes 89 TDSs with lengths ranging from 35 to 2424 metres, with a mean of 560 metres. The TDSs are grouped into 79 blocks and 37 routes. It has no platforms. The timetable of a weekday includes 336 trains, of which 116 high-speed, 129 conventional, and 91 freight trains, with 5 to 13 route alternatives per train.

The Rosny-StEtienne area is a 68-kilometre long corridor of the Paris-Le Havre line with mixed traffic. Figure 3.5b provides a schematic representation of the Rosny-StEtienne corridor. The corridor includes 239 TDSs with lengths ranging from 100 to 2217 metres, with a mean of 740 metres. The TDSs are divided over 152 blocks and 169 routes. It has 10 stations with a total of 39 platform tracks. Its daily timetable features 215 trains, of which 2 high-speed, 122 conventional, 56 freight and 35 empty (including work and test) trains, with 1 to 24 route alternatives each. We note that the Rosny-StEtienne corridor is not a high-speed line, so high-speed trains use the maximum speed of conventional trains.



(a) Gonesse junction with a train route used in Section 3.5.4.



(b) Rosny-StEtienne corridor.

Figure 3.5: Schematic track layouts of the traffic control areas consider in the case studies.

The two areas are equipped with conventional fixed-block signalling systems featuring more than three aspects. We note that the Gonesse junction features signals with different numbers of aspects, and that the blocks in the Rosny-StEtienne corridor are shorter than in the Gonesse area.

For both traffic control areas, 100 delay scenarios are generated based on a one-day timetable, where entry delays between 5 and 15 minutes are imposed on 20% of the trains. Both the selection of the affected trains and the magnitude of their entry delays are determined randomly, using uniform probability distributions. This introduces stochasticity into the scenario generation process and ensures an unbiased variety of delay scenarios, consistent with the approaches of Lusby et al. (2013) and Pellegrini et al. (2014). These delay scenarios are considered in a one-hour period, i.e., the peak hour from 18:00 to 19:00. For the selection of trains entering and/or starting in the traffic control area within the hour, retiming, reordering and rerouting options are taken into consideration.

The case studies are applied in the two versions of RECIFE-MILP: in a DTG/ETCS L2 variant of the enhanced model as well as in the original model for conventional fixed-block signalling. As (only) conventional fixed-block signalling is implemented in the considered areas, we project the DTG principles of ETCS L2 on the case studies. While the same MILP formulation is applied to the two case studies for each of the model versions, the differing characteristics, such as track layout and traffic pattern, lead to distinct models in terms of, e.g., size and outcome. For the two model versions, Table 3.3 presents insights on the model size, including the number of binary variables, continuous variables and constraints for both the junction and the corridor area. Based on the presented numbers, two points are worth noting.

Table 3.3: Number (#) of binary and continuous variables, and constraints in ETCS L2 and original, i.e., conventional fixed-block, model versions.

	ETCS L2	**Original**
Junction		
# binary variables	6186	3784
# continuous variables	28844	28844
# constraints	81407	78476
Corridor		
# binary variables	4649	665
# continuous variables	68902	68902
# constraints	113214	109887

First, the ETCS L2 model is larger than the original. The addition of speed profile options into the model leads to a significant increase in the number of binary variables and a slight increase in the number of constraints. The number of continuous variables stays the same as the number of timing variables is independent of speed modelling.

Second, the two case studies differ substantially in terms of the number of the three model components. Compared to the corridor case study, the junction case study features more binary variables, but significantly less continuous variables and also slightly less constraints. The latter two are in line with the 89 versus 239 TDSs. The binary variables are the route, order and speed variables. Route variables are balanced out with more routes, but less trains. With the difference in number of binary variables being the most substantial in the original model, it is not so much

related to the speed variables, but rather to the order variables. In areas similar to the considered corridor area, the number of order variables can be seriously reduced due to the definition of $\hat{L}_{t,t',l}$, i.e., the set of locations which may be used by trains t, t' such that if t precedes t' on l, then t precedes t' on l'.

3.5.2 Model parameter setup

For the initialisation of the model, we rely on input data from the original fixed-block version of RECIFE-MILP. We make the following assumptions regarding the three enhancements.

First, we discuss the assumptions concerning the track. For ETCS L2, we consider a discretisation of the track based on TDSs, refraining from the conventional fixed blocks. In line with that, reservation and release of the track is modelled at TDS level, and locations refer to the entries of the TDSs. The entry point of a block/TDS containing a switch, is a switch location. Otherwise, it is a track location.

Switch locations lie in the same switch area if they are not separated by a track location. Hence, a switch area consists of a collection of consecutive switch locations, and two switch areas are separated by at least one block/TDS, which can be assumed to be longer than a (passenger) train with a typical length of well over 200 metres.

Second, we discuss the assumptions related to the speed profile options. The maximum speed of a train at a location is the minimum between the maximum track speed at the location and the maximum train speed. The scheduled speed equals the maximum speed in switch areas, and is assumed to be 80% of the maximum speed on the track. Correspondingly, the additional running and clearing times of the scheduled speed profile with respect to the maximum one are 0% in switch areas and at most 25% on the track. We note that this assumption leads to a discontinuity in train speed profiles at the borders of switch areas, which we consider a consequence of the model being a fixed-speed model.

All weight factors in the objective function are set to be 1. With $w_t = 1$ for all trains $t \in T$, no trains are prioritised over other trains. With $w = 1$, it is counted how often a maximum speed profile is assigned. A positive weight ensures that a maximum speed profile is only assigned if it results in some delay recovery. If the weight is higher than 1, the assignment of maximum speed profiles is actively penalised. Then, minimum running times will be less often considered. Such a higher weight would give the ETCS L2 model a disadvantage compared to the conventional fixed-block model and hence, result in an unfair assessment.

Third, we discuss the assumptions related to the DTG blocking times. The formation and release times of a location equal those of the block they belong to in the conventional fixed-block model. These values are mostly based on the presence of trackside train detection, which is still the case in ETCS L2. For the computation of the brake locations (Equation (3.1)), we assume the train braking rates to be constant and the deceleration due to resistance to be negligible. The safety margin is 50 metres (PERFORMINGRAIL, 2022b). We note that these assumptions are made for ease of computation. In practice, they can easily be changed with more accurate values.

Recall that in the original fixed-block model, we use the minimum running and clearing times computed considering maximum speed profiles. By doing so, we are being conservative.

If a bias on delay assessment is introduced, it will be in favour of the original fixed-block model rather than of the enhanced DTG model that we aim to verify.

3.5.3 Comparative analysis

We present the results of the comparative analysis in two steps. First, we focus on the comparison of the ETCS L2 with the conventional fixed-block results in terms of the model objective components. Second, we discuss the underlying rescheduling solutions in more detail. In both steps, we consider the optimisation of the ETCS L2 and the original model for conventional fixed-block signalling as well as their cross-evaluations 'ETCS L2 in original' and 'original in ETCS L2'. By evaluating optimised solutions from one model version, e.g., the original model, in the other model version, e.g., the enhanced model, we analyse how the rescheduling decisions obtained in the context of one signalling system, e.g., conventional fixed block, affect the train delays when implemented under different signalling constraints, e.g., ETCS L2.

The optimisation results of the ETCS L2 and original models in terms of the model objective components are given in the first and last column of Table 3.4, respectively. These are the mean total delay and the mean delay recovery over all scenarios. Note that speed penalties are only included in the objective of the ETCS L2 model. The mean total delays obtained by the ETCS L2 optimisation are 7015 and 5530 seconds for the junction and corridor, respectively. The corresponding delay recovery percentages are 10.89% and 55.33%. With the original optimisation, mean total delays of 7325 and 5576 seconds are obtained for the respective case studies, with corresponding delay recovery percentage of 8.57% and 50.73%. Comparing the results, we observe that the ETCS L2 model significantly improves upon the original model regarding delay. The respective improvements for the two case studies are 4.23% and 0.82% in terms of total delay and 27.07% and 5.13% in terms of delay recovery.

Table 3.4: Mean total delay, delay recovery and number of maximum speed profile assignments (speed penalty) of optimisation and cross-evaluation of the ETCS L2 and original, i.e., conventional fixed-block, models.

	ETCS L2	**Original in ETCS L2**	**ETCS L2 in original**	**Original**
Junction				
Total delay (s)	7015	7062	7599	7325
Delay recovery (%)	10.89	10.39	6.96	8.57
Speed penalty	46	58	–	–
Corridor				
Total delay (s)	5530	5520	5590	5576
Delay recovery (%)	53.33	53.58	53.13	50.73
Speed penalty	10	29	–	–

Table 3.4 also reports the results of the cross-evaluations of the model solutions. First, the evaluation of the ETCS L2 solution in the original model: the scheduling decisions made in the ETCS L2 solutions are applied considering conventional fixed-block signalling. This evaluation is an additional check on the distinctness of the optimised rescheduling decisions of the models. If the ETCS L2 rescheduling decisions lead to the same objective value as

the rescheduling decisions in the original model, then the ETCS L2 solution is an alternative optimal conventional fixed-block solution. Similarly, the evaluation of the original rescheduling decisions in the ETCS L2 model can indicate that the original solution is an alternative optimal solution for the ETCS L2 model. In the case of either of the two situations, we dismiss the solution differences in the further analysis of the scenario.

From the evaluation of the ETCS L2 solution in the original model, it is clear that the ETCS L2 and original solutions contain different rescheduling decisions. With mean total delays of 7599 and 5590 seconds for the different case studies, the ETCS L2 solution performs 8.33% worse in the junction and 1.08% worse in the corridor when applied in fixed block. These differences originate from the different minimum train separation distances. Compared to the original optimisation, the ETCS L2 in original total delays are 3.74% worse in the junction but only 0.25% worse in the corridor. This underlines the occurrence of different optimal rescheduling solutions due to these distinct minimum train separations.

The evaluation of the original solution in the ETCS L2 model provides insight into the extent to which the improvement of ETCS L2 over conventional fixed-block originates from the ETCS L2 blocking times or from the possibly different rescheduling decisions in the respective solutions. The mean total delays of the cross-evaluation of conventional fixed-block in ETCS L2 are 7062 seconds for the junction case study and 5520 seconds for the corridor case study. With this, evaluating the original solution in the ETCS L2 model already improves the total delay with 3.59% and 1.00% for the junction and corridor case studies, respectively. So, most of the ETCS L2 over conventional fixed-block improvement originates from the ETCS L2 blocking times. Notwithstanding the slight improvement of the ETCS L2 solution over the original solution that is due to the different rescheduling decisions (ETCS L2 over original in ETCS L2): from 3.59% to 3.74% for the junction and 1.00% to 1.72% for the corridor. We also note the difference in speed penalty between ETCS L2 and original in ETCS L2. In the ETCS L2 solutions, trains run less often according to maximum speed profiles. Therefore, these solutions will also be beneficial in terms of energy consumption.

To better understand the role of the rescheduling solutions, we take a closer look at the case study instances. Specifically, we are interested in the instances for which the ETCS L2 model finds rescheduling decisions that are different from those in the (conventional) fixed-block solution and that outperform the latter in terms of objective value. As these instances benefit from the ETCS L2 optimisation, they are referred to as 'improved instances'. The percentage of improved instances per case study is reported in Table 3.5, together with the number of different ordering and routing decisions of ETCS L2 relative to conventional fixed block. In the junction case study, 98% of the instances are improved by ETCS L2 relative to conventional fixed block. For the corridor case study, this is the case for 55% of the instances.

Considering all 100 instances for the junction case study, the number of different ordering and routing decisions ranges from 0 to 4 with a mean of 0.31, and from 0 to 13 with a mean of 7.40, respectively. Over the 100 instances in the corridor case study, the number of different ordering and routing decisions range from 0 to 2 with a mean of 0.11 and from 0 to 11 with a mean of 2.87, respectively. So, the mean numbers of different routing decisions are a factor of 24 and 26 higher than the mean numbers of different ordering decisions in the junction and corridor case studies, respectively. On its own, one different rerouting typically offers a minimum improvement. However, some routes seem to be preferable for some trains. Moreover, different routes are becoming attractive due to the shorter blocking times in general.

Table 3.5: Case study results for ETCS L2 relative to conventional fixed block.

Case study	% improved instances	# Δ ordering decisions		# Δ routing decisions	
		max	mean	max	mean
Junction	98	4	0.31	13	7.40
Corridor	55	2	0.11	11	2.87

Overall, the results of the ETCS L2 model are not striking when comparing the rescheduling solutions (in the ETCS L2 model) with mean relative reductions in objective values of 0.85% and 0.17%. However, for specific instances, the model can have a remarkable impact: up to 7.09% reduction in objective value. For all instances, the ETCS L2 optimisation proposes solutions that are at least as good as the conventional fixed-block solution when evaluated in the ETCS L2 model.

From the results, we conclude that (the modelling of) DTG has more impact on CDR in the junction than in the corridor. The percentage of instances improved by ETCS L2 and the relative reduction in objective value is higher. An important factor in this is the traffic density, which is about a factor 1.5 higher in the junction than in the corridor: 336 versus 215 trains on a daily basis and 29 versus 19 trains in the considered hour. Moreover, the difference in track layout affects the attractiveness of alternative routes. The number of alternative routes is higher in the corridor, while the number of different routing decisions is higher in the junction. The corridor features double-track lines for traffic running in two directions, encouraging the separation of the two flows. The junction features overlapping lines in various directions, making flow separation practically impossible.

3.5.4 Detailed analysis of an improved instance

In this section, we further illustrate the model for DTG operations by zooming in on a specific case study instance. We aim to analyse in detail the 'what and why', e.g., related to the DTG modelling principles, of the differences in the rescheduling solutions provided by the model for the different signalling systems, i.e., ETCS L2 and the conventional fixed-block system. We consider a peak hour in the junction area in which four of the 28 trains enter the area with a delay. We note that initially 29 trains were scheduled within the peak hour, but two trains were delayed such that their entry time moved outside the one-hour window while one train from the hour before was delayed such that its entry time fell within the considered hour.

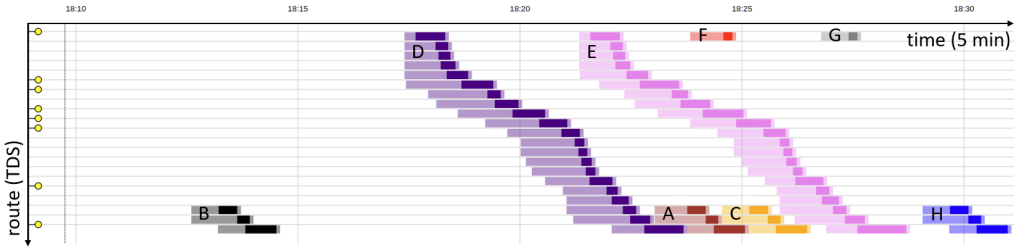
If reordering and rerouting are applied in combination with retiming, the mean total delays are 5186 seconds for ETCS L2 and 5934 seconds for conventional fixed-block signalling, as shown in Table 3.6. Additionally, Table 3.6 gives the total delay of the cross-evaluations, i.e., the ETCS L2 solution evaluated in conventional fixed block (6441 seconds) and the fixed-block solution evaluated in ETCS L2 (5582 seconds). This instance is the one with the highest impact of the ETCS L2 optimisation, showing a 7.09% reduction in objective values in ETCS L2 over conventional fixed block when both evaluated in ETCS L2. For completeness, Table 3.6 also reports the delay recovery and the number of times that the model assigns a maximum speed profile to a train. Note that the mean delay recovery is negative for the conventional fixed-block

and ETCS L2 solutions evaluated in ETCS L2. This indicates that the total exit delay exceeds the total entry delay. As Table 3.6 shows, this different distribution of the delay does not go at the expense of the overall total delay.

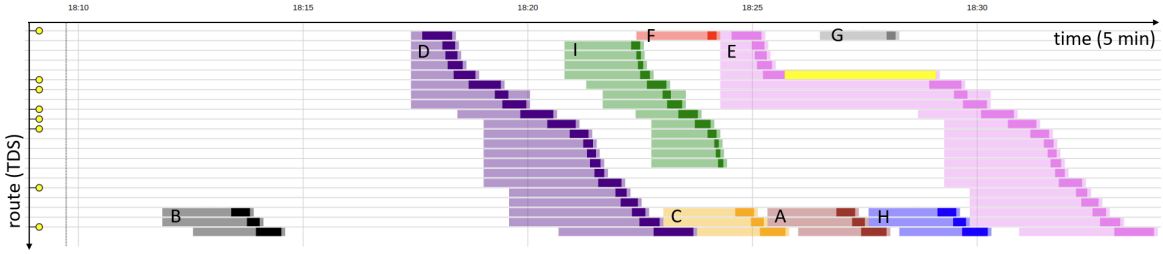
Table 3.6: Total delay, delay recovery and number of maximum speed profile assignments (speed penalty) of ETCS L2 and conventional fixed-block (original) solutions evaluated in the ETCS L2 and original models, for a specific case study instance.

Solution	**In ETCS L2**			**In original**	
	Total delay	Delay recovery	Speed penalty	Total delay	Delay recovery
ETCS L2	5186 s	-4.40%	30	6441 s	+22.64%
Original	5582 s	-5.52%	53	5934 s	+0.20%

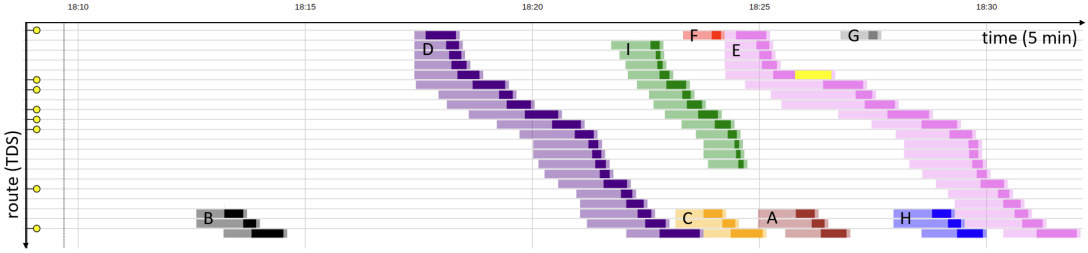
In the following, we focus on a subset of nine trains (A to I) that share a part of their route in the timetable and/or in the rescheduling solutions with train E, which is crucial for the difference between the fixed-block and the ETCS L2 solutions. Trains A to H are labelled alphabetically according to their original scheduled order, while Train I is scheduled to run on a parallel track. Figure 3.6 illustrates the blocking times of the trains in the solutions along the route of train E, which is indicated in Figure 3.5a.



(a) Schematic blocking time diagram of ETCS L2 solution.



(b) Schematic blocking time diagram of conventional fixed-block solution.



(c) Schematic blocking time diagram of conventional fixed-block in ETCS L2 evaluation.

Figure 3.6: Schematic blocking time diagrams of rescheduling solutions along train E's route in terms of TDSs, aligned in time. Physical occupation times and additional delays are indicated by dark colours and yellow, respectively.

Within this subset, trains A and C suffer entry delays of 826 and 466 seconds, respectively. Figures 3.6a and 3.6b illustrate the optimised ETCS L2 and fixed-block solutions by a schematic representation of their blocking times in terms of the TDSs along train E's route. Table 3.7 details the train ordering decisions for ETCS L2 and fixed block, listing only the train pairs that are reordered in at least one of the solutions.

Table 3.7: Orders of reordered trains in the ETCS L2 and original, i.e., conventional fixed-block, model solutions.

ETCS L2	B $\prec$ A	A $\prec$ C	D $\prec$ A	D $\prec$ C	E $\prec$ F	E $\prec$ H
Original	B $\prec$ A	C $\prec$ A	D $\prec$ A	D $\prec$ C	F $\prec$ E	H $\prec$ E

The ETCS L2 solution includes three reordering decisions and one rerouting decision relative to the timetable: B goes before A, D goes before A and C, and D is partially rerouted. The reorderings of A and C with B and D directly follow from them being delayed upon entry, letting B and D run according to schedule. The rerouting of D, which in the timetable only shares the entry and exit blocks with E, frees up some space on the fully packed parallel track. Table 3.8 presents the entry and exit delays of the trains, for the ETCS L2 and the original solutions. Note that only the trains with a positive total delay in at least one of the solutions are included.

Table 3.8: Entry and exit delays of trains with positive total delay in the listed solution evaluations (original is conventional fixed-block). The imposed train entry delay is indicated in the heading.

Delays (s)	**Train A** (+826)		**Train C** (+466)		**Train E** (+0)		**Train I** (+0)		**Total**
	entry	exit	entry	exit	entry	exit	entry	exit	
ETCS L2	826	813	524	537	0	0	48	59	2807
Original	996	988	466	493	177	314	5	0	3439
Original in ETCS L2	965	924	466	453	174	197	0	0	3179

The original solution for conventional fixed block includes the same rescheduling decisions as the ETCS L2 solution. Additionally, A and C are mutually reordered and E is reordered with F and H. Further delaying E allows F and H to run without delays. Actually, the two trains are running ahead of schedule to minimise the delay of train E. The reordering of E and F allows for the rerouting of I away from the same parallel track D was rerouted from.

Figure 3.6c shows the schematic blocking time diagram of the cross-evaluation: we apply the ordering and routing decisions of the conventional fixed-block solution considering ETCS L2. Table 3.8 provides an overview of the entry and exit delays of the trains under 'Fixed block in ETCS L2'. Before diving into the 'why' of the difference between solutions, we take a step back to the overall delays presented in Table 3.6 for the overall picture. Comparing the total delay obtained by evaluating the conventional fixed-block solution in ETCS L2 with the total delay obtained by the ETCS L2 solution (in ETCS L2), indicates an additional delay of 396 seconds due to the different rescheduling decisions. Comparing the total delay of the fixed-block in ETCS L2 evaluation and of the conventional fixed-block solution (in fixed-block), indicates a saving of 352 seconds delay due to the ETCS L2 blocking times. We note that solutions are more sensitive to the signalling system (the model) than to the rescheduling decisions.

For this case study instance, we conclude that due to the significant difference in blocking times under ETCS L2 and conventional fixed-block signalling, the conventional fixed-block solution contains additional rescheduling decisions that lead to a worse performance in ETCS L2. From an operational perspective it is therefore relevant to incorporate DTG principles into CDR model for a more accurate assessment of the performance of CDR under DTG.

3.5.5 Real-time applicability

In this section, we discuss the computational performance of the RECIFE-MILP model enhanced for ETCS L2. Given the real-time nature of the CDR problem, we specifically include the real-time applicability in this discussion. Also, we compare the computational performance of the enhanced with the original RECIFE-MILP model for conventional fixed-block signalling.

With a computation time limit of 3600 seconds, 92% of the junction scenarios and 99% of the corridor scenarios are solved to optimality. The remaining scenarios are solved sub-optimally, with mean optimality gaps of 0.28% and 0.43% for the junction and corridor case study, respectively. In line with these numbers, the time to reach optimality is often far from 3600 seconds, overall averaging 475 seconds for the junction case study and 233 seconds for the corridor case study. The variation in the utilised computation time is illustrated by Figure 3.7, which plots the percentage of scenarios solved to optimality over the computation time. Indeed, the curve's decreasingly increasing nature highlights the relative high likelihood of short utilised computation times.

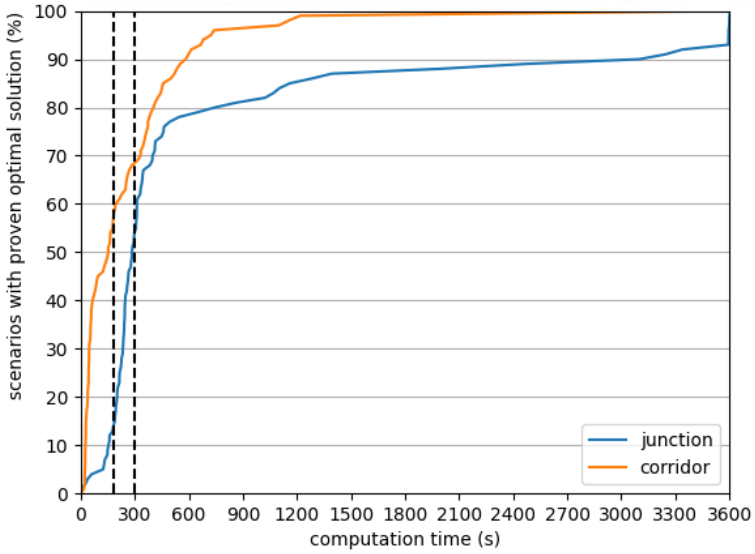


Figure 3.7: Percentage of scenarios with proven optimal solutions over computation time. The dashed lines correspond to practical computation times of 180 and 300 seconds.

In Figure 3.7, the practical computation times of 180 and 300 seconds are highlighted. For the respective case studies, 13% and 56% of the scenarios are optimally solved within 180 seconds, and 54% and 68% within 300 seconds. This shows that the model can perform in real-time.

The model's real-time performance is further discussed based on the additional computational results given in Table 3.9, which include the absolute and relative optimality gaps after

180 and 300 seconds. The absolute optimality gap compares the real-time solution with the real-time lower bound, whereas the relative optimality gap compares the real-time solution with the best lower bound obtained after 3600 seconds (which in most cases corresponds to the optimal solution). As suggested by the differences between the relative and the absolute optimality gap, it is primarily the lower bound that is improved in the further optimisation process. The relative optimality gaps indeed support the real-time applicability of the model, with all mean values below 0.5%, despite occasional maximum relative optimality gaps over 15%.

Table 3.9: Computational performance results for the ETCS L2 and original RECIFE-MILP model.

	ETCS L2 model		**Original model**	
	Junction	Corridor	Junction	Corridor
Mean computation time	475 s	233 s	334 s	25 s
Mean optimality gap of sub-opt solutions	0.28%	0.43%	2.62%	0.00%
% scenarios with proven optimal solution				
In 180 seconds	13%	56%	57%	100%
In 300 seconds	54%	68%	73%	100%
In 3600 seconds	92%	99%	92%	100%
Absolute optimality gap				
At 180 seconds - mean	1.38%	1.45%	–	0%
At 180 seconds - max	8.07%	21.76%	–	0%
At 300 seconds - mean	0.53%	0.88%	–	0%
At 300 seconds - max	4.24%	20.95%	–	0%
Relative optimality gap				
At 180 seconds - mean	0.41%	0.01%	–	0%
At 180 seconds - max	2.57%	15.89%	–	0%
At 300 seconds - mean	0.23%	0.01%	–	0%
At 300 seconds - max	2.57%	0.43%	–	0%

Next to the computational results for the ETCS L2 model, Table 3.9 reports some results for the original RECIFE-MILP model. The original model outperforms the ETCS L2 model in nearly all listed aspects - as can be expected from the different model dimensions reported in Table 3.3. Only the mean optimality gap of the sub-optimal solutions at 3600 seconds for the junction are in favour of the ETCS L2 model. As this only includes one optimality gap for the original model, we cannot take the number as a representative. A notable result of the original model is the clean score for the corridor case study; already optimally solving all scenarios within 180 seconds. To a smaller extent, it also holds for the enhanced model that the corridor area is faster to solve than junction.

Despite being outperformed by the original model, the ETCS L2 model demonstrates robust real-time performance by solving a significant portion of the scenarios to optimality within practical time limits and maintaining minimal optimality gaps for sub-optimally solved scenarios. Thus, the ETCS L2 model shows promise for real-time applications, providing a good balance between computational efficiency and solution quality. We note that the real-time applicability can be improved with a more effective model formulation and solution algorithm.

3.6 Conclusions

In this chapter, we assessed the operational relevance of modelling conflict detection and resolution (CDR) for distance-to-go (DTG) railway signalling. CDR models originally developed for conventional fixed-block signalling can be adapted to DTG signalling by incorporating train- and speed-dependent brake indication points and blocking times to capture the braking curve supervision in DTG signalling. By integrating the DTG modelling approach into the state-of-the-art RECIFE-MILP model, we obtain a CDR model for operations under the European Train Control System Level 2 with Trackside Train Detection (ETCS L2). Application to two real-world case studies, a complex junction with dense mixed traffic and a corridor with regular mixed traffic, shows longer computation times to obtain (near-)optimal solutions under delay scenarios compared to conventional fixed-bock signalling, while maintaining acceptable real-time performance of the enhanced model.

The underlying signalling system causes the CDR model to obtain different rescheduling decisions and total train delay. The ETCS L2 model outperformed the original model with respect to the minimised total train delay. While the overall reduction in delay under ETCS L2 compared to the conventional fixed-block signalling is limited, several case study instances showed significant reductions of up to 7%. Most of the reductions are directly attributable to the shorter, train- and speed-dependent train separation enabled by ETCS L2, while additional reductions arise from different rescheduling decisions, indirectly benefiting from the DTG principles. Furthermore, the inclusion of a penalty for using maximum speed profiles in the enhanced model objective results in trains running less frequently with the maximum speed profile. This provides a secondary benefit in terms of energy consumption.

Future research is needed to investigate the possibility of using the model for other DTG signalling systems such as the ETCS Level 2 variants with onboard train position and train integrity monitoring systems: ETCS Level 2 Virtual Block and ETCS Level 2 Moving Block. The modelling of ETCS Level 2 Virtual Block requires a track discretisation finer than trackside train detection sections, and such finer discretisation can also be used to approximate ETCS Level 2 Moving Block. Other topics to be investigated are the incorporation of additional and/or more advanced speed profile options, as well as the performance of the model in real-time and within a framework including a simulation environment enabling simulation-based validation. Future work should also assess the model's robustness against stochastic factors beyond entry delay that can occur in real-world settings, e.g., dwell times or braking distances. Finally, a valuable extension would be to analyse the trade-off between total delay and the assignment of maximum speed profiles.

Overall, the research topic needs to be further investigated to produce more and more efficient algorithms to enable the use of decision support systems for optimised real-time traffic management which also consider DTG signalling constraints. This includes the fine-tuning of (existing) model formulations and implementations. Promising in terms of computational performance are (geographical) decomposition approaches.

Chapter 4

Impact of track discretisation on conflict detection and resolution under virtual-block signalling

---

A conflict detection and resolution model for fixed-block distance-to-go railway operations was obtained in Chapter 3. Application of the distance-to-go model in two case studies indicated that it can propose different rescheduling decisions than the conventional fixed-block variant of the model. Chapter 3 only considered the current state-of-practice ETCS Level 2 with TTD as (fixed-block) distance-to-go signalling variant. Hence, conflict detection and resolution for the next-generation distance-to-go systems of virtual block and moving block remain unaddressed.

In this chapter, the conflict detection and resolution model for fixed-block distance-to-go signalling is adapted to describe ETCS Level 2 with Onboard TIM. The obtained model specifically describes ETCS Level 2 Virtual Block but can also approximate ETCS Level 2 Moving Block. A main feature of ETCS Level 2 Virtual Block (compared to ETCS Level 2 with TTD) is the digital discretisation of the track into shorter virtual blocks. This chapter assesses the impact of the track discretisation granularity on conflict detection and resolution under virtual-block signalling.

Apart from minor changes, this chapter has been published as:

Versluis, N. D., Pellegrini, P., Quaglietta, E., Goverde, R. M. P., & Rodriguez, J. (2025). Impact of track discretisation on conflict detection and resolution under ETCS with onboard train integrity monitoring. *Journal of Rail Transport Planning & Management*, 35, 100533.

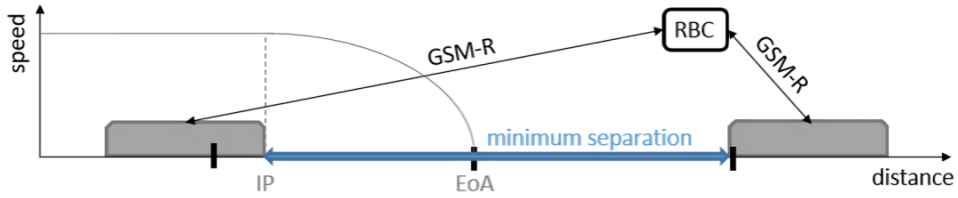
4.1 Introduction

On the railway network, safe operations are ensured through railway signalling. On the one hand, trains are separated by providing and supervising a movement authority (MA) and corresponding dynamic speed profile, i.e., the permission to move to a specific location under distance and speed monitoring, to the trains. On the other hand, train movements are protected by setting and locking train routes over movable track elements such as switches. Train separation and route protection characteristics vary depending on the implemented signalling system.

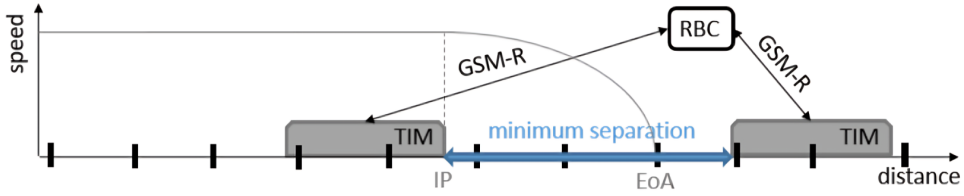
In Europe, railway signalling systems are developed in the context of the European Train Control System (ETCS) within the European Railway Traffic Management System (ERTMS) specifications with interoperability throughout Europe as overall objective. The current state-of-practice is ETCS Level 2 with Trackside Train Detection (TTD), which is implemented on freight and passenger railways throughout the world (European Rail Supply Industry Association, 2022). ETCS Level 2 with TTD is a radio-based fixed-block distance-to-go signalling system which relies on TTD for train position and train integrity monitoring. In fixed-block signalling systems, the track is partitioned into blocks (or sections) of fixed lengths which can be occupied by at most one train at a time. In conventional systems, block entries are protected by trackside multi-aspect signals which indicate whether an approaching train can proceed, needs to start braking or is required to stop. Additionally, distant signals can be used to indicate the approach of a restrictive main signal in two-aspect signalling. Radio-based distance-to-go signalling systems, instead, feature radio-based cab signalling and continuous braking curve supervision from the train front position to the section exit corresponding to the end of MA. With this, brake indications can be given to a train independent of its position on the track. As a consequence, train separation distances become speed-dependent due to the approach distance to the first occupied section being based on the train's absolute braking distance.

In the next-generation distance-to-go signalling systems ETCS Level 2 Virtual Block (VB) and ETCS Level 2 Moving Block (MB), TTD is replaced by onboard train position and train integrity monitoring (TIM). These systems are collectively referred to as ETCS Level 2 with Onboard TIM, which was known as ETCS Level 3 until 2023. ETCS Level 2 VB is still a fixed-block system, but the block boundaries are defined virtually instead of through trackside devices. In ETCS Level 2 MB, the track partitioning is eliminated such that the MA can be given up to a safety margin behind the train in front. Depending on the section lengths within the track partitioning, (ETCS Level 2) VB can be considered as an approximation of (ETCS Level 2) MB (Versluis et al., 2024; Furness et al., 2017). In addition, the railway industry is currently developing ETCS Hybrid Train Detection (HTD), previously known as ETCS Hybrid Level 3 (EEIG ERTMS Users Group, 2024). That is, ETCS with limited TTD and virtual subsections, such that the MA of a train following a train not equipped with TIM will be based on TTD sections, while it will be based on the virtual subsections when it follows a train with onboard TIM.

Figure 4.1 illustrates the minimum train separation in fixed-block distance-to-go signalling with a coarse track partitioning, e.g., ETCS Level 2 with TTD, and with a fine track partitioning, e.g., ETCS Level 2 VB. A comparison of the two demonstrates the effect of section length in fixed-block distance-to-go systems. Shorter sections lead to shorter train separation distances, which means more efficient operations, e.g., in terms of capacity.



(a) Coarse track discretisation granularity, corresponding to ETCS Level 2 with TTD.



(b) Fine track discretisation granularity, corresponding to ETCS Level 2 VB.

Figure 4.1: Schematic layout of the minimum separation between two trains in a speed-distance diagram under fixed-block distance-to-go signalling systems with different track discretisation granularities. EoA stands for end of movement authority and IP for the indication point of the braking curve to the EoA.

In terms of train separation distances, the conventional signalling systems relate to ETCS Level 2 with TTD and ETCS Level 2 VB as follows. In legacy fixed-block systems, train separation is determined by the fixed block design and the braking capabilities of the train, which together define the admissible speed and safe stopping distances. As a result, separation distances are generally larger than those in ETCS Level 2 with TTD. Under ETCS HTD, the minimum separation depends on whether the preceding train is equipped with onboard TIM. If so, the separation distance is as under ETCS Level 2 VB. If not, the separation distance is as under ETCS Level 2 with TTD. Under ETCS Level 2 MB, the train separation is shorter than under ETCS Level VB due to the independence of track discretisation.

In case of disturbed railway operations, railway traffic management is responsible for the efficient use of the available capacity. The problem of taking effective rescheduling measures in order to minimise train delays in case of disturbances can be mathematically described in conflict detection and resolution models. There are, however, very few conflict detection and resolution models for next-generation distance-to-go signalling systems such as ETCS Level 2 with Onboard TIM (Versluis et al., 2024).

In previous work, we have addressed this gap by proposing an approach to enhance existing conflict detection and resolution models to describe fixed-block distance-to-go railway operations (Versluis et al., 2025a). The incorporation of the enhancements into the state-of-the-art rescheduling model RECIFE-MILP (Pellegrini et al., 2015) resulted in a verified conflict detection and resolution model for ETCS Level 2 with TTD. A comparison of the enhanced with the original model indicated that other rescheduling decisions can be proposed due to the exploitation of distance-to-go operations to minimise train delay.

In this chapter, we continue to address conflict detection and resolution under distance-to-go signalling. Building upon our previous work, we now shift the focus towards ETCS Level 2 with Onboard TIM, specifically ETCS Level 2 VB. By considering ETCS Level 2 with Onboard TIM parameters as input for the enhanced RECIFE-MILP, we obtain a conflict detection and

resolution model for next-generation distance-to-go signalling. Additionally, we reformulate the train separation constraints in the bottleneck areas around switches to decrease their criticality in operations. The model describes ETCS Level 2 VB, but can also serve as an approximation of ETCS Level 2 MB. For both, the granularity of the track discretisation is a crucial modelling aspect. While the impact of track discretisation on railway capacity is well addressed in the literature, e.g., Cuppi et al. (2021), Ranjbar et al. (2022) and Knutsen et al. (2024), its effect on conflict detection and resolution has only been scarcely, if at all, investigated. This work aims to fill this gap by assessing how the track discretisation granularity affects the conflict detection and resolution model for ETCS Level 2 with Onboard TIM. Specifically, we focus on the effects on rescheduling decisions, complemented by an analysis of the trade-off between solution quality and computation time.

With this, the chapter's main contributions are:

- A reformulation of train separation constraints for switches based on minimum switch blocking times.
- A conflict detection and resolution model for ETCS with onboard train integrity monitoring.
- A sensitivity analysis of the conflict detection and resolution model for ETCS with onboard train integrity monitoring on track discretisation granularity.

The chapter is organised as follows. In Section 4.2, related literature is reviewed. Section 4.3 introduces the approaches related to the ETCS signalling parameters, the track (re)discretisation and the reformulation of the switch blocking times. In Section 4.4, the computational setup and experimental results are presented and discussed. The chapter is finalised with concluding remarks in Section 4.5.

4.2 Literature review

To the best of our knowledge, there is no available literature on the impact of track discretisation on conflict detection and resolution under ETCS Level 2 (or other next-generation distance-to-go signalling systems). Therefore, we divide our literature review into the following two subtopics. We address conflict detection and resolution under distance-to-go signalling in Section 4.2.1 and track discretisation in railway signalling in Section 4.2.2.

4.2.1 Towards conflict detection and resolution under ETCS Level 2

Existing conflict detection and resolution models mostly refer to conventional fixed-block multi-aspect signalling systems (Versluis et al., 2024). Two state-of-the-art models are ROMA (D'Ariano et al., 2008) and RECIFE-MILP (Pellegrini et al., 2015). They represent two main categories of conflict detection and resolution models: alternative graph-based, e.g., Mazzarello & Ottaviani (2007), Corman et al. (2009) and Janssens (2022), and MILP-based models, e.g, Törnquist & Persson (2007), Pellegrini et al. (2014) and Luan et al. (2018), respectively. Actually, all MILP-based models mentioned so far are disjunctive MILP models. Another class of MILP-based models are time-indexed models, as proposed by, e.g., Lusby et al. (2013), Reynolds et al. (2020) and Bettinelli et al. (2017). Some other modelling approaches applied to

conflict detection and resolution are constraint programming (Rodriguez, 2007; Marlière et al., 2023) and model predictive control (Caimi et al., 2012; Pochet et al., 2016).

The different approaches each have their own strengths and limitations in terms of, e.g., inclusion of speed, suitability for rerouting, computational performance and flexibility in objective function (Versluis et al., 2024). For example, classic alternative graph models can only feature graph-based objectives, typically the minimisation of the maximum secondary train delay, while MILP models can feature any objective as long as it can be represented by a linear function. Moreover, MILP models are more flexible in terms of rerouting. Disjunctive MILP models, however, rely on the general big-M method, resulting in a relatively weak linearisation which makes them hard to solve to optimality. The other MILP-based approach of time-indexed models is an alternative in that aspect. The time-indexed model formulation has, however, the drawback of a large model size due to the explicit need of time-space resources.

Another way to counter the limitations of big-M formulations is by decomposition of the conflict detection and resolution problem, e.g., Lamorgese & Mannino (2015), Yi et al. (2023) and Lippes (2024). Lamorgese & Mannino (2015) consider an exact macro/micro decomposition method inspired by Benders decomposition (Benders, 1962). They propose an iterative approach between the overall problem which is solved at a macroscopic level and the subproblems which are searching feasible train routes through stations. Yi et al. (2023) also consider a macroscopic and a microscopic level of the problem. At the microscopic level, for each control area separately, rescheduling decisions are proposed by the RECIFE-MILP model. At the macroscopic level, coordination of the local decisions is performed by a time-indexed MILP model. This iterative coordination framework is independent from the approach used for the rescheduling at the microscopic level. Lippes (2024) proposes an iterative distributed approach to describe moving-block conflict detection and resolution. The model used for the conflict resolution in subareas is based on the disjunctive MILP formulation of Törnquist & Persson (2007). Lippes (2024) solves the subproblems sequentially with coordination intervention in case of dependencies. The works show significant improvements in computational performance compared to centralised approaches.

The sparse literature on conflict detection and resolution models for ETCS Level 2 or, more general, distance-to-go signalling includes the work of Pochet et al. (2016), Xu et al. (2017, 2021), Liu et al. (2021), Janssens (2022) and Versluis et al. (2025a). Pochet et al. (2016) address the topic of moving-block (CBTC) conflict detection and resolution by applying a model predictive control approach on a suburban railway network. Without notion of speed, few practical scenarios of disturbed operations are optimised in terms of punctuality. The approach is incorporated in a microscopic simulation tool of the French train operator SNCF. The works of Xu et al. (2017) and Liu et al. (2021) are in the context of the Chinese high-speed railways with homogeneous traffic under fixed-block distance-to-go signalling (CTCS-3). Both propose disjunctive MILP formulations, with Xu et al. (2017)'s model being based on the alternative graph model underlying ROMA (D'Ariano et al., 2008). Xu et al. (2017) consider speed-dependent train separation, with a discrete set of speed levels included in an extension (Xu et al., 2021), similar to the modelling of speed by Liu et al. (2021). Janssens (2022) also takes the ROMA model formulation as starting point, extending the alternate graph model to describe moving-block operations. Actually, an approximation of moving block by considering fixed blocks of train length is proposed - without a notion of speed or rerouting options. That leaves our previously mentioned work leading to a fixed-block distance-to-go reschedul-

ing model (Versluis et al., 2025a). More specifically, a conflict detection and resolution model for ETCS Level 2 with TTD obtained from the original RECIFE-MILP model for conventional fixed-block signalling. The presented MILP model considers speed-dependent train separation, given two speed profile options related to maximum and scheduled speed. The combination of a disjunctive MILP model with a commercial solver makes it hard to prove optimality of solutions. However, the real-time performance is shown to be acceptable.

For more information, we refer to the review papers of Cacchiani et al. (2014) and Versluis et al. (2024). In the latter, research gaps and challenges related to the modelling of conflict detection and resolution under moving-block signalling are identified, and research steps to address those are proposed. These include the investigation of the approximation of moving-block signalling by fixed-block distance-to-go signalling with a fine discretisation of the track.

4.2.2 Track discretisation in railway signalling

Track discretisation has been previously studied in the literature, primarily focusing on capacity improvement. Recently, ETCS HTD has renewed research interest in this area. The virtual subsection lengths typically considered for ETCS HTD are 25 to 500 metres (Knutsen et al., 2024). The minimum length of 25 metres is put forward in the ETCS HTD specification document as the length that should provide a capacity similar to a moving-block system (EEIG ERTMS Users Group, 2024). This is somewhat in contrast to the assertion by Furness et al. (2017), who state that with virtual block or subsection lengths of 200 metres, moving-block performance can be achieved, assuming speeds of circa 160 km/h on the open track.

The impact of section length on the railway (signalling) system is typically considered in terms of (static) capacity. While Knutsen et al. (2024) consider several performance indicators to measure capacity effects, capacity consumption/occupation rate is typically considered as key performance indicator. That goes, for example, for the impact assessment studies of Jansen (2019) and Vergroesen (2020). In both works, the capacity impact of virtual section lengths on ETCS HTD is assessed using simulation. Jansen (2019) considers virtual subsection lengths of 500 metres and 100 metres, which lead to a decrease in capacity consumption compared to the legacy (fixed-block) system of 16.2% and 20.1%, respectively. Guided by the requirements of the Dutch infrastructure manager ProRail, it was proposed to have section lengths of 200 metres with 100 metres only at critical infrastructure. Interestingly, this work also includes a test analysis of some simple delay scenarios, i.e., a single train having a 10- or 30-minute departure delay. Serious decreases in secondary delay and also significantly shorter recovery times are reported for ETCS HTD compared to the legacy system. The benefits of the smaller section lengths are much smaller, but still significant with ca. 15% decrease in secondary delay and 2% to 10% decrease in recovery time. We note that only retiming is applied, no reordering or rerouting. Building upon the work of Jansen (2019), Vergroesen (2020) specifically considers the impact on ETCS HTD in combination with automatic train operation. By subdividing TTD sections into virtual (sub)sections, the occupation rate is reduced by 16 percent point.

Cuppi et al. (2021) and Ranjbar et al. (2022) assess the capacity benefits of ETCS HTD in Italian and Swedish studies, respectively. Cuppi et al. (2021) consider virtually dividing the existing blocks of the Roman railway node into subsections with lengths of 350 to 450 metres. From a simulation-based comparison with the current multi-aspect signalling system, capacity

increases of at least 56% are reported for the stations in the area. Ranjbar et al. (2022) compare the capacity consumption of ETCS HTD with ETCS Level 2 with TTD and the Swedish legacy signalling system on a single line with homogeneous traffic. For ETCS HTD with virtual subsection lengths of 100 to 20 metres, microscopic simulation showed 9% and 14% capacity gain compared to ETCS Level 2 with TTD and the legacy system, respectively. The authors propose further analysis of the effects of ETCS HTD with heterogenous traffic considering delays as future work.

We conclude that the effects of track discretisation on capacity under virtual-block signalling are well-established. However, to the best of our knowledge, no more than preliminary work can be reported about the consequent effects of track discretisation on conflict detection and resolution.

4.3 Methodology

With the aim to obtain a conflict detection and resolution model for next-generation distance-to-go signalling such as ETCS Level 2 with Onboard TIM, we build upon the earlier presented conflict detection and resolution model for fixed-block distance-to-go signalling (Versluis et al., 2025a). The further development of the model is captured in three steps. First, we introduce the ETCS Level 2 with Onboard TIM signalling parameters and their relation with the minimum train separation in Section 4.3.1. Second, we address the track discretisation granularity by presenting a 'rediscretisation procedure' departing from the existing track description in Section 4.3.2. Third, in Section 4.3.3, we propose a reformulation of the blocking time constraints for switch areas to better align with the short train separation under next-generation distance-to-go signalling.

4.3.1 Train blocking times for ETCS Level 2 with Onboard TIM

The signalling system in place determines the safe train separation. A well-known concept for the modelling of minimum train separation is blocking time theory (Hansen & Pachl, 2014). The blocking time components are setup, reaction, approach, running, clearing and release time. Figure 4.2 illustrates the relation between the blocking time components and the minimum train separation under fixed-block distance-to-go signalling, for a specific section.

The setup time is the time required to set a route, to update the MA and translate it into a dynamic speed profile on the driver machine interface onboard. This time is dependent on the type of section that is requested. Specifically, whether the section contains a switch that possibly needs to be set and locked. In any case, there is a value to be derived as minimum setup time. For that value, the following subcomponents of the setup time are considered: the trackside processing time, the trackside-to-train communication time and the onboard computation time. Under ETCS, these subcomponents correspond to the Radio Block Centre (RBC) processing time, the RBC-to-train communication time and the European Vital Computer (EVC) computation time. From PERFORMINGRAIL (2022b), we obtain the respective values of 0.5 seconds, 1.0 second and 1.5 seconds.

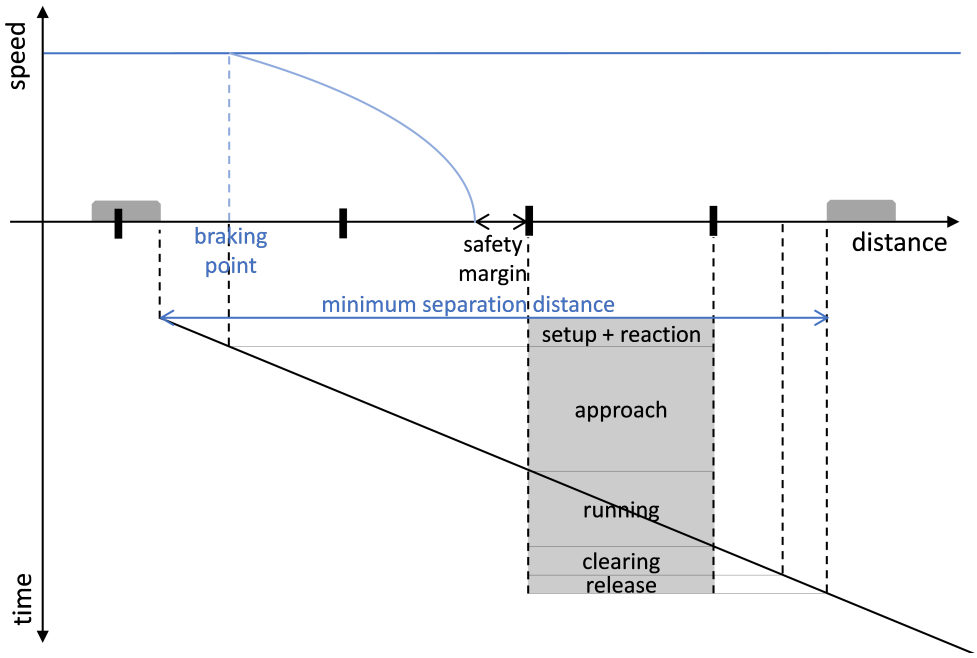


Figure 4.2: Blocking time components related to minimum train separation under fixed-block distance-to-go signalling.

For switch sections, the minimum setup time does not suffice given the additional safety requirements concerning movable track elements. Therefore, we consider the following two additional subcomponents: the switch position evaluation time and the switch turning time. The first component captures the time needed to verify the position in which the switch is set, which is around one second. The position evaluation time is in place in all cases of a train requesting a route over a switch. The second component captures the time needed for a switch to be set in the required position. In practice, different types of switches have a different turning time, ranging from three up to 16 seconds for high-speed switches. The switch turning time is practically only in place in case a train requests a route over a switch that is set in the wrong position. We derive a switch turning time of 6.0 seconds as representative value as it corresponds to the most common type of switch.

The reaction time is the time from the moment the updated speed profile appears on the driver machine interface until the train starts to decelerate after the train driver has applied the brakes, if needed. Hence, it includes a subcomponent for driver reaction time as well as an additional brake build-up time. Given this split, we obtain 12.0 seconds as value for the reaction time. That is, 8.0 seconds for the driver reaction time (PERFORMINGRAIL, 2022b) and 4.0 seconds for the brake build-up time (European Union Agency for Railways, 2020).

The approach time is the time for the train to run from the braking point to the EoA. More specifically, it is the time to run over the braking distance (plus an additional safety margin), as also visualised in Figure 4.2. This distance-to-go approach distance, with its speed- and train-dependency, is formalised in Versluis et al. (2025a).

The running time and the clearing time together represent the physical occupation time of the train on the section. As first occupation time component, the running time captures the time to run from the entry to the exit of the section with the front of the train. Subsequently, the clearing time is the time from the moment the front of the train exits the section until the rear of

the train exits as well. The occupation time is not directly affected by the signalling principles. However, it is clearly dependent on the speed and the featured length of the sections.

The release time is the time from the moment a section is physically cleared by a train until it can be set for another train. Hence, the release time relates to the detection and the communication of the section clearance. In this component, there is a difference between ETCS Level 2 with TTD and with Onboard TIM due to the replacement of trackside train detection with onboard train position and integrity monitoring. In line with Shift2Rail (2023), we assume the implementation of Global Navigation Satellite System (GNSS) as train localisation and integrity monitoring technology. Based on PERFORMINGRAIL (2022b), we obtain 2.0 seconds as GNSS-based release time. This value is derived from the train positioning report (TPR) update time of 1.0 second and the train-to-trackside, i.e., train-to-RBC, communication time of 1.0 second. Note that the TPR contains both train position and train integrity information. Table 4.1 summarises the ETCS Level 2 with Onboard TIM parameters and their values relevant for the train blocking times.

Table 4.1: ETCS Level 2 with Onboard TIM parameter values for train blocking times.

Component	**Value**
Minimum setup time	*3.0 s*
Trackside processing time	0.5 s
Trackside-to-train communication time	1.0 s
Onboard computation time	1.5 s
Switch setup time	*1.0 s / 7.0 s*
Switch position evaluation time	1.0 s
Switch position turning time	6.0 s
Reaction time	*12.0 s*
Driver reaction time	8.0 s
Brake build-up time	4.0 s
Release time	*2.0 s*
GNSS-based TPR update period	1.0 s
Train-to-trackside communication time	1.0 s

4.3.2 Track discretisation procedure

As introduced in Section 4.1, the minimum train separation under fixed-block distance-to-go signalling depends on the track discretisation. Given that we build upon a model describing ETCS Level 2 with TTD, we consider the existing track discretised into TTD sections, which we simply refer to as *sections*. In practice, these sections have varying lengths, for example ranging from 35 to 2424 metres and from 100 to 2217 metres within the control areas presented in Section 4.4.1. However, for next-generation distance-to-go signalling systems such as ETCS HTD, virtual section lengths in the range of 25 to 500 metres are commonly considered (Section 4.2.2).

We propose a procedure to virtually discretise the existing track into shorter sections based on the sections in place (see Algorithm 4.1). Sections exceeding a given maximum length are

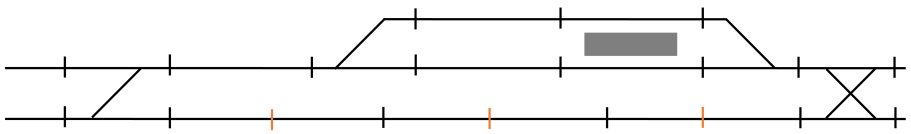
discretised into subsections of equal length. The number of subsections is determined by the ratio of the original section length over the maximum length, rounded up to the nearest integer. This ensures that subsection lengths are between half the maximum length and the maximum length. For *switch sections*, i.e., sections containing a switch, the original sections are kept to ensure compliance with safety regulations. Given the varying section lengths, the resulting discretisation will also be non-uniform. However, we obtain a more uniform discretisation than originally.

Algorithm 4.1 Track discretisation procedure.

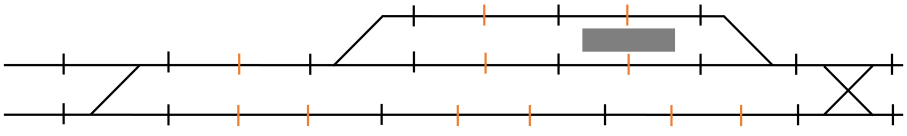
Input: original sections, switch sections, maximum length $l_{\max}$
Output: discretised sections

for all original sections **do**
  if original section is not a switch section and section length $> l_{\max}$ **then**
    $n = \lceil$ section length / $l_{\max} \rceil$ $\{$*number of subsections*$\}$
    $l =$ section length / n $\{$*subsection length*$\}$
    discretise original section into n subsections of length l
  end if
end for

For illustration of the track discretisation procedure, we consider a fictive double-track line, as shown in Figure 4.3. The line connects two switch areas on the left and right, with a side track around a station platform. The bottom track features original sections with lengths of around 600 metres, while the top and side track feature original sections of around 360 metres. We apply the discretisation procedure given two different maximum lengths: 400 metres and 200 metres. For the 400-metre maximum length, only the original sections on the bottom track between the switch areas are longer and therefore discretised into two subsections of 300 metres each (see Figure 4.3a). For the 200-metre maximum length, all non-switch sections are discretised. The sections on the bottom track are divided into three subsections of 200 metres each, while the ones on the top and side track are divided into two subsections of 180 metres each (see Figure 4.3b).



(a) Discretisation for maximum length of 400 m.



(b) Discretisation for maximum length of 200 m.

Figure 4.3: Illustration of track discretisation procedure, with original section boundaries in black and additional section boundaries due to discretisation in orange.

4.3.3 Train separation at switches

In the modelling of train separation, we need to differentiate between sections with and without a switch (Versluis et al., 2025a). Recalling that train separation is modelled in terms of train blocking times, we know from Section 4.3.1 that a main difference between switch and non-switch sections lies in the switch setup time. For ease of the formulation (see Section 3.4.2), only the minimum setup time is considered in the blocking times. The switch setup time is, instead, included as an additional blocking time component.

In the following, the reformulation of train separation for switches is presented. First, the new and other relevant model components are introduced. Subsequently, the constraints related to the train blocking times are described. Finally, the model objective function is recalled.

For the model formulation of the earlier version of the conflict detection and resolution model, i.e., for ETCS Level 2 with TTD, as presented in (Versluis et al., 2025a), we refer to Section 3.4.2. In the model, we consider track locations, i.e., the entry points of track sections, rather than track sections themselves. Here, we provide a brief description of the model's main ideas and concepts. The model considers two speed profile options in line with the scheduled and maximum speed. With a preference for the former through the penalisation in the model objective of the latter, the speed profiles are assigned to trains per stretch of track between switch areas along their route. Hence, trains can only transition between speed profile options after switch areas where speed is limited and assumed to be the same for the two speed profile options. The speed profile option assigned affects the train running and clearing times on non-switch sections. The running and clearing times together form the (physical) occupation time. The model allows for an extension of the occupation time if it is not possible to exactly adhere to a specific speed profile, including station dwell times. Similar to how the speed profile options are assigned, train orderings are fixed on the level of track stretches, namely per sequence of track sections that are shared by the routes of a pair of trains.

Sets, parameters and variables

To update the MILP model in line with the assumptions, new notation is introduced, see Table 4.2. For a description of the previously defined sets, parameters and variables, we refer to Versluis et al. (2025a). Here, we describe the new elements per category.

First, we introduce two subsets of routes for switch locations, i.e., the entry points of sections containing a switch. R_l^0 contains the routes using the switch in straight position 0, and R_l^1 contains the routes using the switch in diverging position 1.

Second, we introduce two parameters related to the switch setup time. sw^{eval} indicates the time to evaluate the position of a switch, while sw^{turn} represents the time to turn a switch from one position to the other. Regarding the switch turning time, we assume the simultaneous turning of switches that are set together under the route-locking interlocking system.

Third, we introduce binary variables related to required switch positions. $I_{t,t',l}$ indicates whether trains $t, t' \in T$ are assigned routes that use common switch $l \in L_t \cap L_{t'}$ with $s_l = 1$ in opposite position. Note that these 'indicator' variables are auxiliary variables as their values directly follow from the values of the route assignment variables.

Table 4.2: Notation for updated blocking time constraints.

Element	Description
T	ordered set of trains
R	set of routes
L	set of locations
$L_t \subset L$	set of locations which can be used by train $t \in T$
$OL_{t,r,l} \subset L^r$	set of locations along route $r \in R_t$ such that if train $t \in T$ starts occupying it, the train has not yet cleared location $l \in L^r$, $l \notin OL_{t,r,l}$
$\hat{L}_{t,t',l} \subset L$	set of locations $l' \in L_t \cap L_{t'}$ which may be used by both trains $t,t' \in T$ such that if t precedes t' on l, then t precedes t' also on l'
$\sigma_{r,l} \in L^r$	succeeding location of location $l \in L^r$ along route $r \in R$
$\rho_{r,l} \in \mathrm{P}^r$	speed assignment location associated with location $l \in L^r$ along route $r \in R$
$s_l \in \{0,1\}$	$= 1$ if location $l \in L$ lies in a switch area
$for_{r,l} \in \mathbb{R}_+$	formation time, i.e., setup and reaction time, of location $l \in L^r$ along route $r \in R$
$rel_{r,l} \in \mathbb{R}_+$	release time of location $l \in L^r$ along route $r \in R$
$\Delta ct_{t,r,l} \in \mathbb{R}_+$	additional clearing time for train $t \in T$ of location $l \in L_t$ along route $r \in R_t$ in case of scheduled speed profile
$ref^{\mathrm{s}}_{t,r,l} \in L^r$	reference brake location for location $l \in L^r$ along route $r \in R_t$ for train $t \in T$ approaching according to scheduled speed profile
$ref^{\mathrm{m}}_{t,r,l} \in L^r$	reference brake location for location $l \in L^r$ along route $r \in R_t$ for train $t \in T$ approaching according to maximum speed profile
$lag^{\mathrm{s}}_{t,r,l} \in \mathbb{R}_+$	time by which blocking of location $l \in L^r$ by train $t \in T$ running according to scheduled speed profile along route $r \in R_t$ can be postponed after passing $ref^{s}_{t,r,l}$
$lag^{\mathrm{m}}_{t,r,l} \in \mathbb{R}_+$	time by which blocking of location $l \in L^r$ by train $t \in T$ running according to maximum speed profile along route $r \in R_t$ can be postponed after passing $ref^{m}_{t,r,l}$
$M \in \mathbb{R}_+$	a large constant
$y_{t,t',l} \in \{0,1\}$	$= 1$ if train $t \in T$ blocks location $l \in L_t \cap L_{t'}$ before train $t' \in T$
$x_{t,r} \in \{0,1\}$	$= 1$ if train $t \in T$ uses route $r \in R_t$
$v^{\mathrm{s}}_{t,r,l} \in \{0,1\}$	$= 1$ if train $t \in T$ passes speed assignment location $l \in \mathrm{P}^r$ along route $r \in R_t$ according to scheduled speed profile
$o_{t,r,l} \in \mathbb{R}_+$	occupation start time of train $t \in T$ on location $l \in L^r$ along route $r \in R_t$
$o^{+}_{t,r,l} \in \mathbb{R}_+$	extended occupation time of train $t \in T$ between locations $l \in L^r$ and $\sigma_{r,l} \in L^r$ along route $r \in R_t$
$b^{\mathrm{s}}_{t,l}, b^{\mathrm{e}}_{t,l} \in \mathbb{R}_+$	time at which train $t \in T$ starts/ends blocking location $l \in L_t$
New notation	
$R^i_l \in R$	set of routes over switch location $l \in L$ requiring position $i \in \{0,1\}$
$sw^{eval} \in \mathbb{R}_+$	switch position evaluation time
$sw^{turn} \in \mathbb{R}_+$	switch position turning, i.e., setting and locking time
$I_{t,t',l} \in \{0,1\}$	= 1 if trains $t,t' \in T$ are assigned routes that use common switch $l \in L_t \cap L_{t'}$ in opposite position

Constraints

Constraints (4.1) to (4.9) are the model constraints related to the train blocking times. They can be classified into four categories as follows. Constraints (4.1) to (4.3) are the start and end blocking constraints. Constraints (4.4) and (4.5) are the original disjunctive constraints, now

only for non-switch locations. Constraints (4.6) and (4.7) are the disjunctive constraints for switches. Constraints (4.8) and (4.9) are the 'indicator' constraints.

$$b^{\mathrm{s}}_{t,l} \leq \sum_{\substack{r \in R_t: \\ l \in L^r}} \left(o_{t,r,ref^{\mathrm{s}}_{t,r,l}} + \left(lag^{\mathrm{s}}_{t,r,l} - for_{r,l} \right) x_{t,r} \right) \qquad \forall t \in T,\ l \in L_t, \quad (4.1)$$

$$b^{\mathrm{s}}_{t,l} \leq \sum_{\substack{r \in R_t: \\ l \in L^r}} \left(o_{t,r,ref^{\mathrm{m}}_{t,r,l}} + \left(lag^{\mathrm{m}}_{t,r,l} - for_{r,l} \right) x_{t,r} + M\, v^{\mathrm{s}}_{t,r,\rho_{r,ref^{s}_{t,r,l}}} \right) \qquad \forall t \in T,\ l \in L_t, \quad (4.2)$$

$$b^{\mathrm{e}}_{t,l} = \sum_{\substack{r \in R_t: \\ l \in L^r}} \left(o_{t,r,\sigma_{r,l}} + \left(ct_{t,r,\sigma_{r,l}} + rel_{r,\sigma_{r,l}} \right) x_{t,r} + \Delta ct_{t,r,\sigma_{r,l}} v^{\mathrm{s}}_{t,r,\rho_{r,\sigma_{r,l}}} + \sum_{\substack{l' \in L^r: \\ l' \in OL_{t,r,\sigma_{r,l}}}} o^{+}_{t,r,l'} \right)$$

$$\forall t \in T,\ l \in L_t, \quad (4.3)$$

$$b^{\mathrm{e}}_{t,l} - M\left(1 - y_{t,t',\hat{l}}\right) \leq b^{\mathrm{s}}_{t',l} \qquad \forall t,t' \in T, t \prec t', l,\hat{l} \in L_t \cap L_{t'} : l \in \hat{L}_{t,t',\hat{l}} \wedge s_l = 0, \quad (4.4)$$

$$b^{\mathrm{e}}_{t',l} - M\, y_{t,t',\hat{l}} \leq b^{\mathrm{s}}_{t,l} \qquad \forall t,t' \in T, t \prec t', l,\hat{l} \in L_t \cap L_{t'} : l \in \hat{L}_{t,t',\hat{l}} \wedge s_l = 0, \quad (4.5)$$

$$b^{\mathrm{e}}_{t',l} - M\left(1 - y_{t,t',\hat{l}}\right) \leq b^{\mathrm{s}}_{t,l} - sw^{eval} \left(\sum_{r \in R^0_l \cap R^1_l} x_{t,r} + \sum_{r' \in R^0_l \cap R^1_l} x_{t',r'} - 1 \right) - sw^{turn} I_{t,t',l}$$

$$\forall\, t,t' \in T, t \prec t',\ l,\hat{l} \in L_t \cap L_{t'} : l \in \hat{L}_{t,t',\hat{l}} \wedge s_l = 1, \quad (4.6)$$

$$b^{\mathrm{e}}_{t',l} - M\, y_{t,t',\hat{l}} \leq b^{\mathrm{s}}_{t,l} - sw^{eval} \left(\sum_{r \in R^0_l \cap R^1_l} x_{t,r} + \sum_{r' \in R^0_l \cap R^1_l} x_{t',r'} - 1 \right) - sw^{turn} I_{t,t',l}$$

$$\forall\, t,t' \in T, t \prec t',\ l,\hat{l} \in L_t \cap L_{t'} : l \in \hat{L}_{t,t',\hat{l}} \wedge s_l = 1, \quad (4.7)$$

$$I_{t,t',l} \geq \sum_{r \in R^0_l} x_{t,r} + \sum_{r' \in R^1_l} x_{t',r'} - 1 \qquad \forall\, t,t' \in T, t \prec t',\ l \in L_t \cap L_{t'} \wedge s_l = 1, \quad (4.8)$$

$$I_{t,t',l} \geq \sum_{r \in R^1_l} x_{t,r} + \sum_{r' \in R^0_l} x_{t',r'} - 1 \qquad \forall\, t,t' \in T, t \prec t',\ l \in L_t \cap L_{t'} \wedge s_l = 1. \quad (4.9)$$

Constraints (4.1) to (4.3) are the start and end blocking constraints. Constraints (4.1) and (4.2) describe the speed-dependent blocking start times of track locations per train. Note that the track locations refer to the entry points of sections, as introduced in Versluis et al. (2025a). Constraints (4.1) ensure that, when a train is approaching a location according to a scheduled speed profile, the location is blocked for that train at the latest the setup and reaction (together 'formation') time before the train passes the 'scheduled-speed' braking distance plus a safety margin. This is captured by the start occupation time of the route-dependent 'scheduled-speed' brake reference location $ref^{s}_{t,r,l}$ and the corresponding reservation lag time $lag^{s}_{t,r,l}$. Similarly, Constraints (4.2) ensure the timely blocking in case of a maximum speed profile. With the 'maximum-speed' braking point lying before the 'scheduled-speed' braking point, a big-M term is added to relax the constraint in case of a scheduled speed profile. We set the same start blocking constraints for all locations, independently of the presence of switches. We note that this is different from Versluis et al. (2025a), where the start of blocking of 'succeeding switch locations' is set together with the preceding switch location's blocking start time. Constraints (4.3) set the blocking end times. The blocking of a location lasts until the train has fully passed the succeeding location along its route, plus the release time. In line with that, the speed-dependent clearing time is included in the blocking time. Additionally, if the

train is long enough to keep occupying a location when the front is at the end of the following ones (included in set $OL_{t,r,l}$), also the extended occupation times of the train for these locations has to be accounted for.

Constrains (4.4) and (4.5) are the disjunctive ordering constraints for non-switch locations. With these disjunctive constraints, it is decided which of a pair of trains is passing first. They are the disjunctive constraints as they were set for all locations in the original RECIFE-MILP model for fixed-block as well as in the enhanced ETCS Level 2 with TTD version. With these constraints, it is ensured that the blocking times of a location by any pair of trains do not overlap.

Constraints (4.6) and (4.7) are the disjunctive ordering constraints for switch locations. The original disjunctive constraints are updated to capture the earlier introduced switch blocking time variants. Beforehand, we note that the constraints are only more restrictive than the original constraints for the combination of a train pair and a switch location if both trains are assigned routes that use the switch. This follows from the values of the 'switch parameter factors', i.e., the 'switch turning' indicator $I_{t,t',l}$ and the 'switch evaluation' expression $\sum_{r\in R_l^1\cap R_l^2} x_{t,r} + \sum_{r'\in R_l^1\cap R_l^2} x_{t',r'} - 1$. If it is not the case that both trains use the location, the indicator takes value 0 and, hence, the switch setting and locking time is dismissed. Similarly, the expression will then be at most 0, with which the switch position evaluation time is effectively cancelled out. In case both trains use the switch, then the switch evaluation time is added as separation of the train blocking times. The 'switch evaluation' expression will equal 1 as both trains then have a route assigned that uses the switch in either position 0 or 1. In case both trains use the switch in different positions, then also the switch turning time is added because if so, the switch turning indicator takes value 1. If the trains use the switch in the same position, the additional switch separation times is solely the position evaluation time as the indicator's value is 0.

Constraints (4.8) and (4.9) ensure that the 'switch turning indicator' takes the right value. Constraints (4.8) sets the indicator value to 1 if trains t and t' are assigned routes that use switch l in positions 0 and 1, respectively, and Constraints (4.9) if vice versa.

Objective function

The main objective of the conflict detection and resolution model is to minimise the total delay. Additionally, we want to enforce the assignment of scheduled speed profiles where possible. In the formulation of the objective function, we make use of the following notation. The sets T, S_t, R_t and P^r contain respectively the trains, the stations where train $t \in T$ has a scheduled stop, the routes which can be used by train $t \in T$, and the speed assignment locations along route $r \in R$. The auxiliary variables $z_t \in \mathbb{R}_+$ and $z_{t,s} \in \mathbb{R}_+$ capture the final delay and the delays at scheduled stop $s \in S_t$ of train $t \in T$, while decision variable $v^{\mathrm{m}}_{t,r,l} \in \{0,1\}$ indicates whether train $t \in T$ runs according to the maximum speed profile over speed assignment location $l \in P^r$ along route $r \in R_t$. With this notation, the objective function is formulated as follows:

$$\text{minimise} \sum_{t\in T} \Big((z_t + \sum_{s\in S_t} z_{t,s}) + \sum_{\substack{r\in R_t:\\ l\in \mathrm{P}^r}} v^{\mathrm{m}}_{t,r,l} \Big).$$

4.4 Computational experiments

Computational experiments are carried out with the aim to assess the impact of the track discretisation granularity on the conflict detection and resolution model for ETCS Level 2 with Onboard TIM. The assessment is focused on the impact in terms of rescheduling decisions, specifically reordering and rerouting, but also covers computational complexity aspects. For this assessment, six different discretisation granularities are considered. From the original partitioning of the track into TTD sections, alternative descriptions of the infrastructure in terms of shorter sections are obtained by applying the track discretisation procedure presented in Section 4.3.2. The considered maximum section lengths are 800, 400, 200, 100 and 50 metres. These maximum lengths are in line with the commonly considered virtual section lengths in ETCS HTD and ETCS Level 2 VB of 25 to 500 metres (Knutsen et al., 2024). 25 metres is mentioned as minimum length that provides a capacity comparable to moving block (Knutsen et al., 2024; EEIG ERTMS Users Group, 2024), however, already from 200 metres down the performance of ETCS Level 2 VB and MB are expected to be similar due to the communication delays in ETCS (Furness et al., 2017; Knutsen et al., 2024).

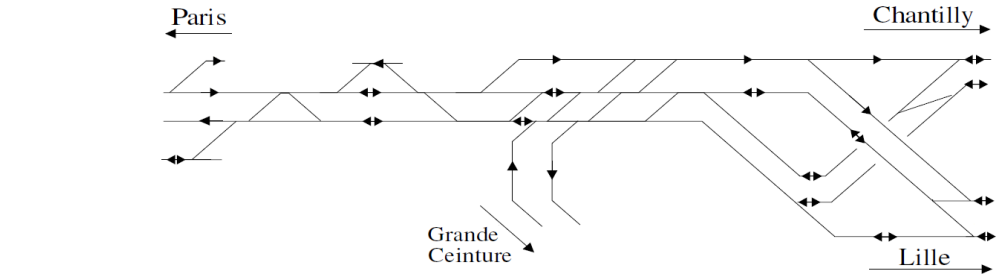
The experimental setup and results are presented in the following sections. Two case studies are described in Section 4.4.1. In Section 4.4.2, insights into model complexity are obtained by looking into the effects of the number of sections on the model dimensions. Section 4.4.3 presents the results of the computational experiments for the two case studies. Concludingly, Section 4.4.4 discusses the rescheduling decisions across the discretisations.

4.4.1 Case studies: Gonesse junction and Rosny-StEtienne corridor

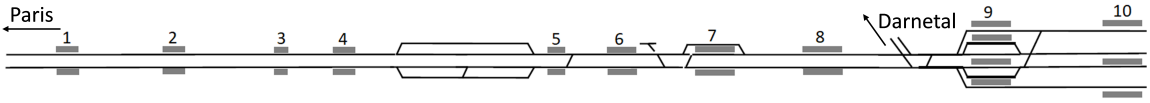
The computational experiments are performed in two case studies representing traffic control areas in France: the Gonesse junction and the Rosny-StEtienne corridor. The Gonesse area is a 17-kilometre long complex junction north of Paris with dense mixed traffic. Figure 4.4a provides a schematic representation of the Gonesse junction. The junction area includes 89 TTD sections, 38 of which contain a switch, with lengths ranging from 35 to 2424 metres, with a mean of 560 metres. The TTD sections are grouped into 79 blocks and 37 routes. The area has no station platforms. The timetable of a weekday includes 336 trains, of which 116 high-speed, 129 conventional, and 91 freight trains, with 5 to 13 route alternatives per train.

The Rosny-StEtienne area is a 68-kilometre long portion of the Paris-Le Havre line with mixed traffic. Figure 4.4b provides a schematic representation of the Rosny-StEtienne corridor. The corridor includes 239 TTD sections, 44 of which contain a switch, with lengths ranging from 100 to 2217 metres, with a mean of 740 metres. The TTD sections are divided over 152 blocks and 169 routes. The area has 10 stations with a total of 39 platforms. The daily timetable features 215 trains, of which 2 high-speed, 122 conventional, 56 freight and 35 empty (including work and test) trains, with 1 to 24 route alternatives each. We note that the Rosny-StEtienne corridor is not a high-speed line, so high-speed trains use the maximum speed of conventional trains.

For both traffic control areas, 25 delay scenarios are obtained from a one-day timetable, where an entry delay between 5 and 15 minutes is imposed on 20% of the trains. The affected trains and the delays are randomly selected following uniform probability distributions to ensure



(a) Gonesse junction.



(b) Rosny-StEtienne corridor.

Figure 4.4: Schematic track layouts of the considered control areas.

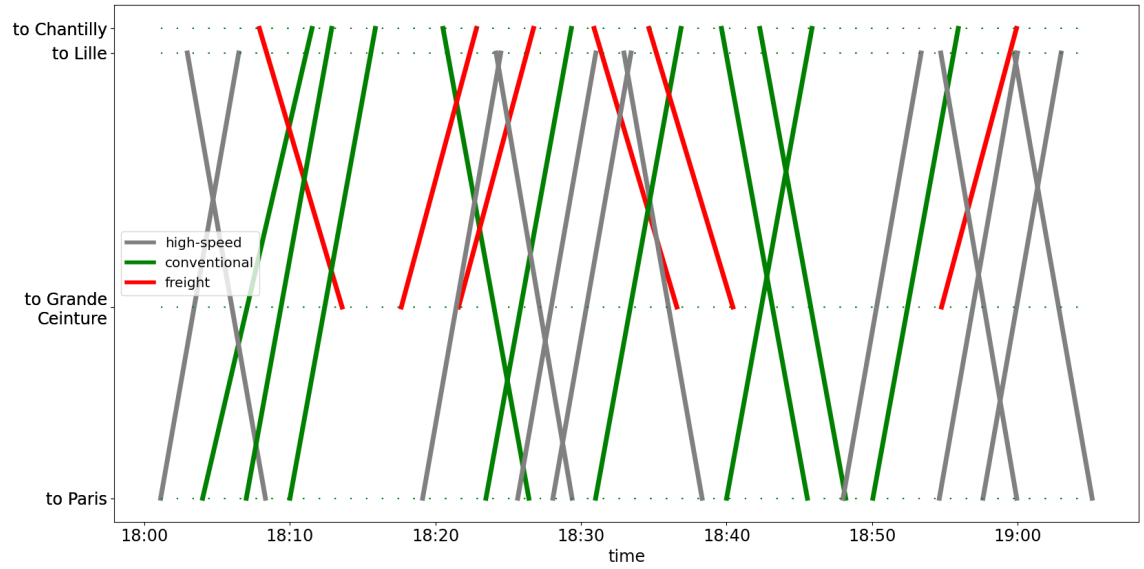
a variety of delay scenarios following the approaches of Lusby et al. (2013) and Pellegrini et al. (2014). These delay scenarios are considered in a one-hour period, i.e., the peak hour from 18:00 to 19:00. Details of the timetable and delays for the one-hour scenarios in both case studies are given in Table 4.3. We note that these timetables are actual timetables and hence not designed for ETCS Level 2 with Onboard TIM.

Table 4.3: Details of the case study scenarios in terms of timetable and delays.

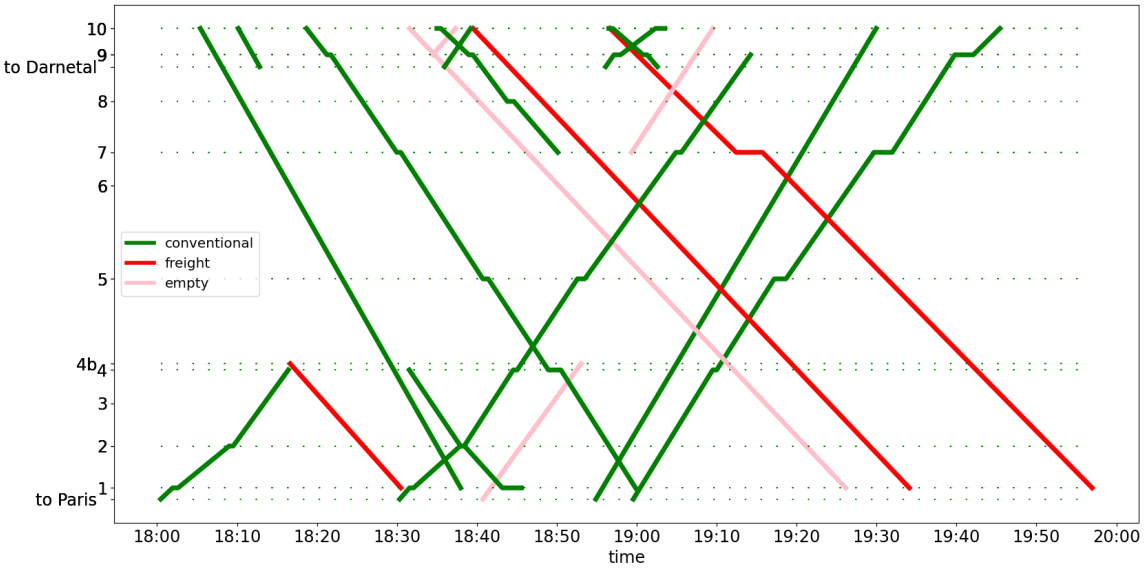
	Junction	**Corridor**
Timetable		
# trains	28	19
- # high-speed trains	12	–
- # conventional trains	10	12
- # freight trains	6	3
- # empty trains	–	4
Scheduled headway	180 / 120 s	210 / 270 s
Buffer time	105 / 40 s	21 / 73 s
Dwell time supplement	–	10 - 30 s
Travel time supplement	60 s	90 - 150 s
Delays		
# delayed trains - median	6	4
# delayed trains - range	3 - 9	1 - 7
Total delay - range	1244 - 5510 s	693 - 4863 s
Mean train delay - range	415 - 663 s	464 - 817 s

In the junction area, 28 trains are scheduled to either enter or start their service between 18:00 and 19:00. Of the 28 trains, there are 12 high-speed, 10 conventional and 6 freight trains. The scheduled paths of these trains are shown in Figure 4.5a. The scheduled headway is generally three minutes, with two-minute headways occurring occasionally. As indicators of the timetable's sensitivity to delays, we report the buffer time, i.e., the minimum time between train blocking times along their shared path, and travel time supplement, i.e., the difference in travel

time through the area between running at maximum speed and at scheduled speed. Buffer times of 105 and 40 seconds are included in the three-minute and two-minute headways, respectively, when evaluated under ETCS Level 2 with Onboard TIM with the original discretisation into TTD sections. The travel time supplement is approximately one minute for all trains.



(a) Train path diagram of timetable considered for Gonesse junction.



(b) Train path diagram of timetable considered for Rosny-StEtienne corridor.

Figure 4.5: Train path diagrams of case study timetables.

In the corridor area, 19 trains are scheduled to either enter or start between 18:00 and 19:00. Of the 19 trains, there are 12 conventional, 3 freight and 4 empty trains. The scheduled paths of these trains are shown in Figure 4.5b. The scheduled headway varies significantly, although 3.5 and 4.5 minutes occur regularly. As for the 'timetable sensitivity indicators' of buffer and supplement times, we report the following. The 3.5-minute and 4.5-minute headways respectively include 21 and 73 seconds of buffer time. The dwell time supplements are 10 to 30 seconds per stop. For the travel time supplement, there is a potential difference in exit time of 1.5 to 2.5 minutes.

The input delays for the considered hour are directly taken from the delay scenarios defined for the daily timetable. As a result, the number of delayed trains show some variation around the imposed 20%. In the junction case study, the number of delayed trains ranges from 3 to 9, with a median of 6. The total train delay ranges from 1244 to 5510 seconds, while the mean train delay varies between 415 and 663 seconds. In the corridor case study, the number of delayed trains ranges from 1 to 7, with a median of 4. The total train delay ranges from 693 to 4863 seconds, while the mean train delay varies between 464 and 817 seconds.

4.4.2 Impact of methodological changes on model size

Before diving into the experimental results, we consider the effects of the methodological changes presented in Section 4.3 on the model size. As a first step, we consider the number of sections in the six considered track discretisations corresponding to different granularities, as a result from the track discretisation procedure of Section 4.3.2. Table 4.4 presents the various number of sections for the two case studies, with the original number of sections (89 and 239, respectively) in the first column. Additionally, Table 4.4 includes the growth rates. A growth rate indicates the increase in number of sections relative to the number of sections in the one step coarser discretisation. With the maximum section length being halved over the discretisations (from 800 to 50 metres), the growth rate is maximum two. A rate of two will, however, not be reached given that both case studies contain switches whose sections are never split.

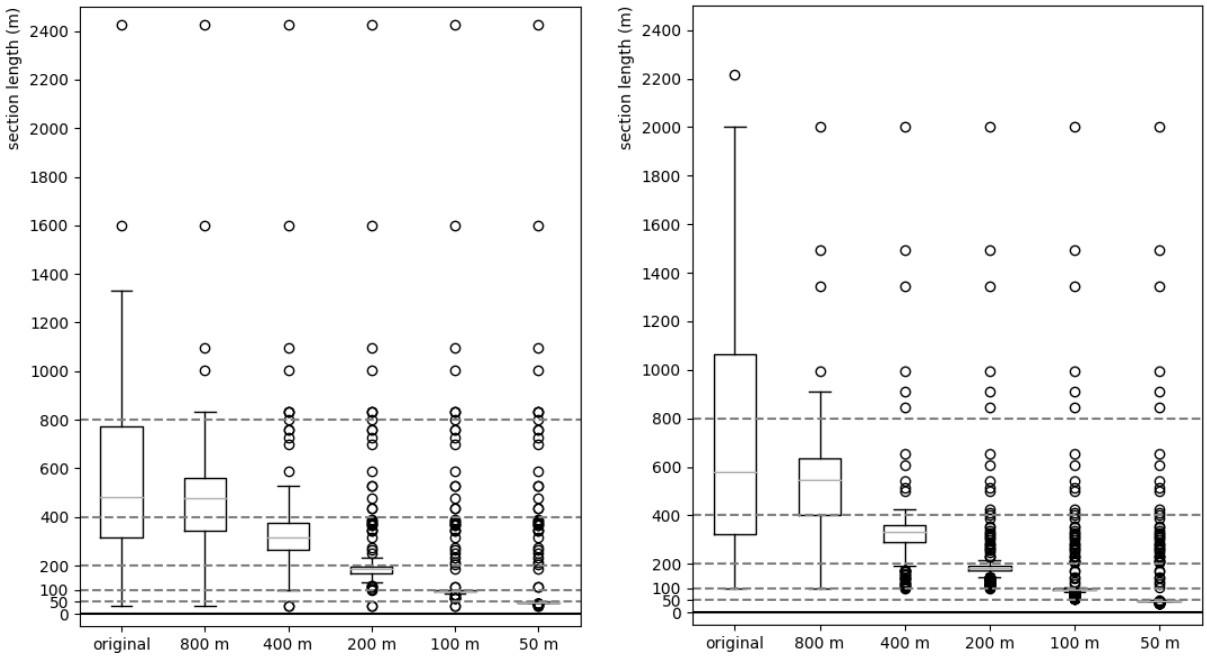
Table 4.4: Number (#) of sections per case study for the different discretisations, represented by their maximum section lengths. In brackets, the growth rate with respect to one step coarser discretisation.

# sections	**original**	**800 m**	**400 m**	**200 m**	**100 m**	**50 m**
Junction	89	101 ($\times$1.13)	135 ($\times$1.34)	203 ($\times$1.50)	345 ($\times$1.70)	630 ($\times$1.83)
Corridor	239	340 ($\times$1.42)	530 ($\times$1.56)	912 ($\times$1.72)	1692 ($\times$1.86)	3252 ($\times$1.92)

The finer the discretisation, the closer the growth factor gets to two as more sections need to be split that have not yet been split due to their original length being shorter than the previously considered maximum length. So, it shows the presence of a significant amount of original sections with a length even shorter than 50 metres. Also, we see a remarkable difference between the two case studies. Not only the original number of sections is much higher in the corridor than in the junction area, but also the growth rate. The difference in growth rate from, for example, the original length to 800 metres shows the relatively high number of sections longer than 800 metres in the corridor. This is not surprising given the characteristics of the other area as junction. That is, the junction area has relatively more switch sections and generally shorter sections compared to the corridor area.

The difference in section lengths between the two case studies is illustrated in Figure 4.6. The distribution of the section lengths is given for the original track discretisation in terms of TTD sections as well as for the rediscretisations. Figure 4.6a and Figure 4.6b clearly demonstrate the wide range of section lengths in the original representation of the junction (35 to 2424 metres) and corridor case study (100 to 2217 metres), respectively. Due to original section lengths averaging between 500 and 600 metres, the median section lengths are practically the

same in the 800-metre discretisation. In the finer discretisations, this median value decreases significantly. The median value stays below the maximum discretisation length, despite not splitting (long) switch sections, which is clearly visible by the spread in the section length beyond the maximum discretisation length. That is due to the limited amount of switch sections, originally short sections, and the fact that by the discretisation procedure, the sections are split into subsections with lengths between half the maximum and the maximum length.



(a) Junction. *(b) Corridor.*

Figure 4.6: Distributions of section lengths for the different discretisations in the case studies.

Related to the section lengths, we include notes on the section length variability and the switch sections. Overall, the variation in section lengths within the discretisations make it less straight forward in terms of drawing conclusions. An alternative approach would be to consider uniform section lengths, though in its turn that would not capture the practicality of having shorter sections close to stations and switches, i.e., (capacity) bottlenecks. Moreover, dividing TTD sections into virtual subsections is in line with the ETCS HTD concept. Regarding switch sections, we initially assumed they would be relatively short and therefore did not partition them further. However, we observed some longer switch sections due to extended TTD sections, in which a switch was included without additional boundary partitioning. Since these sections did not emerge as critical in our results, we opted to retain them as they were in our models. Alternatively, these sections could be partitioned into a pure switch section and adjacent TTD sections, where the latter could be partitioned into parts of maximum lengths like the other TTD sections.

With the increase in the number of sections over the discretisations, the dimensions of the mathematical model are negatively affected. This is clearly visible in Table 4.5, which reports the number of binary variables, continuous variables and constraints of the MILP for the two case studies. While the overall model size increases for finer discretisations, the number of binary variables is stable. The independence of the number of binary variables on the discretisation is explained by the binary model variables being the switch turning indicator variables

as addition to the 'original' ordering, routing and speed variables, which are (in terms of sections) defined for either a switch or an 'open line stretch', see also Versluis et al. (2025a). The increase in numbers of continuous variables and constraints across the discretisations is close to the growth rate of the number of sections as reported in Table 4.4.

Table 4.5: Number (#) of binary and continuous variables, and constraints of model initialised without delays for the different discretisations.

	Versluis et al. (2025a)	**This work**				
		original	800 m	400 m	200 m	100 m	50 m
Junction							
# binary variables	6186	12857	12857	12857	12857	12857	12857
# continuous variables	28844	28844	32544	42202	62030	102346	184221
# constraints	81407	94749	105567	132452	189982	304094	547369
Corridor							
# binary variables	4649	5275	5275	5275	5275	5275	5275
# continuous variables	68902	68902	97228	151210	258569	480938	923402
# constraints	113214	114466	159643	246042	420732	782561	1502517

In addition to showing the trends in model dimensions across discretisations, Table 4.5 highlights the impact of the reformulation of the switch blocking time constraints on the model size. This impact is assessed by comparing the model dimensions using the old and new formulation for the same discretisation. Specifically, we compare the model dimensions from Versluis et al. (2025a) with those in this work considering the original section lengths. This comparison tells us that the reformulation leads to a significant increase in the number of binary variables and constraints, while the number of continuous variables are not affected. This is in line with the addition of the switch turning indicator variables and their 'value setting constraints'. Remarkable is the growth rate of the number of binary variables in the junction case study. Indeed, the number of switches is proportional significantly larger (38 out of 89 TTD sections versus 44 out of 239 TTD sections), but the increase in number of constraints is comparable for the two case studies. We do note that in the presolve phase, the model dimensions are already significantly cut down by CPLEX.

4.4.3 Analysis of conflict detection and resolution model for ETCS Level 2 with Onboard TIM

The experiments are performed on the two case studies described in Section 4.4.1, considering two aspects. First, the effects of the model adjustments to the earlier presented conflict detection and resolution model are analysed. We compare the (rescheduling) results obtained by, on the one hand, the model for ETCS Level 2 with Onboard TIM as presented in Section 4.3 with, on the other hand, the model describing ETCS Level 2 with TTD, as presented by Versluis et al. (2025a). For this comparison, we fix the track discretisation to correspond to the TTD sections in place.

Second, we assess the impact of the track discretisation on the enhanced conflict detection and resolution model for ETCS Level 2 with Onboard TIM. For the junction and

the corridor case study, the sensitivity analysis is carried out in two steps. In the first step, the model is run to find optimised (rescheduling) solutions for the 25 delay scenarios in each of the six discretisations. The solutions are evaluated (and compared) in terms of objective value, computation time, optimality gap and rescheduling decisions. In the second step, the optimised solutions are cross-evaluated. That is, the solution obtained with a certain discretisation is evaluated given a different discretisation.

The experiments are run on an Intel(R) Xeon(R) CPU Gold 6226R CPU @ 2.90GHz, 16 cores, 256 GB RAM. The implementation uses IBM ILOG CPLEX Concert Technology for C++, version 20.1. For the optimisation, the computation time limit is set to 3600 seconds. With this, the time limit does not correspond to a real-time application, but it allows for a thorough evaluation of the model. For an idea on the real-time applicability, we refer to Versluis et al. (2025a). For the earlier model version for ETCS Level 2 with TTD, Versluis et al. (2025a) conclude that despite the difficulties in obtaining optimally proved solutions, the real-time performance is acceptable. Given that the latter is due to the short computation time needed to find high-quality solutions, also for the larger model presented in this chapter, reasonable real-time performance can be expected. However, in general, more efficient solution algorithms and/or decomposition need to be considered.

Effects of modelling adjustments on optimisation solutions

In this chapter, the fixed-block distance-to-go conflict detection and resolution model of Versluis et al. (2025a) is adjusted to describe ETCS Level 2 with Onboard TIM. The two adjustments introduced in Sections 4.3.1 and 4.3.3 are directly related to the minimum train separation, being the incorporation of onboard TIM parameter values (such as the GNSS-based train position report update period) and the reformulation of the switch blocking time constraints. With these adjustments, the minimum train separation is modelled differently. On the open line, the train blocking times are shorter. At the switches, however, the train blocking times have increased due to the explicit consideration of the switch setup time. This was needed to differentiate between the route relations to account for general interlocking rules.

In Table 4.5, we have seen that the new constraints lead to serious increases in the model dimensions, specifically for the junction case study. From the overview of the (mean) results of the models with and without the model adjustments in Table 4.6, it is clear that the larger model size is directly reflected in the optimisation process. The computation time used on average increases from 782 to 3468 seconds for the junction and from 377 to 1136 seconds for the corridor. The increase is significantly larger for the junction, for which also the model size increase was a lot larger. Without the adjustment, the model was able to solve all but two scenarios to optimality within the computation time limit. With the adjustments, only two were solved to optimality. Despite the poor performance in terms of 'optimality convergence', the obtained solutions are of (relatively) good quality - with a maximum optimality gap of 1.01% and a mean gap of 0.22%. We end up with a minimal drop in mean objective value, from 6866 to 6861, for the junction area. Underlying, we have that in about 50% of the cases, the ETCS Level 2 with TTD model gives the best objective value and in the other half of the cases the ETCS Level 2 with Onboard TIM model does so. The fact that without the model adjustments a better solution can be obtained is due to the addition of the switch setup time. We do remark that the speed component of the objective value, counting the number of times a maximum speed

profile is assigned to a train between switches/on an open line track, is larger in the adjusted model. This means that in the total delay of the two solutions, there is some more difference, in favour of the enhanced model.

The corridor case study has similar results, but there are also differences. As a result of the less increased model dimensions, the corridor case study is solved to optimality for the same number of scenarios in the two model versions (19/25). Despite the longer mean computation time, the mean optimality gap is lower (0.01% versus 0.02%). This does not help, as we end up with a mean objective value that is higher for the enhanced model. Here too, there is a higher speed penalty, but also the total delay alone is higher, though minimally. Apparently, the effects of the additional switch setup time is larger here than in the junction. There are more switches, though less relative to the number of sections in the case study.

More relevant is the impact to the modelling of the conflict detection and resolution. We have seen that we get different results for the model versions with and without the adjustments. For conflict detection and resolution, different results would mean different rescheduling decisions. So we also looked into the solutions themselves. For the case study's scenarios, we compared the two solutions in terms of reordering and rerouting decisions.

For the junction scenarios, no solution was the same. At least, one different route was assigned. The number of different routes goes up to 11, while only in one scenario one different order was proposed. Also for the corridor scenarios, only one time one different order was proposed. Differently from the junction, in nine scenarios, the same solution was proposed. In the other cases, one to seven different routes were found.

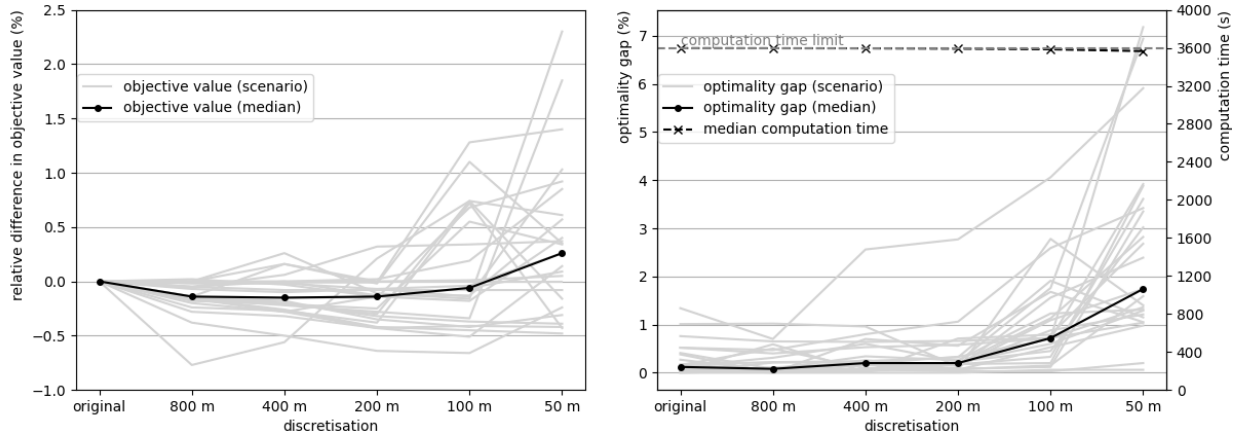
Table 4.6: Mean results of model optimisation with (w/) and without (w/o) model adjustments.

	Junction		**Corridor**	
	w/	w/o	w/	w/o
Objective value	6861	6866	4764	4753
Speed penalty	62	46	16	9
Computation time (s)	3468	782	1136	377
Optimality gap (%)	0.22	0.01	0.01	0.02
# optimal solutions	2	23	19	19

Impact of track discretisation in the junction case study

In this section, we present and discuss the experimental results for the junction case study. Two key indicators for the optimisation step are the objective value and the optimality gap obtained within the maximum computation time of 3600 seconds. The trends in objective values and optimality gaps across the solutions obtained with the different discretisations are illustrated in Figure 4.7. In Figure 4.7a, the objective values of the 25 scenarios are reported relative to the objective value obtained with the original discretisation, complemented with the median relative objective value. As the model describes a minimisation problem, negative percentages indicate an improvement, while positive percentages indicate worse objective values. On the one hand, Figure 4.7a tells us that, for the most part, the discretisations of 800, 400, 200 and 100 metres lead to an improvement over the original discretisation. On the other hand, the 50-metre discretisation often performs worse. Looking at Figure 4.7a, the number of scenarios

with a worse performance is already significant for the 100-metre discretisation. Given that the shorter the sections the shorter the separation distance, the reason for the worse performance is that the increased model dimensions result in less optimised solutions. As Figure 4.7b illustrates, the median optimality gaps after the maximum computation time are significant for the two finest discretisations, while the median optimality gaps for the coarser discretisations are relatively close to 0%. However, Figure 4.7b also shows that the median computation time used is close to the computation time limit from the original discretisation on. Only two scenarios are solved to optimality in the original discretisation. For the 800-, 400-, 200-, 100- and 50-metre discretisations, respectively three, one, one, zero and zero obtained solutions are proven to be optimal.



(a) Objective values relative to original ones. *(b) Optimality gaps and computation times.*

Figure 4.7: Results of the junction case study: trends in objective value and optimality gap across discretisations.

The relatively poor quality of the solutions of the 100- and 50-metre discretisation is also clear in Figure 4.8, in which the absolute objective values of the cross-evaluation are plotted. That is, the means of the objective values obtained by evaluating a solution optimised for one discretisation in another discretisation. While the lines in the top of the plot correspond to the 100- and 50-metre solutions evaluated with the various maximum section lengths, it is the 800-metre solution which results in the bottom line. Hence, typically, the 800-metre solution is the best throughout the discretisations. Given the small difference in objective value, this is mainly a result from the suboptimality in the case of finer discretisations. Figure 4.8 illustrates the trend in objective value across the evaluations of the different solutions in the different discretisations. The overall trend is clear: the finer the discretisation, the lower the objective value of a solution. We note that, for the junction area, the maximum difference in objective values between evaluations of a solution is 21 seconds, corresponding to a minimal 0.3%. The gain in objective value for a specific solution purely indicates the effect of the shorter train separation due to the shorter sections in a finer discretisation.

At the level of individual scenarios, we do not necessarily see the same trend in objective value of the solution cross-evaluations. Occasionally, a best solution for one discretisation does not yield the minimum objective value observed for another discretisation. In the junction case study, this occurs in seven of the 25 scenarios. Typically in these cases, for the original discretisation a unique best solution is found which performs worse in one or more finer discretisations compared to their optimised solution. We note that this concerns the 800, 400 and 200 metres

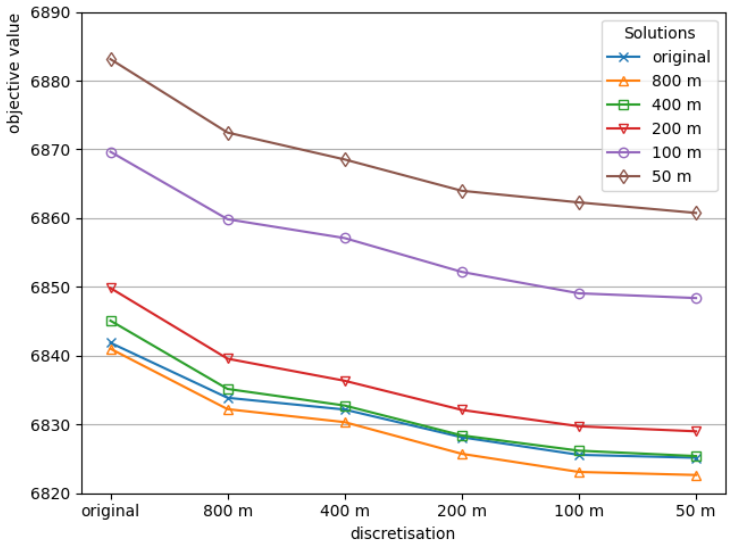


Figure 4.8: Mean objective values of optimised solutions evaluated in all discretisations.

since the 100- and 50-metre solutions have too large optimality gaps. In turn, the solutions optimised for these finer discretisations give a higher objective value than the optimised one in the original discretisation. These differences, both in the original and the finer discretisations, are a matter of a couple of seconds.

Despite these differences being minimal, there is something significant behind it. Underlying these differences are different rescheduling decisions. For the seven scenarios in the junction case study, the differences in decisions are three to nine route assignments. There is an exception, for one scenario there is a different order decided in the exclusively best solutions. That exceptional scenario is also the only one for which the best solution for the original discretisation is not its own solution; the two best solutions are from the 800 and 200 metre discretisations. This is possible due to the overall slightly better performance of the 800-metre compared to the original discretisation, which is illustrated by the lowest objective values in the cross-evaluation (Figure 4.8) as well as the lower median computation time and optimality gap (Figure 4.7b).

Concerning the different route assignments, they are sometimes simply the result of a shorter or faster route. In other cases, it is inferred from the interaction between trains. For example, the first train of a conflicting train pair can be detoured in order to allow a free track for the second train. This can go together with a 'maximum speed assignment'. Typically, the two routing solutions have more effect in the original discretisation than in the finer one. Some of the trains are not affected by the difference in routes in the finer discretisation, while they are in the original discretisation. Note that there is always at least one train for which it does matter in the finer discretisation, otherwise the scenario would not be considered here. The potentials of finer discretisations in terms of conflict detection and resolution are better illustrated by the different train ordering observed. It is an exemplary result of what shorter train separation through shorter sections (in combination with the switch blocking times) can lead to.

Overall, we can conclude that in the considered junction case study, the impact of track discretisation granularity on conflict detection and resolution is limited. This outcome is primarily due to the model's complexity as a result from the model formulation and the traffic density.

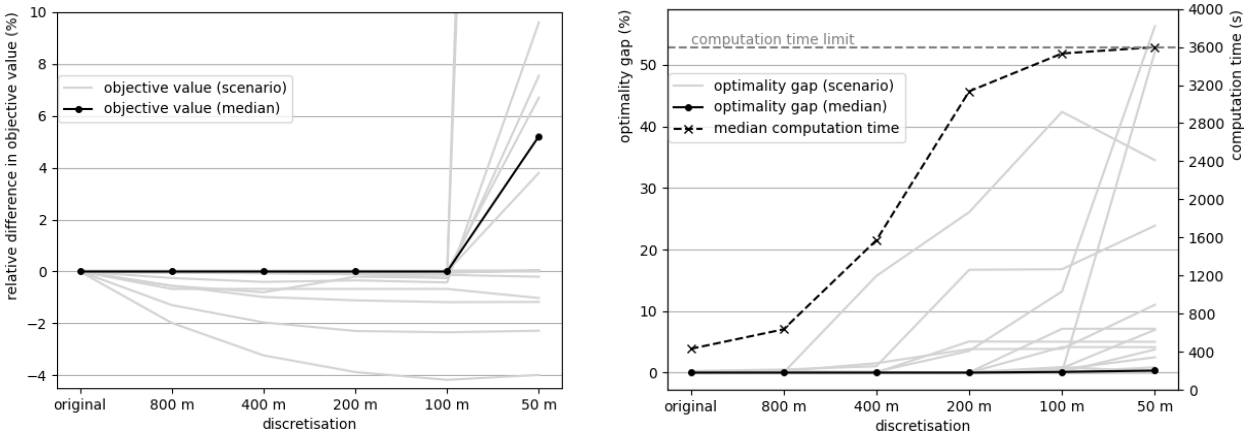
Additionally, the track layout with relatively many switches reduces the effects. While the overall impact is minimal, there are clear indications of the potential impact of track discretisation granularity on conflict detection and resolution. Notably, the original discretisation into TTD sections does not have the overall best performance in balancing solution quality and computational complexity. Instead, the discretisation with a maximum length of 800 metres (and a median length of 500 metres) performs better. Looking further, the 200-metre discretisation can be a viable option. Due to the high complexity, its performance is unstable, but promising in terms of objective value. With some (more) computational improvement, the 100-metre option might also become interesting as there is still some gain to be obtained from 200 to 100 metres (see cross-evaluation). For a better assessment, it is essential to reduce the computational complexity of the conflict detection and resolution model.

Impact of track discretisation in the corridor case study

For the corridor case study, the trends in objective value and optimality gap across the solutions obtained with the different discretisations are illustrated in Figure 4.9, with the objective values relative to the objective value obtained with the original discretisation in Figure 4.9a and the optimality gaps and used computation times in Figure 4.9b.

Aside from the 50-metre discretisation, the median objective values across the discretisations are the same, as a consequence of that being the case for 16 of the 25 scenarios. The higher median objective value for the 50-metre discretisation is due to bad performance of one specific scenario with an above-median value in the 50-metre discretisation while having below-median values in the other discretisations.

Consequently, the median optimality gap is practically stable at 0% across the discretisations. Only for the 100-metre and 50-metre discretisations, the median optimality gap increases very slightly. The optimality gap of some individual scenarios starts to increase more significantly from the 400-metre discretisation on. This goes hand-in-hand with the rise in the median computation time used, which increases from the 400-metre discretisation to become significantly higher for the 200-, 100- and 50-metre discretisations.



(a) Objective values relative to original ones. *(b) Optimality gaps and computation times.*

Figure 4.9: Results of the corridor case study: trends in objective value and optimality gap across discretisations.

Figure 4.9a shows that for all the scenarios with differing objective values across the discretisations, the objective values actually show a decreasing trend except for the 50-metre discretisation (up to -4.17%). Hence, for this case study, we see that despite the increasing computational complexity, the results of a discretisation are at least as good as those of the coarser discretisations (with the exception of the 50-metre discretisation). Few other exceptions can be found, for example when going from 400 to 200 metres. These are the results of suboptimality of the solution obtained for the finer discretisation, as they indeed correspond to the lines in Figure 4.9b with the increasing optimality gaps.

As Figure 4.8 did for the junction case study, Figure 4.10 shows the results of the cross-evaluation for the corridor case study. By plotting the mean of the absolute objective values from evaluating the solutions optimised using the different discretisations in the other discretisations, we can (better) assess the 'overall' quality of the obtained solutions. The plot shows three groups of trend lines. The first group consists of the original, the 800- and the 400-metre solution, the second one contains the 200- and the 100-metre solution, and the third corresponds to the 50-metre solution. The three groups describe very similar trends, i.e., the declining decrease in objective value due to shorter train separation when considering finer track discretisations. The difference between them lies in the exact values. The second and third group obtain values slightly (ca. 0.1%) or significantly (ca. 5.5%) higher, respectively, compared to the first group due to the higher suboptimality in the optimisation step for the finer discretisations, visible by both the increased median computation time and the higher number of positive optimality gaps in Figure 4.9b. The limited number of positive optimality gaps for the coarser discretisations results in minimal values in the cross-evaluation. The offset between the groups is the result of a limited number of scenarios for which the final solution for the finer discretisation(s) is worse than the objective value of the evaluation of the other solutions. Within the first group, the 800-metre solution is the best by a minimal difference overall. However, for the evaluation in the 400-metre discretisation, the 400-metre solution proves to be best. Within the second group, it does not hold that the solutions have the lowest objective in their respective discretisation, again due to the couple of scenarios with a worse final objective value. The third group of the 50-metre discretisation, does not play a role due to the frequent and significant suboptimality.

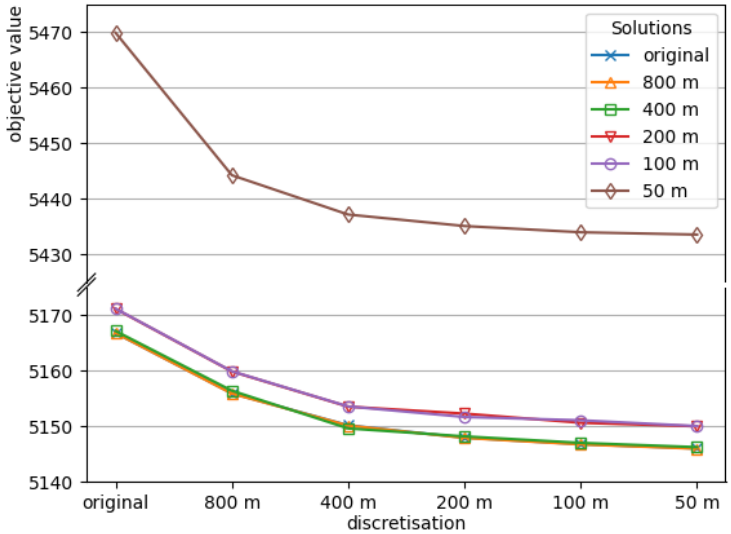


Figure 4.10: Mean objective values of optimised solutions evaluated in all discretisations.

For the corridor case study, there is less to say on the level of individual scenarios than for the junction case study. As already said, for 16 of the 25 scenarios, the same objective values are found across the different discretisations. In six of these 16 scenarios, the same rescheduling decisions are proposed; in the remaining ten, alternative solutions of the same quality are found.

There is a second group for which either the same solutions or solutions of the same quality are found. For the six scenarios in this group, the same rescheduling decisions result in different objective values due to the shorter train separation enabled by a finer track discretisation. The reduction in objective value reaches up to 4.17%.

This leaves three scenarios to be discussed. For these scenarios, both different objective values and significantly different solutions have been obtained. Unfortunately, in all three cases it holds that the objective value increases for a finer discretisation, meaning that the different solution is a suboptimal solution as a consequence of having a positive final optimality gap.

Given the above classification of the scenarios based on their solutions, we have to conclude that for the corridor case study the impact of the track discretisation on conflict detection and resolution is non-apparent in terms of rescheduling decisions. With the mean decrease in objective values over the discretisations being similar to that of the junction case study, we identify track layout and traffic density as indicators for the impact of track discretisation on conflict detection and resolution.

4.4.4 Discussion of rescheduling decisions for different discretisations

In line with the focus on the effects of track discretisation on rescheduling decisions, we conclude the experimental section with a discussion of how these decisions differ across the considered discretisations for our implementation of ETCS Level 2 with virtual subsections. In Table 4.7, we provide an overview of the number of proposed reordering and rerouting decisions per discretisation that are different with respect to the decisions proposed for the next coarser one. So, for the 800-metre discretisation the comparison is made with the original discretisation, while for the finer discretisations, i.e., 400, 200, 100 and 50 metre, it is made with the 800-, 400-, 200- and 100-metre discretisation, respectively. Additionally, Table 4.7 reports the number of scenarios for which the solution was proven optimal, as well as the number of scenarios for which the solution of the considered discretisation was equally good or better in terms of objective value compared to the solution of the next coarser one (when evaluated for the considered discretisation).

A first observation is that the corridor case study appears quite stable with respect to rescheduling decisions. This is evident from having only one scenario for one discretisation with an improved objective value. Furthermore, the number of corridor scenarios with a proven optimal solution is relatively high (up to 19 of 25 compared to up to 3 of 25 for the junction case study), while the number of scenarios with solutions resulting in a worse objective value is relatively low (from 0 to 9 of 25 compared to from 3 to 16 of 25 for the junction case study). In contrast, the junction case study seems more sensitive to discretisation changes. For each discretisation, there are three to seven scenarios for which a better objective value was found due to different decisions compared to the next coarser discretisation.

A second observation is that the number of different routing decisions gives an indication rather than a complete picture of the impact. Not all different route assignments are critical, i.e.,

Table 4.7: Number of differing order and route decisions (# Δ o & r) per discretisation compared to the next coarser one, along with counts of scenarios with proven optimal solution and same or improved objective values.

	# scenarios with proven optimal sol.	**Same objective value**			**Improved objective value**		
		# scenarios	# Δ o & r decisions median	# Δ o & r decisions mean	# scenarios	# Δ o & r decisions median	# Δ o & r decisions mean
Junction							
800 (vs orig)	3/25	15/25	0 & 3.5	0 & 3.4	7/25	0 & 7	0 & 6.6
400 (vs 800)	1/25	16/25	0 & 4	0 & 3.6	3/25	0 & 6	0.3 & 6.7
200 (vs 400)	1/25	11/25	0 & 3	0 & 3.4	6/25	0 & 7.5	0 & 7.4
100 (vs 200)	0/25	7/25	0 & 4	0 & 4.8	3/25	0 & 3	0 & 3.7
50 (vs 100)	0/25	3/25	0 & 6	0 & 4.3	6/25	1 & 3.5	0.8 & 4.8
Corridor							
800 (vs orig)	19/25	24/25	0 & 2	0 & 2.5	1/25	1 & 8	1 & 8
400 (vs 800)	15/25	24/25	0 & 4	0 & 3.4	0/25	–	–
200 (vs 400)	14/25	23/25	0 & 3	0 & 2.9	0/25	–	–
100 (vs 200)	6/25	23/25	0 & 3	0 & 2.8	0/25	–	–
50 (vs 100)	0/25	16/25	0 & 0.5	0 & 1.7	0/25	–	–

contributing to the improvement of the objective value. This is illustrated by the number of different routes assigned in scenarios with solutions given the same objective value. In the junction and corridor case studies, respectively circa four and three of the route assignments seem to be interchangeable with respect to total train delay. Considering rescheduling practice, in which there is a general reluctance to apply (local) rerouting, introducing a penalty for assigning an alternative route can be included to only allow rerouting if it results in some delay recovery.

Overall, we conclude that for the corridor case study, the 800-metre discretisation performs best, as it has a non-negative and even some positive impact on the quality of the solution in terms of rescheduling decisions. While finer discretisations are unlikely to provide further benefits in this context, they also rarely lead to worse outcomes. For the junction case study, the impact of discretisation is less straightforward. Considering both the number of scenarios with an improved objective value and those with a worsened objective value, we conclude that the 800-, 400- or the 200-metre discretisations are all valid options.

4.5 Conclusions

In this chapter, we assessed the impact of the track discretisation granularity on conflict detection and resolution for next-generation distance-to-go signalling. The assessment considers both the mathematical optimisation model as such and the computational results obtained with it. A particular focus is on the impact on rescheduling decisions included in the model solutions.

The considered conflict detection and resolution model is based on an earlier developed model for fixed-block distance-to-go signalling. The model is updated from European Train Control System (ETCS) Level 2 with Trackside Train Detection (TTD) to ETCS Level 2 with Onboard Train Integrity Monitoring (TIM) by considering onboard TIM parameters and intro-

ducing a track (re)discretisation procedure. Additionally, as a general update of the model, a reformulation of train separation at switches is considered.

The impact assessment is conducted on two case studies featuring different types of control areas: a heavily trafficked complex junction and a less densely trafficked corridor. The discretisations considered are derived from the original track partitioning into TTD sections and are characterised by maximum section lengths in the range from 50 to 800 metres. These (virtual) section lengths are in line with those commonly considered in research on ETCS Hybrid Train Detection, which currently has the industry's focus.

Over the discretisations, from coarse to fine, the dimensions of the mixed integer linear programming (MILP) model grow linearly with the number of sections. The main increase in model dimensions is, however, due to the reformulation of the train separation constraints at switches. Given the relatively large model dimensions and the exploration phase of the research, a computation time limit of one hour is set. Nevertheless, for the finer discretisations and the junction case study in general, the model is not solved to optimality in most cases.

Regardless of the more frequent suboptimality for the finer discretisations, the corridor case study practically features non-increasing objective values (up to -4.17%). They are, however, solely due to shorter train separation distances as a result from the shorter sections. Any effects of the track discretisation on the underlying rescheduling solutions is non-apparent. The results from the junction case study show some effects of the track discretisation on conflict detection and resolution. Though by far the most benefit in objective value between solutions is due to the shorter train separation resulting from the shorter sections, in some cases (delay scenarios) different rescheduling decisions proved to be better for different discretisations. Mostly, the different decisions concern a limited number of route assignments, but occasionally also a train ordering. These are enabled by the short separation in general, enforced by the new separation constraints at switches.

The reported results indicate maximum section lengths of 400 or 200 metres as possible thresholds in terms of the balance between solution quality and complexity. Focusing on the rescheduling decisions, the 800-metre discretisation already performs well, providing stable and reliable solutions, particularly for the corridor case study. In general, the appropriate discretisation threshold strongly depends on the case study and its track layout. Hence, a next step is to look into case study characteristics such as track layout, traffic density and delays to identify possible predictors for the effects of track discretisation on conflict detection and resolution.

On the way to practical implementation of conflict detection and resolution models for next-generation distance-to-go signalling systems, further research is needed. For example, into the type of track discretisation; whether uniform sections lengths or 'location-specific' lengths would result in a better balance in computational complexity and solution quality. Furthermore, it would be interesting to assess the impact of track discretisation on other conflict detection and resolution algorithms, both within and beyond MILP-based approaches. Specially, it is recommended to look into alternative methods for the implementation of the problem, especially with an eye on the real-time nature of the problem.

Chapter 5

Early-warning hazard prediction for railway traffic management under radio-based distance-to-go signalling with onboard train integrity monitoring

---

Chapter 4 considered a conflict detection and resolution model for radio-based distance-to-go railway signalling systems with onboard train integrity monitoring (TIM). For implementing onboard TIM, satellite-based technologies are a promising solution. However, degradation in satellite signals might lead to hazardous traffic conditions that require specific operational management.

This chapter presents an early-warning hazard prediction framework to support railway traffic management in mitigating hazardous traffic conditions arising from radio and satellite signal degradation. The core railway traffic management framework is enhanced by integrating real-time monitoring and prediction of onboard train position and TIM parameters including radio communication delays and satellite positioning errors.

Apart from minor changes, this chapter has been accepted for publication as:

Versluis, N. D., Quaglietta, E., Garcia-Fernandez, M., & Goverde, R. M. P. (in press). Early-warning hazard prediction for railway traffic management under radio-based signalling with satellite-based train integrity monitoring. *International Journal of Transportation Science and Technology*.

5.1 Introduction

To fulfil the future need for higher railway capacity as a consequence of the forecasted increase in railway demand (European Environment Agency, 2021), traditional trackside functions are being moved from the trackside to onboard systems, such as train positioning and train integrity monitoring (TIM). TIM is a safety-critical function responsible for monitoring the physical integrity of the train, i.e., checking that all cars remain properly coupled during operations. This requires continuous and reliable monitoring of the actual train length by means of accurate measurements of both the front and rear position of the train. Traditionally, the role of TIM is fulfilled by trackside train detection systems such as axle counters or track circuits. However, migrating TIM to onboard systems reduces the reliance on trackside train detection as in fixed-block signalling systems, and enables the deployment of higher-capacity radio-based distance-to-go signalling systems with virtual or moving blocks.

Various wired and wireless solutions for onboard TIM have been explored in international research initiatives, such as Shift2Rail projects X2Rail-2 (2020) and X2Rail-4 (2022). In particular, satellite-based TIM has been investigated for its potential to enhance train integrity verification, as demonstrated in Shift2Rail project PERFORMINGRAIL (Shift2Rail, 2023) and Europe's Rail project R2DATO (Mayolle et al., 2025). Within the context of the (radio-based) European Train Control System (ETCS), it is proposed to equip trains with Global Navigation Satellite System (GNSS) receivers at both the front and rear end to allow for automatic train length estimation, typically updated every one to two seconds.

The implementation of satellite-based TIM introduces dependencies on the quality and availability of satellite signals, which then become critical factors for safe and efficient train operations under advanced radio-based signalling such as ETCS Level 2 Virtual Block, ETCS Level 2 Moving Block or ETCS Hybrid Train Detection (European Rail Supply Industry Association, 2025). Signal degradation due to environmental conditions e.g., tunnels, urban or naturel canyons (PERFORMINGRAIL, 2022a), may compromise the reliability of train length estimation. Consequently, it is necessary to identify such potential hazards and asses their impact on railway safety and operational performance.

Besides GNSS, radio communication systems are another potential source of safety hazards. Within the context of ETCS, this relates to the currently incorporated Global System for Mobile Communication - Railways (GSM-R) or the Future Railway Mobile Communication System (FRMCS) as the intended successor of GSM-R (International Union of Railways, 2024).

With railway traffic management in charge of supporting the execution of railway operations, both under nominal and disturbed conditions, human dispatchers will play a crucial role in mitigating these risks. To support dispatchers in managing potentially hazardous traffic conditions under radio-based distance-to-go signalling with satellite-based TIM, we present a non-vital early-warning hazard prediction framework, with 'hazard' referring to conditions that could negatively affect performance, such as by increasing uncertainty in train length estimation.

Our proposed framework for railway traffic management under radio-based distance-to-go signalling with satellite-based TIM enhances the core railway traffic management functionalities including traffic state monitoring and prediction, and conflict detection and resolution. We extend these with a tripartite early-warning module. In the first part, real-time values of onboard train position and TIM parameters such as radio communication delays and GNSS positioning

errors are collected. In the second part, future values of the same parameters are predicted for scheduled train runs. If any predicted value exceeds its critical threshold, defined as a value beyond which performance may be impacted, then it is sent to the conflict detection and, if needed, conflict resolution module. In the third and last part, the predicted parameter values are combined with the detected conflicts, if any, to generate an early warning to support the dispatcher in reviewing the updated traffic plan proposed by the conflict resolution module.

The chapter's main contributions are:

- An early-warning hazard prediction module for radio-based distance-to-go signalling with satellite-based train integrity monitoring, integrated into the railway traffic management framework.
- A method for the (monitoring and) prediction of positioning errors in case of GNSS hazards.
- An illustration of the proposed early-warning framework through a real-world use case of degraded GNSS and radio signals due to atmospheric disturbances.

The chapter is organised as follows. In Section 5.2, relevant aspects of (railway traffic management under) radio-based ETCS with GNSS-based TIM are introduced. Section 5.3 presents the early-warning hazard prediction framework for railway traffic management, including a method for predicting train positioning errors in case of GNSS hazards. Section 5.4 illustrates the framework through a case study for a representative use case. Finally, Section 5.5 concludes the chapter.

5.2 Radio-based ETCS with GNSS-based TIM

Within the Shift2Rail research programme of the European Commission, the radio-based distance-to-go signalling system with satellite-based TIM considered most is ETCS Level 2, formerly 3, Moving Block. This holds in particular for projects within Innovation Programme 2 (Advanced Traffic Management and Control Systems) such as the X2Rail projects and PERFORMINGRAIL (Shift2Rail, 2024), which consider GNSS-based TIM for radio-based ETCS.

This section introduces aspects which are relevant for railway traffic management under radio-based ETCS with GNSS-based TIM. First, Section 5.2.1 and Section 5.2.2 introduce train position reporting and GNSS-based TIM for radio-based ETCS, respectively. Then, Section 5.2.3 presents GNSS hazards. Section 5.2.4 concludes with railway traffic management under radio-based ETCS with GNSS-based TIM.

5.2.1 Train position reporting in radio-based ETCS

In radio-based ETCS, i.e., ETCS Level 2, continuous train supervision is maintained through radio (GSM-R/FRMCS) communication between the train and the Radio Block Centre (RBC). Train movement authorities (MAs) are sent from the RBC to the trains, indicating the location to which the train can safely move under supervision. The provision of MAs is based on train position reports (TPRs) sent from the trains to the RBC, indicating the safe position and integrity status of the trains. More specifically, as shown in Figure 5.1, the TPR includes the

estimated front-end position, the confidence interval, the direction, the estimated speed and the train integrity information (European Union Agency for Railways, 2016).

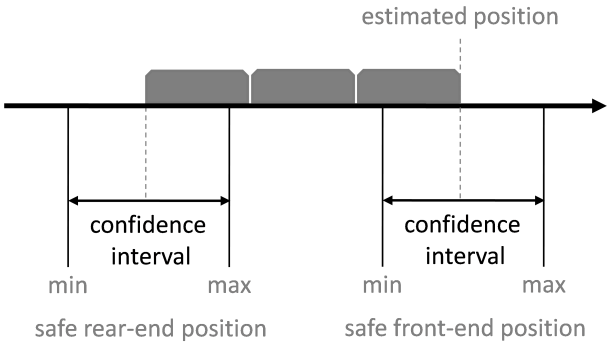


Figure 5.1: Safe train positions for train position report.

The main train position is the estimated location of the train's front end, which is considered within a confidence interval. This interval is bounded by the minimum and maximum safe front-end positions. The safe front-end position can be determined using an onboard train odometer and/or satellite signals, depending on the specific localisation function adopted within the onboard system (PERFORMINGRAIL, 2022a). When using an odometer onboard, the train's position is determined by measuring the distance travelled, usually via wheel rotation sensors. For the safe positions, the odometer measurement of the train position is considered to be affected by a linear error of 5% from the last passed (Euro)balise. These trackside balises have the passive function of recalibrating the train's onboard odometer with the absolute location reference coordinates.

From the safe front-end positions, the safe rear-end positions can be determined. The minimum and maximum safe rear-end positions are obtained by subtracting the train length from the respective front-end positions. That is, at the time that the train's integrity is confirmed and updated. For radio-based distance-to-go signalling systems with reduced or no trackside train detection such as ETCS Level 2 Virtual Block, ETCS Level 2 Moving Block and ETCS Hybrid Train Detection, the train integrity information is provided by onboard TIM. Given its safety-critical role, the integrity information provided by onboard TIM must be highly reliable and accurate.

5.2.2 GNSS-based TIM for radio-based ETCS

Within the PERFORMINGRAIL projects, GNSS-based TIM for radio-based ETCS is assumed. PERFORMINGRAIL (2021) proposes the placement of GNSS receivers at both the front and the rear of each train, as shown in Figure 5.2. The baseline between the two receivers can be computed in real time and at a high rate, position estimated within one second or faster, with an accuracy of a few centimetres. This will allow close monitoring of the baseline distance between the ends of the train. Assuming that the nominal length of the train is known for a given trip, if this estimated baseline exceeds a certain threshold, a warning or alarm can be issued - as indicator that the integrity of the train has been broken.

For the provision of TIM using GNSS, a critical aspect is not only the baseline estimate but also the estimated error of the baseline estimate. Providing an accurate positioning error esti-

mate is a crucial function of GNSS receivers, as it allows the inclusion of a confidence interval in the TPR (see Section 5.2.1). Usually, the positioning error estimate is determined as the formal error of the positioning engine, estimated through the post-fit covariance filter or through the analysis of the post-fit residuals generated by the positioning engine (PERFORMINGRAIL, 2023).

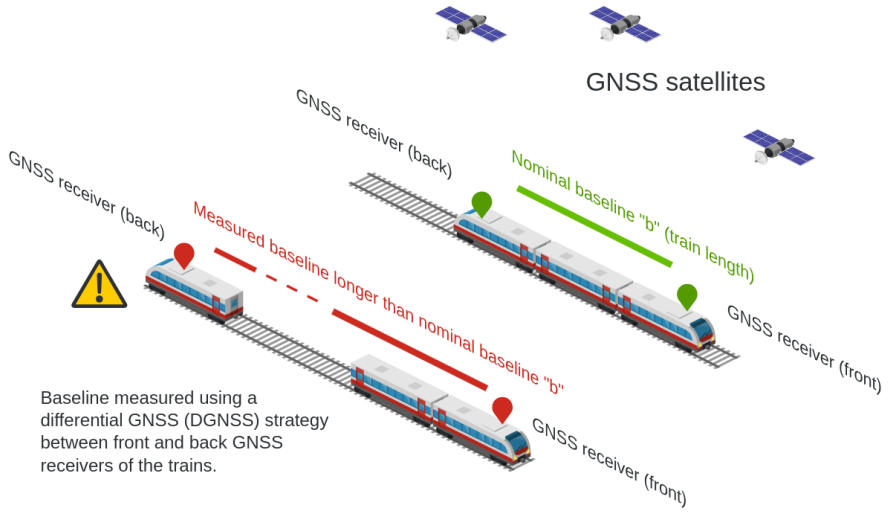


Figure 5.2: Diagram illustrating the working principle of GNSS-based TIM.

5.2.3 GNSS hazards

Within the X2Rail-3 project, several critical hazardous traffic conditions associated with radio-based distance-to-go signalling and onboard TIM are identified (X2Rail-3, 2020). These include errors in train position and train length, as well as other situations involving the erroneous clearance of tracks and switches. Assuming the use of GNSS-based TIM, PERFORMINGRAIL (2022a) has identified and analysed specific GNSS risks which could potentially trigger those conditions. Specifically, the following GNSS hazards are included:

- **Partial lack of GNSS visibility** occurs when less than four satellites are available. This generally is the case in urban or natural canyons, e.g., around central stations or in mountain valleys. Under partial lack of visibility, the estimated position of GNSS receivers includes an increased error.
- **Total lack of GNSS visibility** typically occurs in tunnels. With no GNSS satellites available, no estimation of GNSS receiver positions can be done; the provision of a position is denied.
- **Interference** is the natural or accidental disturbance of GNSS signals by noise emitted by other devices at the same frequency. The effect of interference on the GNSS signal quality is similar to lack of GNSS visibility.
- **Jamming** is the intentional interference of GNSS signals. As interference variant, its effect on train positioning is comparable to a lack of GNSS visibility.
- **Spoofing** corresponds to deceiving the GNSS receiver by forging GNSS signals, forcing the GNSS receivers to compute a wrong position. Also for spoofing, the effect on the train position is similar to lack of GNSS visibility.

- **Multipath propagation** corresponds to GNSS signals not reaching the GNSS receiver directly, due to nearby structures such as buildings or other trains. The signals might be reflected from the surrounding structures, reaching the receiver with a delay and/or disturbance. Depending on the severity, multipath propagation can lead to an increase of the positioning error.
- **Loss of lock** occurs when the receiver fails to maintain continuous tracking of the carrier phase of a satellite signal (also see Section 5.3.3), causing discontinuities in the carrier phase measurement known as cycle slips. Loss of lock is common in environments with degraded GNSS signal quality, such as urban canyons and vegetation canopies, and reduces the accuracy and reliability of the position estimate.
- **Atmospheric disturbances** introduce errors in GNSS signals that revert into positioning errors. In particular, ionospheric irregularities due to interaction between solar radiation and the Earth's geomagnetic field, affect the signal travelling from the satellites to the Earth's surface. Similar to the multipath hazard, the effects of ionospheric irregularities on train positioning range from increased positioning errors to a complete inability to provide a position. The latter occurred, for example, during the Mother's Day ionospheric storm in May 2024, when GNSS receivers used in farming, e.g., tractors, experienced complete service outages.

5.2.4 Railway traffic management under radio-based ETCS with GNSS-based TIM

In radio-based distance-to-go signalling systems with onboard TIM, real-time traffic management becomes essential for the mitigation of hazardous traffic conditions which may result from degraded performance in train positioning and communication systems.

Although not a vital component of the railway system, traffic management supports railway operations under both nominal and disturbed conditions. Its overall aim is to maintain conflict-free operations while minimising the impact on scheduled railway traffic. As such, traffic management is responsible for monitoring and predicting the traffic state, as well as the detection and resolution of track conflicts.

Track conflicts occur whenever a track section is required by more than one train during overlapping time intervals. Assuming an a priori conflict-free timetable, such conflicts only occur in case of shifted or extended train blocking times. That is, the time intervals during which a track section is reserved for a train and hence blocked for other trains. While shifted blocking times are usually due to train delays, extended blocking times can have various causes. For example, differences in running speed, delays in (radio) communication, e.g., related to the MA and TPR, or issues with the (GNSS-based) confirmation of train position and/or train integrity for TPR.

In practice, conflict detection and resolution is largely handled manually by human dispatchers. Mathematical optimisation models have been developed that can support dispatchers in taking optimised rescheduling decisions in case of disturbances, see, e.g., D'Ariano et al. (2007b), Corman & Meng (2015) and Pellegrini et al. (2015). Little of the literature on conflict detection and resolution models is however aimed at radio-based distance-to-go signalling, let

alone the more specific radio-based distance-to-go signalling with satellite-based TIM (Versluis et al., 2024).

In earlier work, we have enhanced the conflict detection and resolution model for conventional fixed-block signalling of Pellegrini et al. (2015) to describe train operations under different configurations of radio-based distance-to-go signalling including ETCS Level 2 with and without onboard TIM (Versluis et al., 2025a,b). In the model, train separation is modelled through blocking time theory (Hansen & Pachl, 2014). Hence, radio- and GNSS-related parameters such as communication times and positioning errors are incorporated in the blocking time components. Specifically, RBC-to-train communication of the MA is considered in the setup time, train-to-RBC communication of the TPR in the release time, and train positioning errors in the clearing time. Since the latter only accounts for predicted positioning errors, the concept of safety margin of the approach time remains applicable for unpredicted errors and other uncertainties. This model will be used in the illustration of the early-warning hazard prediction framework in Section 5.4.

5.3 Early-warning hazard prediction framework for railway traffic management

In this section, we propose an early-warning framework for hazardous railway traffic conditions under radio-based distance-to-go signalling with satellite-based TIM. The aim of the framework is to warn dispatchers ahead in time of potentially hazardous traffic conditions caused by increased signalling parameter values.

As shown in Figure 5.3, the early-warning framework extends the general railway traffic management framework. In addition to the core traffic management modules and the operational components of railway operations and dispatchers, the framework includes three dedicated early-warning modules: signalling state monitoring, signalling state prediction and early-warning generator. Before detailing these early-warning modules in Sections 5.3.2 to 5.3.4, we shortly describe the other framework components in Section 5.3.1. Within Section 5.3.3, we present a method for predicting train positioning errors in case of GNSS hazards.

5.3.1 Core traffic management modules

The core of railway traffic management frameworks, such as the ON-TIME framework (Quaglietta et al., 2016), consists of four modules representing the main tasks of traffic management and their connection with railway operations and dispatchers on the operational side. The traffic state monitoring module estimates the current traffic state, i.e., the current position and speed of every train in the network. This estimation relies on track occupancy data collected from real-time operations. The traffic state prediction module forecasts train positions and speeds over a given horizon, e.g., an hour (Quaglietta et al., 2016). Based on the predicted traffic state, the conflict detection module identifies potential track conflicts by projecting it onto the real-time traffic plan. If conflicts are detected, the conflict resolution module provides a set of rescheduling measures, i.e., retiming, reordering and/or rerouting, to minimise the impact of the detected

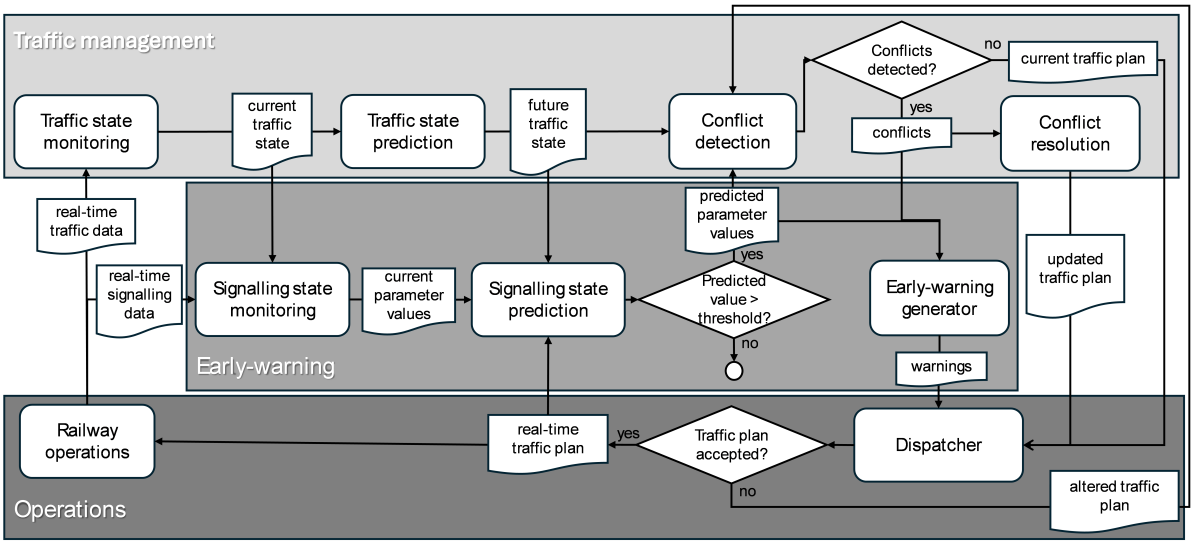


Figure 5.3: Early-warning framework for railway traffic management including core traffic management, new early-warning, and operations components.

conflicts. The proposed measures are integrated into the traffic plan and sent to the dispatcher for review and possible implementation during operations.

5.3.2 Signalling state monitoring

The signalling state monitoring module is responsible for collecting current values of signalling parameters, specifically the ones relevant for onboard train position and integrity monitoring. As such, it requires real-time signalling data from railway operations as well as the current traffic state. Onboard the trains, the real-time signalling data is combined with the train's position and speed to determine the train- and location-specific signalling parameter value. The obtained real-time signalling parameter values are transferred to the signalling state prediction, where they are used to validate predicted signalling parameter values.

5.3.3 Signalling state prediction

The signalling state prediction module forecasts signalling parameter values such as radio communication delays and GNSS positioning errors. The module combines information on the future traffic state and the real-time traffic plan to estimate train positions during the prediction horizon. Along the estimated path, effects of known indicators of radio and/or GNSS disturbances are projected to obtain a prediction of the relevant signalling parameter values. These indicators include occurrence of causes of GNSS hazards as described in Section 5.2.3, or similarly for radio hazards.

The prediction of signalling parameter values can be enhanced using current signalling parameter vales as a result of signalling state monitoring as follows. If the same increase in communication delay and/or positioning error is predicted for trains crossing a certain area, and for the first of these trains the value monitored deviates from the predicted value, then the monitored value can be set as the predicted value for the trains yet to reach the area.

In a subsequent step, it is checked whether any predicted parameter value exceeds the corresponding predetermined critical threshold. These critical values can be determined, for instance, through a sensitivity analysis with a model describing railway operations under radio-based signalling with GNSS-based TIM as has been done within PERFORMINGRAIL with a model of ETCS Level 2 Moving Block using stochastic activity networks (PERFORMINGRAIL, 2022b; Aoun et al., 2024). If the threshold value is exceeded, the predicted parameter values are sent to the conflict detection module. In its turn, the detected conflicts, if any, are sent to the conflict resolution as well as to the early-warning generator.

Given that the primary function of the signalling state prediction module is the forecasting of signalling parameter values such as GNSS positioning errors, we dedicate the remainder of this section to the presentation of a simulation-based approach for predicting train positioning errors under GNSS hazards.

Within PERFORMINGRAIL, a GNSS software-in-the-loop simulator has been implemented (PERFORMINGRAIL, 2021; Mazini et al., 2022) in order to provide a tool for predicting train positioning errors under the effect of various GNSS hazards. This GNSS Location Simulator (GLS), emulates the behaviour of a GNSS system and can be integrated in a larger simulator of systems that require GNSS functionality, such as a railway (signalling) simulator. In particular, GLS simulates the different stages and processing steps of a GNSS receiver: the measurement engine, in charge of satellite signal tracking and carrier phase and pseudorange measurement generation, and the positioning engine, that transforms the measurements into position and speed estimates.

The architecture of GLS, with the external interfaces and building blocks is shown in the block diagram in Figure 5.4. The communication interface between the external entity, e.g., a railway signalling simulator (RSS) and the GLS is based on a transmission control protocol (TCP) where data (in our case positions) are sent using a comma separated value (CVS) format. The measurement engine (`argos`) simulates and sends the GNSS range measurements to the positioning filter (`rift`), which estimates the position of the train based on the simulated input measurements.

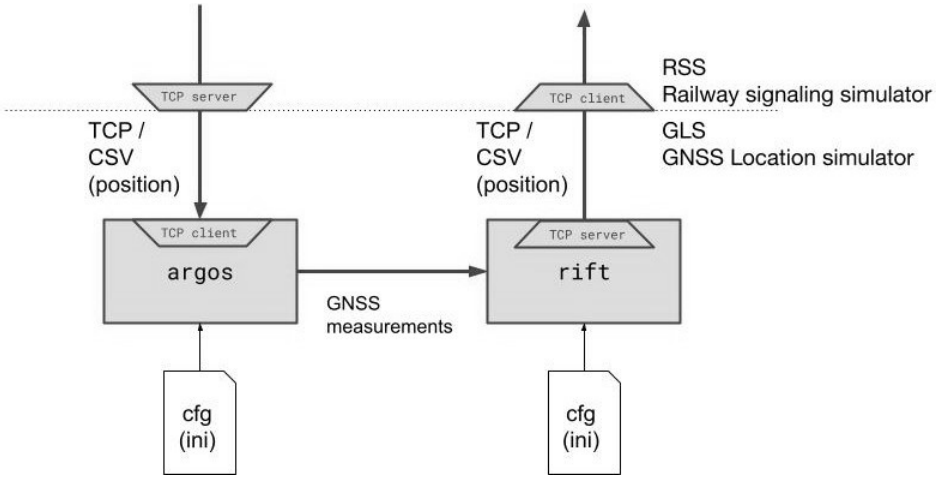


Figure 5.4: Diagram of the GNSS Location Simulator (GLS) main building blocks and interface with the Railway Signalling Simulator.

The working principle of GLS is as follows:

1. The calling entity, i.e., an RSS or an actual real-life railway environment, sends a position to GLS.
2. GLS uses the incoming positions of simulated GNSS receivers of the train. These positions are used by the measurement engine to calculate the expected GNSS measurements at these positions. To this end, various models for the GNSS pseudorange components such as satellite orbits, propagation delay due to the atmosphere and receiver thermal noise, are used to compute the GNSS measurements, e.g., code pseudorange, carrier phase delay, in line with ZhG (2021). Moreover, these signal models can also include potential signal hazards, such as an increase of noise due to multipath propagation or signal obstructions.
3. The GNSS measurements are forwarded to the positioning engine (`rift`), which computes the position as would be computed by a GNSS receiver.
4. The computed position, i.e., the GLS position, is sent back to the calling entity.

In the ideal case of no errors, the position delivered by GLS would be identical to the input position. However, in a more realistic scenario that considers errors and signal disruptions, the GLS position will differ from the input position. GLS incorporates a model for the GNSS pseudoranges that includes several potential hazards and can be used to simulate various scenarios and assess the potential impact of the GNSS hazards in position estimates as well as its associated uncertainty estimation.

As an example of GLS usage, consider an RSS assessing multipath hazards in urban canyons. For each simulation epoch, RSS sends the ideal train position, e.g., the estimated position of a train through the projection of the predicted traffic state onto the real-time traffic plan, to GLS. GLS calculates GNSS measurements including multipath errors, and the positioning engine computes the resulting position. This simulated position is returned to RSS, allowing the impact of GNSS hazards on train positioning to be evaluated.

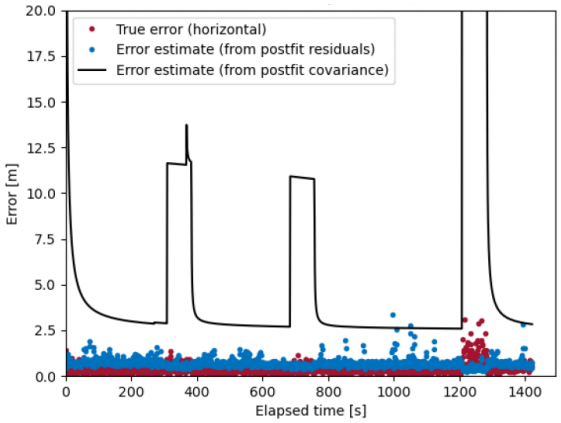
The following scenarios based on the GNSS hazards listed in Section 5.2.3 have been simulated with GLS: lack of visibility, multipath, loss of lock due to signal degradation, and a combination of multipath and loss of lock (PERFORMINGRAIL, 2023). For the simulation, the Tuxford test track, operated and managed by United Kindgom's Rail Innovation and Development Centre (RIDC), was used as baseline position for the different scenarios.

The main output of the simulations is the formal errors of the positioning system. This output allows to establish an empirical relationship between the formal error, as estimated by the filter, and the actual positioning error, i.e., the difference between the input position to GLS by RSS and the position resulting from the simulation performed by GLS, sent back to RSS. An example of this comparison, focusing on the Tuxford test track with three identified 'affected' areas as illustrated in Figure 5.5, is presented in Figure 5.6. The results show that formal errors experience a sudden increase as a result of the presence of positioning hazards. This particularly holds for formal errors derived from the postfit covariance matrix, indicated with black lines in Figure 5.5.

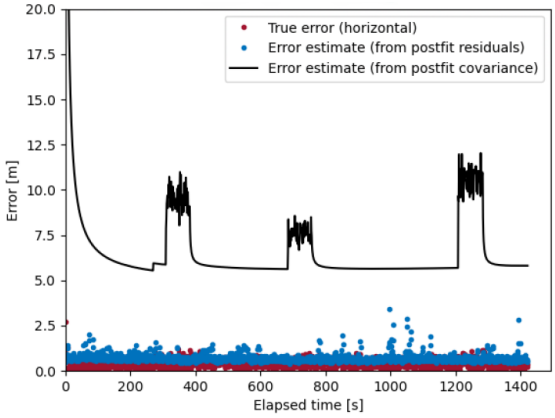
Moreover, since the scenarios are simulated and thus the true position is known, it is possible to obtain the actual positioning error and compare it with the formal error, represented as red dots in the panels of Figure 5.6, and establish a relationship between the formal and actual errors. Besides the formal error computed from the post-fit covariance, an error estimate based on the post-fit residual analysis can be established. These post-fit residual error estimates are

shown as blue dots in Figure 5.6. In general, the error estimate using the post-fit residuals generated by the filter (blue dots) is a better estimate of the actual positioning error (red dots) and hazard indicator, compared to the formal error (black lines). See, for instance, the error values in the multipath scenario (Figure 5.6c), where the formal error barely notices the impact of the hazard, but the post-fit residual error captures the event.

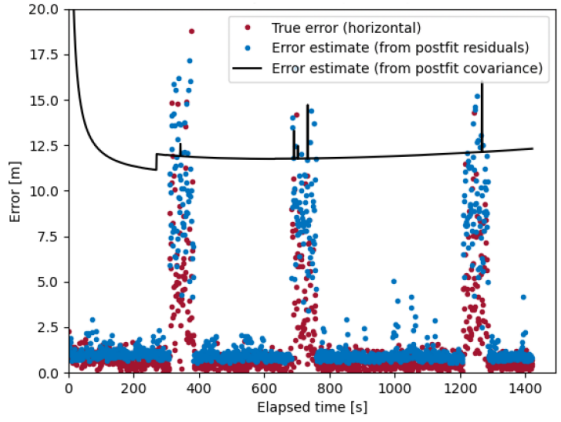
Figure 5.5: Example of a scenario simulating degraded visibility in three segments on the Tuxford test track (PERFORMINGRAIL, 2023).



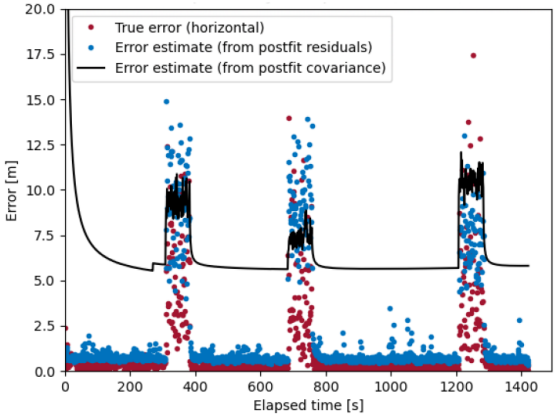
(a) Lack of visibility.



(b) Loss of lock.



(c) Multipath.



(d) Multipath + loss of lock.

Figure 5.6: Positioning errors in GNSS hazard scenarios (PERFORMINGRAIL, 2023).

A reliable strategy to estimate the positioning error is important because it allows to establish a methodology to assess the trustworthiness of the formal error reported by the positioning system. This is also known as positioning integrity, and is crucial in safety-of-life applications such as aeronautics and road transport. Traditionally, GNSS position integrity concepts were developed for aviation, but their direct application in urban environments is challenging due to multipath and reduced satellite visibility (Zhu et al., 2018).

One of the main outputs that can be obtained from the simulations is the error characterisation and, in particular, the possibility to forecast the positioning error. Despite the fact that GNSS hazards are of random nature, especially those related to receiver thermal noise, some considerations can be made. As introduced in Section 5.2.3, multipath errors are typically the most critical GNSS signal errors in railway applications, particularly in urban or obstructed environments, while atmospheric errors can dominate in open-sky conditions for single-frequency receivers. Future hardware improvements, such as the wider availability of dual- or multi-frequency GNSS receivers, may help mitigate these effects (Blatnik & Batagelj, 2025).

Given the fact that the path of the train is well known, it is possible to forecast the positioning error due to multipath and/or lack of visibility if the distribution of obstructions are known. There are various public databases such as Microsoft or Google Open Building databases that could be used to predict the multipath error along the track that will be traversed by the train. A similar reasoning can be done in areas with limited visibility as well as error increase due to dense foliage, i.e., forests. For these cases, publicly available land cover maps can be used. Finally, the ionosphere is a slow varying medium that can be easily predicted in the short-term, provided that there are no ionospheric storms or periods with high geomagnetic activity. Therefore, the error associated with it along the train path can also be forecasted. This approach aligns with recent railway studies on GNSS-based solutions, which emphasize the estimation of local errors for different environments to support safety and operational reliability, while also highlighting the inherent difficulties of such estimation (Mayolle et al., 2025).

5.3.4 Early-warning generator

The early-warning generator is the final stage in the early-warning hazard prediction process. In this stage, relevant information originating from the signalling state prediction is combined and sent to the dispatcher. For the generation of early-warnings to the dispatcher, the module requires the output from the conflict detection module given the predicted signalling state. Namely, the trains and locations involved in track conflicts detected due to threshold exceeding signalling parameter values. Additionally, details about the triggering of signalling parameters should be provided. This includes the type of parameter, e.g., communication delay or positioning error, the associated train and the (predicted) location.

The main function of the early-warning generator is to generate an early-warning message upon receiving the notification that a conflict has been detected due to predicted signalling parameter values exceeding the corresponding threshold. The warning is populated with information on the trains and locations of the detected conflicts, as well as the train and location to which the underlying predicted signalling parameter value correspond.

5.4 Illustration of early-warning hazard prediction framework

We illustrate the early-warning hazard prediction framework through a case study based on a specific use case. The case study and the use case are detailed in Sections 5.4.1 and 5.4.2, respectively. Section 5.4.3 shows the results of the application of the framework. For the detection and resolution of conflicts in the application of the framework, we use the state-of-the-art RECIFE-MILP rescheduling algorithm enhanced for ETCS Level 2 with Onboard TIM as presented in Versluis et al. (2025b).

5.4.1 Case study: RIDC Melton with disturbed operations

As case study, we consider disturbed traffic conditions on the Melton test track managed by the United Kingdom's Rail Innovation and Development Centre (RIDC). As represented in Figure 5.7, the track connects three stations: Melton on the south end, Tollerton on the north end, and Old Dalby in the middle. The track configuration consists of single-track segments on both ends, with a central double-track stretch extending from Old Dalby to the north.

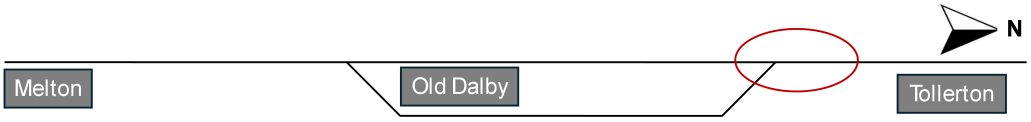


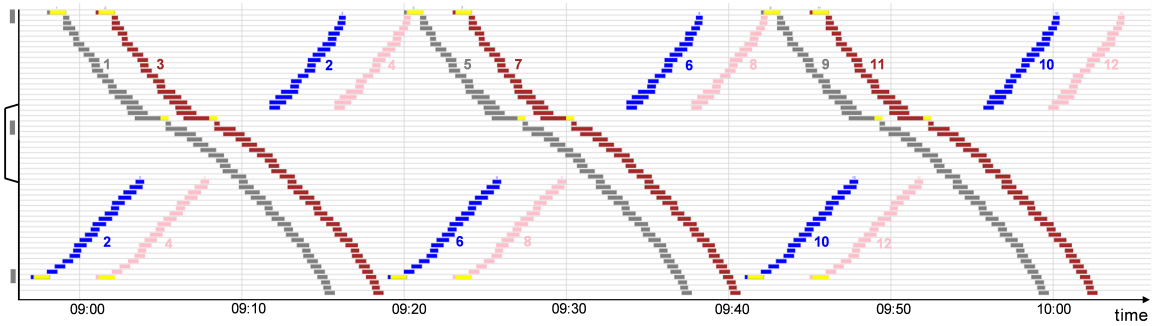
Figure 5.7: Schematic layout of RIDC Melton, with degraded area considered in Section 5.4.2. Based on NetworkRail (2015).

Under the assumption of ETCS Level 2 Virtual Block, the timetable covers a 65-minute period during which 12 trains run in three cycles. Each cycle involves four trains, two per direction, that meet on the central double tracks. The trains are numbered 1 to 12: odd-numbered trains run the northbound, from Melton to the north end, with scheduled stops at Melton and Old Dalby; even-numbered trains run southbound, from Tollerton to Melton, with scheduled stops at Tollerton, Old Dalby and Melton. The timetable is provided in terms of arrival and departure times (at the level of seconds) in Table 5.1 and blocking time diagram in Figure 5.8a. The blocking time diagram indicates the time periods during which track sections along the northbound route are assigned for specific trains. That is, from the moment a train requests an MA for that section that is accepted by the interlocking until the section is released after the train has traversed it. The total blocking time consists of the following components: setup, reaction, approach, running, clearing and release time (Hansen & Pachl, 2014). In Versluis et al. (2025b), the blocking times under ETCS Level 2 with Onboard TIM are detailed.

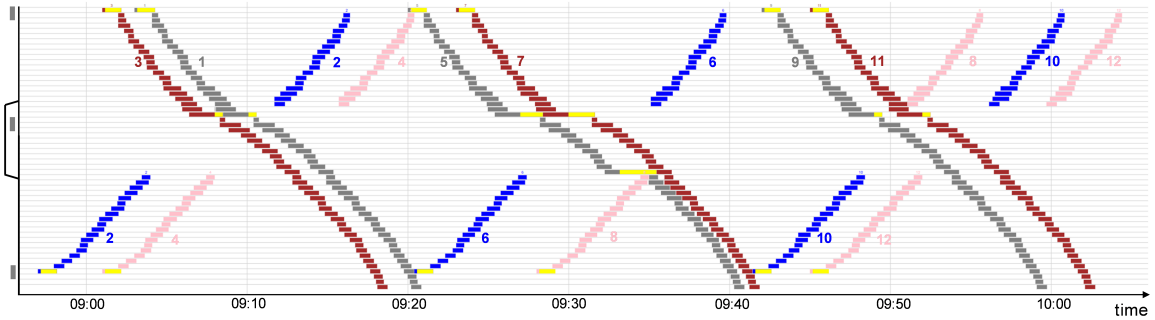
For simplicity, we assume that the timetable does not include any time supplements and that there are minimum buffer times between the cycles (35 seconds at the south end and 43 seconds at the north end). Pairs of two trains running in the same direction have a scheduled headway of three minutes (northbound) or four minutes (southbound). The minimum buffer time occurs at Old Dalby is 54 seconds for northbound trains and 76 seconds for southbound trains.

Table 5.1: Timetable for RIDC Melton (hh:mm:ss), with train delays (+mm:ss) as in real-time traffic plan for 5-minute entry delay of trains 1 and 8.

Northbound trains		**1**$^{+05:00}$	**3**	**5**	**7**	**9**	**11**
Melton	Arrival	08:58:00$^{+05:00}$	09:01:00	09:20:00	09:23:00	09:42:00	09:45:00
	Departure	08:59:00$^{+05:06}$	09:02:00	09:21:00	09:24:00	09:43:00	09:46:00
Old Dalby	Arrival	09:04:44$^{+05:06}$	09:07:44	09:26:44	09:29:44	09:48:44	09:51:44
	Departure	09:05:14$^{+05:06}$	09:08:14	09:27:14	09:30:14	09:49:14	09:52:14
North end	Exit	09:15:34$^{+05:06}$	09:18:34	09:37:34$^{+02:44}$	09:40:34$^{+00:41}$	09:59:34	10:02:34
Southbound trains		**2**	**4**	**6**	**8**$^{+05:00}$	**10**	**12**
Tollerton	Arrival	08:57:00	09:01:00	09:19:00$^{+01:14}$	09:23:00$^{+05:00}$	09:41:00$^{+00:26}$	09:45:00
	Departure	08:58:00	09:02:00	09:20:00$^{+01:14}$	09:24:00$^{+05:00}$	09:42:00$^{+00:26}$	09:46:00
Old Dalby	Arrival	09:09:57	09:13:57	09:31:57$^{+01:14}$	09:35:57$^{+05:00}$	09:53:57$^{+00:26}$	09:57:57
	Departure	09:10:57	09:14:57	09:32:57$^{+01:14}$	09:36:57$^{+05:00}$	09:54:57$^{+00:26}$	09:58:57
Melton	Arrival	09:16:06	09:20:06	09:38:06$^{+01:14}$	09:42:06$^{+13:37}$	10:00:06$^{+00:26}$	10:04:06
	Departure	09:17:06	09:21:06	09:39:06$^{+01:14}$	09:43:06$^{+13:37}$	10:01:06$^{+00:26}$	10:05:06



(a) Original traffic plan according to timetable, along route of northbound trains.



(b) Real-time traffic plan for considered delay scenario, along route of northbound trains.

Figure 5.8: Schematic blocking time diagrams of traffic plans. Yellow indicates scheduled stops and additional waiting delays.

We focus on a specific delay scenario rather than the nominal situation, as delay scenarios typically involve more critical traffic conditions. Under such scenarios, the assumptions of no time supplements and minimum buffer times are justified. As indicated in Table 5.1 (by +05:00), trains 1 and 8 are both experiencing an entry delay of 5 minutes. The real-time traffic plan for this scenario, generated by the conflict detection and resolution algorithm from Versluis et al. (2025b), is illustrated by the time-distance diagram in Figure 5.8b. The delays for all trains with respect to the original traffic plan, i.e., the timetable, is indicated in minutes and seconds in Table 5.1. The 5-minute delay of train 1 causes it to reorder with train 3, following it with a

minimum 2-minute headway. Due to the lack of time supplements and the long blocking time at Old Dalby, train 1 arrives 5 minutes and 6 seconds late. This delay propagates to train 6, which travels in the opposite direction and is consequently delayed by 1 minute and 14 seconds at both departure and arrival.

The 5-minute delay to train 8 has more impact on the traffic plan. Train 8 departs from Tollerton 7 minutes and 36 seconds after train 6, due to the knock-on delay of train 1 on train 6. As a result, trains 5 and 7 running in the opposite direction have to wait for train 8 to clear the northern single track. Hence, trains 5 and 7 arrive at the north end with a delay of 2 minutes and 44 seconds and 41 seconds, respectively. The exit delay of train 7 affects train 10, which travels with a minimum delay of 26 seconds along its route. To prevent delays to trains 9 and 11 at the southern single track, train 8 has to wait after stopping at Old Dalby to enter the double track. Consequently, train 8 arrives at its destination 13 minutes and 37 seconds behind schedule, so with an additional 8 minutes and 37 seconds of delay compared to its initial delay.

5.4.2 Use case: Atmospheric disturbances causing GNSS and radio signal degradation

As use case, we consider a scenario in which atmospheric disturbances impact railway traffic conditions through GNSS and radio signal degradation. The atmosphere includes the tropospheric and the ionospheric layers (see Figure 5.9), each introducing different types of interference that can cause signal degradation. The troposphere is the lowest layer of the atmosphere, where most weather phenomena occur, such as rain, wind and fog. Tropospheric conditions affecting signal degradation can be predicted using meteorological data such as temperature, pressure, humidity and precipitation (Saastamoinen, 1972; Akinwole & Yussuf, 2023). The ionosphere is the ionised part of the upper atmosphere, which is affected by interactions between solar radiation and the Earth's geomagnetic field. Ionospheric conditions can be forecasted around 24 hours in advance, with prediction methods relying on historical data of previously observed ionospheric activity and its effects on satellite and radio signals (Sizun, 2005).

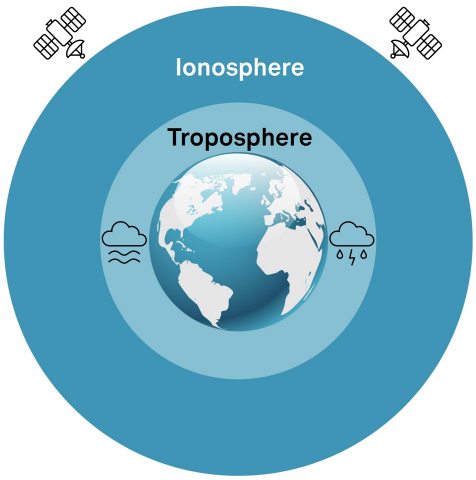


Figure 5.9: Atmospheric layers of troposphere and ionosphere. Adjusted from Jeffrey & Munro (2023).

GNSS signals travel from satellite to receivers on the Earth's surface, passing through the ionosphere and troposphere. As they propagate, these signals are subject to various atmospheric effects that can degrade their positioning accuracy. GNSS positioning errors are particularly dependent on ionospheric conditions. Under nominal conditions, these errors are typically around two metres (Jeffrey & Munro, 2023), as shown in Figure 5.6. However, ionospheric irregularities, caused by phenomena such as solar flares or geomagnetic storms, can introduce significant inaccuracies. While GNSS receivers can effectively mitigate many of these effects, particularly in terms of relative positioning as used for train integrity confirmation, absolute positioning errors can reach 10 to 50 metres during periods of intense ionosphere activity such as ionospheric scintillation (Space Weather Prediction Center, 2025). In addition to ionospheric effects, the troposphere also contributes to GNSS positioning errors. The troposphere can cause signal delay due to refraction, which depends on temperature, pressure and humidity. While these delays are generally more stable, they can vary significantly with local weather conditions, especially in the presence of heavy rain, fog, or rapidly changing atmospheric pressure. Although tropospheric errors are typically smaller than ionospheric ones, they can still introduce several metres of positioning errors, particularly during severe weather events (Xu & Xu, 2016).

Operating at low-frequency radio signals, GSM-R communication is less affected by ionospheric distances than GNSS signals. However, severe ionospheric scintillation can lead to temporary GSM-R signal degradation or loss of lock. Typical message update intervals in GSM-R systems can increase from a nominal 500 ms to several seconds during periods of interference or signal loss (GSM-R Operators Group, 2015). Moreover, tropospheric disturbances such as storm, heavy rain or dense fog can cause signal delays and even temporary outages (European Union Agency for Cybersecurity, 2024).

Based on the discussion of atmospheric disturbances and their impact on signal reliability, the following values have been selected for the illustration of the early-warning framework. The GNSS positioning error is set to 50 metres in the degraded area, corresponding to an off-set of 48 metres compared to the default positioning error of two metres. This value reflects conditions during severe atmospheric disturbances, such as ionospheric scintillation. For GSM-R communication, a delay of five seconds is assumed, reflecting potential temporary loss of lock caused by sever weather conditions such as storm, heavy rain or dense fog.

5.4.3 Application of early-warning framework

We consider the atmospheric conditions and corresponding GNSS positioning errors and GSM-R communication delays along the Melton track to be forecasted in advance of operations. We assume the positioning errors of up to 50 metres and communication delays of up to five seconds in the degraded area around the northern switch. For simplicity, we assume the lack of buffer times at critical sections in the real-time traffic plan. Therefore, it is obvious that the predicted increased positioning error and communication delay exceed the safe threshold. Hence, the predicted values are send to the conflict detection module.

The increased positioning error and communication delay lead to extended blocking times. The clearing time is affected by the positioning errors as it effectively extends the safe, estimated

train length. The setup and release times are affected by the communication delays due to the required communication for MA and TPR updates, respectively. While the communication delay of five seconds is constant, the effect of the positioning error depends on the train speed. In the degraded area, the train speed is 80 km/h. At 80 km/h, an extended train length of 48 metres leads to just over two seconds of extra clearing time. With the setup and release times both extended by five seconds, we assume blocking time extension of twelve seconds in the degraded area. That is, for trains traversing the area while running 80 km/h. As also shown in Figure 5.8b, train 5 is the exception to that, having an unplanned stop in the degraded area in the real-time traffic plan.

To determine the extra clearing time of the sections in the degraded area for train 5, we consider its speed profile in the area as shown in Figure 5.10. For each section other than the one with the unplanned stop, the extra clearing time is calculated as the difference in time required for the train front to traverse its actual train length (150 metres) plus a positioning error of two or 50 metres from the section exit. This gives additional clearing times of maximum 4.8 seconds, which is the case at the end of the first section after the stop, i.e., the switch of 141 metres long. So, the blocking time of the switch for train 5 is extended by a total of 15 seconds. For the section with the unplanned stop, we add the time it takes to accelerate from standstill over the first 48 metres, i.e., twelve seconds, to the ten seconds of communication delay. In Figure 5.11a, the extended blocking times are projected onto the real-time traffic plan.

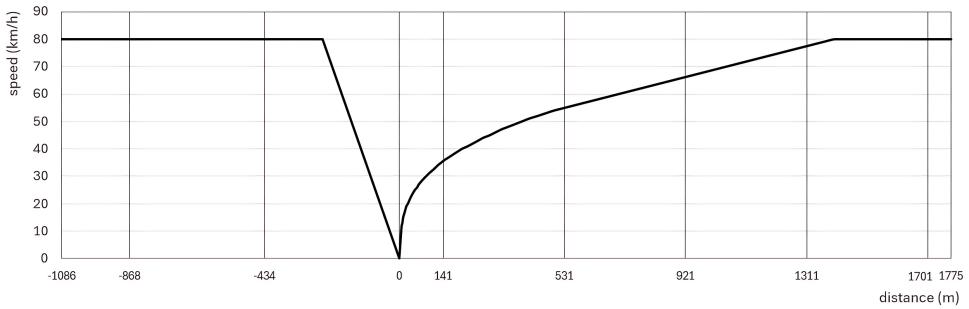
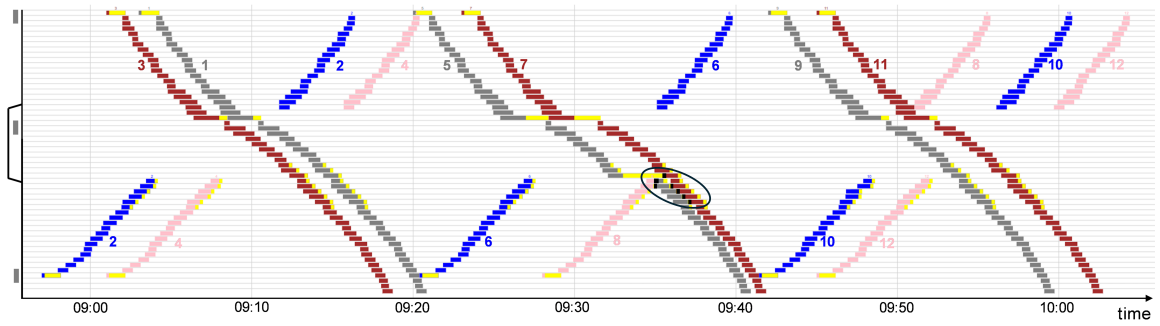


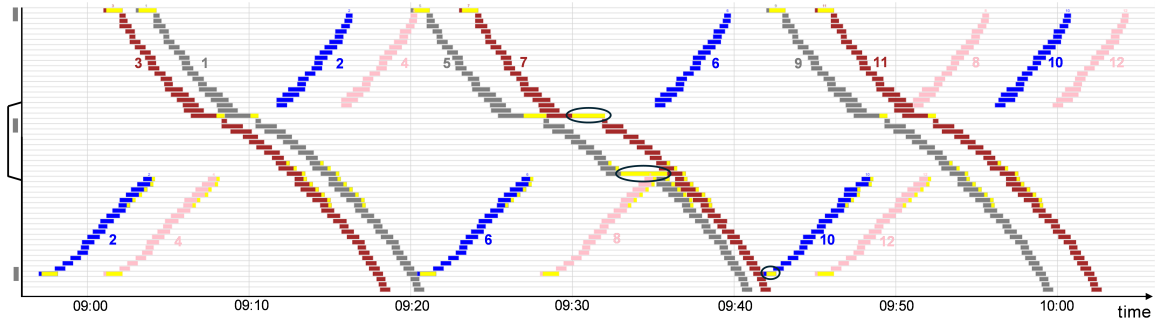
Figure 5.10: Speed-distance diagram for train 5 over the virtual sections in degraded area. Distances correspond to the relative position of section boundaries with respect to the unplanned stopping point.

Taking into account the extended blocking times, the conflict detection module detects conflicts between trains 5, 7, and 8 around the switch in the degraded area. The detected conflicts are sent to the early-warning generator, as well as to the conflict resolution module. A warning is issued to the dispatcher stating that trains 5, 7 and 8 are involved in track conflicts at the northern switch due to increased positioning errors in combination with communication delay, together with the updated conflict-free traffic plan provided by the conflict resolution module.

In the updated traffic plan the dispatcher receives, train 5's unplanned stop is extended and the departure of train 7 from Old Dalby is delayed such that it can continue its journey afterwards without further restrictions. The latter, however, causes train 7 to arrive late at the northern end of the track, which has caused train 10 to depart late. The dispatcher, however, decides to shorten the dwell time of train 10 at Tollerton, such that it can depart on time. This results in the final real-time traffic plan to be implemented, shown in Figure 5.11b.



(a) Traffic prediction and conflict detection based on real-time traffic plan, with overlapping blocking times indicated, along route of northbound trains.



(b) Real-time traffic plan implemented by dispatcher after receiving early-warning, with the three changes indicated, along route of northbound trains.

Figure 5.11: Schematic blocking time diagrams of traffic plans with extended blocking times due to increased GNSS positioning errors and GSM-R communication delays in degraded area. Yellow indicates scheduled stops and additional delays.

After operations started, train 2 is the first to cross the area with the predicted increased positioning errors and communication delays. The real-time values that the train is monitoring while running through the area confirm that there is an increase in positioning errors, compared to the first part of its route. However, we assume that the monitored values are lower than the predicted value: circa 20 metres instead of 50 metres. This data is fed from the signalling state monitoring to the signalling state prediction module, which consequently updates the positioning error prediction for trains 5 to 12. Due to the overruling communication delay that is also measured in real time, the monitoring-based predicted errors are dealt with in the same way as the predicted errors. That is, according to the real-time traffic plan in Figure 5.11b.

For comparison, we consider the case study scenario without the availability of the early-warning framework. Without the framework in place, the degradation of GNSS and GSM-R signals and the corresponding effects on scheduled operations would not have been predicted. As a result, trains entering the area affected by the atmospheric disturbances would experience unexpected delays in receiving their MA and sending TPR updates, along with a sudden increase in reported safe train length. In most cases, the resulting additional blocking times of up to a couple of tens of seconds would be absorbed by buffer times in the traffic plan. However, this buffer may be insufficient in case of delays or tightly schedule operations. In our case, train 7 would have departed from Old Dalby at its scheduled time, reaching the northern switch before it was released due to the extended blocking times of train 8 and 5. Consequently, train 7 would have had to either reduce its speed significantly or make an unplanned stop. In either case,

overall operational efficiency would be compromised. These last minute interventions can lead to substantial reduction in operational performance.

We conclude with the quantification of the impact of the framework for the considered scenario of GNSS and GSM-R signal degradation during already disturbed operations, using several performance indicators related to operational efficiency. The scope of impact is indicated by the number of affected trains and the impact duration, i.e., the period during which the operations deviate from the (original) real-time traffic plan, and is reflected in the total arrival and departure delays experienced by the affected trains within that period. Besides total delays, the time loss due to unplanned braking, required to respect safety-critical MAs, is highlighted. Additional operational effects are considered to illustrate how the framework influences railway traffic management. These include the number of extended and unplanned stops, the number of rescheduling decisions (retiming only in this scenario), and the number of resolved conflicts. The indicator values with and without the application of the framework are listed in Table 5.2. The first two columns report values relative to the base scenario of disturbed operations without signal degradation, for which the real-time traffic plan is given in Figure 5.8b. The third column shows the impact of the framework as the difference in indicator values between 'with framework' and 'without framework'.

Table 5.2: Railway traffic management performance indicators for use case scenario with and without framework, relative to the real-time traffic plan.

Performance indicator	**With framework**	**Without framework**	**Difference (with − without)**
Number of affected trains	3	3	0
Impact duration (mm:ss)	10:59	20:17	−09:18
Total arrival delay (s)	60	106	−46
Total departure delay (s)	36	56	−20
Time loss due to restricted MA (s)	0	26	−26
Number of extended stops	2	1	+1
Number of unplanned stops	0	1	−1
Number of rescheduling decisions	2	3	−1
Number of resolved conflicts	3	3	0

While the number of affected trains stay the same (trains 5, 7 and 10), application of the framework reduces the impact duration by 48%, from just over 20 minutes to practically 11 minutes. Similarly, the overall impact in terms of train delay is significantly lower with application of the framework. With the framework, the total arrival delay of the three affected trains amounts to 60 seconds (12 seconds for train 5, 24 seconds for train 7 and 24 seconds for train 10), and the total departure delay is 36 seconds (12 seconds for train 5 and 24 seconds for train 7). Without the framework, an additional 46 seconds of arrival delay (26 seconds for train 7 and 20 seconds for train 10) and 20 seconds of departure delay (train 10) was observed. Of this additional arrival delay, 26 seconds are due to a restricted MA (train 7). This effect is also reflected in the number of unplanned stops. Without the framework, one extra unplanned stop occurs. With the framework, it was replaced by a less critical extended stop. Finally, the number of rescheduling decisions and the number of resolved conflicts show that, with the framework, one less intervention was required to resolve the same number of conflicts. These results demonstrate that, for the considered scenario, application of the framework im-

proves the management of railway operations under signal degradation, enabling proactive and efficient conflict detection and resolution.

Overall, the illustration of the framework shows its (potential) applicability to use cases comparable to degraded GNSS and/or GSM-R signals due to atmospheric conditions, e.g., lack of GNSS visibility due to dynamic factors beyond fixed environments such as natural of urban canyons.

5.5 Conclusions

In this chapter, we proposed an early-warning hazard prediction framework for the effective management of real-time railway operations under radio-based distance-to-go signalling with satellite-based train integrity monitoring. The framework extends the general railway traffic management framework with an early-warning module with the aim to warn dispatchers ahead in time of potentially hazardous traffic conditions due to signalling parameters attaining values beyond critical thresholds. These parameters include radio communication delays and GNSS positioning errors. Concerning the latter, a method for the monitoring and prediction of positioning errors in case of GNSS hazards is described.

The early-warning framework is illustrated through its application to a case study based on the Melton test track at the United Kingdom's Rail Innovation and Development Centre. Specifically, the framework is applied to the use case of atmospheric disturbances leading to degraded GNSS and radio signals. Thanks to the framework's ability to monitor and predict GNSS positioning errors and radio communication delays, hazardous conditions were identified well in advance. This early detection enabled the dispatcher to respond proactively, preventing further compromise of operational efficiency.

Next research steps include the further refinement of GNSS positioning error prediction under atmospheric conditions and its application to other use cases, i.e., other dynamic conditions potentially leading to degraded GNSS and/or radio signals. A general next step is to set up a railway traffic management loop including the early-warning modules in a simulation environment. In such an environment, real-life scenarios based on historical data can be used to establish a validated early-warning hazard prediction system for railway traffic management under radio-based signalling with satellite-based train integrity monitoring.

Chapter 6

Conclusions

The research in this thesis contributes to the intersection of railway signalling and railway traffic management, focusing on the safety and efficiency aspects of real-time operations. With the aim to develop models that support the effective management of real-time railway operations under radio-based signalling with distance-to-go automatic train protection, this thesis includes a literature review and research agenda (Chapter 2), a conflict detection and resolution model (Chapter 3), an impact assessment of track discretisation (Chapter 4), and an early-warning framework (Chapter 5).

This chapter contains conclusions and recommendations concerning the included work. First, the main findings in relation to the research questions posed in Chapter 1 are shared in Section 6.1. Then, recommendations for practice and for future research are given in Section 6.2 and Section 6.3, respectively.

6.1 Main findings

To achieve the aim of this thesis, five research questions were formulated and addressed. This section provides the answers to these questions.

1. What are the current research gaps and challenges in modelling conflict detection and resolution under moving-block as next-generation radio-based distance-to-go railway signalling system?

In Chapter 2, a comprehensive review of the literature on conflict detection and resolution models is presented. These models have been developed to support railway traffic management in taking train rescheduling decisions, such as retiming, reordering and rerouting, in order to minimise the impact of disturbances that lead to conflicting operations. Most existing models focus on conventional fixed-block signalling, in which trains are separated by a predefined number of blocks, each with a fixed length and protected by trackside signals, and supervised by automatic train protection systems to ensure compliance to the multi-aspect block signalling. In contrast, under moving-block and other radio-based distance-to-go railway signalling systems, train separation is based on train- and speed-dependent braking distances, facilitated by radio-

based cab signalling and continuous braking curve supervision, i.e., distance-to-go automatic train protection.

Consequently, the main research gap is identified as the lack of models that accurately describe railway operations considering train separation based on braking distances for main-line railways. In moving-block systems, both the brake indication point and the end of movement authority (EoA), i.e., the location to which the train is authorised to run, are dynamically calculated and can be located anywhere along the track. While in fixed-block distance-to-go signalling system, the EoA is bound to lie at predefined block boundaries, the associated braking curve is calculated backwards from this point and may therefore begin anywhere on the track. Due to the definition of both EoA and brake indication at block entries under conventional signalling, typical modelling approaches rely on the discretisation of the track into fixed blocks and a simplified dependency of train separation on speed. Correspondingly, the main research challenges are the modelling of the brake indication point and/or EoA in continuous space, in combination with the incorporation of dynamic, speed-dependent train separation.

2. How can conflict detection and resolution under radio-based distance-to-go railway signalling be modelled?

In Chapter 2, a research agenda is presented for modelling conflict detection and resolution under moving block as next-generation radio-based distance-to-go railway signalling system. The proposed steps are based on a comparative analysis of relevant modelling approaches and aim to address the previously identified research gaps regarding the lack of models to describe railway operations with train- and speed-dependent train separation. On the one hand, a shift to continuous-space dynamic system approaches is suggested, as these have been proven effective for modelling the train control problem under moving-block signalling. However, such models tend to be highly complex in multi-train scenarios and, as a result, lack a clear connection to conflict detection and resolution. On the other hand, it is proposed to explore the applicability of the established (discrete-space) conflict detection and resolution approaches of alternative graph and (disjunctive) mixed integer linear programming (MILP) to approximate moving-block and less advanced radio-based distance-to-go operations with short (virtual) blocks.

In Chapter 3, an approach is proposed to enhance existing conflict detection and resolution models, originally developed for conventional fixed-block signalling, to describe radio-based fixed-block signalling with distance-to-go automatic train protection. First, a distinction is made between interlocking areas and the open line. Within interlocking areas, movable track elements such as switches require additional safety protection, and are therefore considered as fixed (virtual) blocks - unlike the open line, for which (shorter) virtual or moving blocks are considered. Second, alternative speed profiles are modelled corresponding to a maximum and a scheduled speed profile according to the planned timetable, including stops. These speed profiles are defined between interlocking areas, and the assigned speed profile affects the train running and blocking times on the open line. Maximum speed profiles should only be assigned if they contribute to delay recovery. Third, train separation is based on (train- and) speed-dependent braking distances. The braking point required to stop before a specific (virtual) block is calculated based on the train's braking characteristics for both speed profile options. Fourth, the setup time, i.e., the time to set and lock a route for a train and update the train's movement authority (its permission to move to a specific location under supervision), is adapted to account for the transition from trackside to cab signalling based on radio communication.

The proposed enhancements are applied to the state-of-the-art RECIFE-MILP rescheduling model, which allows for the retiming, reordering and rerouting of trains in order to minimise the total train delay. The enhanced model is verified for the radio-based fixed-block distance-to-go signalling system of ETCS Level 2 with Trackside Train Detection (TTD), where the fixed blocks correspond to the track sections in which the presence or absence of a train is automatically detected.

In Chapter 4, the RECIFE-MILP model is further enhanced to describe virtual-block signalling, specifically ETCS Level 2 Virtual Block, which is characterised by onboard train position and train integrity monitoring (TIM). First, TTD parameters are replaced by onboard TIM parameters in the calculation of minimum train separation through blocking times. In particular, the release time, i.e., the time between the physical clearance of the section by a train until the section can be assigned to another train, is affected. This release time now depends on the update and communication time associated with reporting the train's position and integrity information from onboard systems to the trackside. Second, train separation in interlocking areas is redefined to become route-dependent, reducing the operational criticality of switches under virtual-block signalling. In particular, the setup time, i.e., the time to request, set and lock a train route over a switch, is based on the current and required position of the switch.

3. What are the effects of radio-based fixed-block distance-to-go signalling over conventional fixed-block signalling in conflict detection and resolution?

In Chapter 3, a comparative analysis is conducted between the enhanced RECIFE-MILP model for radio-based fixed-block distance-to-go signalling and the original RECIFE-MILP model for conventional fixed-block signalling. The models are applied to two distinct case studies: one featuring a complex junction with dense mixed traffic, and another representing a corridor with regular mixed traffic. For each case study, one hundred uniformly generated delay scenarios are considered in which 20% of the trains are imposed an entry delay between 5 and 15 minutes during a peak hour. A cross-evaluation is performed between the rescheduling solutions generated by the two model versions, respectively representing ETCS Level 2 with TTD and conventional multi-aspect fixed-block signalling, to assess the effects and modelling implications of radio-based distance-to-go versus conventional fixed-block signalling.

As main focus, the effects on conflict detection and resolution are analysed in terms of rescheduling decisions and total train delay. While a reduction in total delay of radio-based distance-to-go over conventional fixed-block signalling is generally observed, the magnitude of these reductions is typically limited. However, several case study scenarios show significant improvement, with up to 7% less total delay. Most of this reduction is a direct result of the shorter train separation enabled by radio-based distance-to-go signalling. Additional improvements arise from different rescheduling decisions that occur as an indirect consequence of the shorter train separation. Since the difference in train separation is relatively small, there are mainly different routing decisions, besides the more obvious retimings. Overall, radio-based distance-to-go signalling has a greater impact on conflict detection and resolution in the junction than in the corridor area. This can be attributed to the higher traffic density and more complex track layout.

A secondary point concerns the effect of the conflict detection and resolution model for radio-based distance-to-go signalling with respect to speed profiles. By considering two speed

profile options, one according to the maximum speed and one to the scheduled speed, it is shown that under ETCS Level 2 with TTD, maximum speed profiles can be used less frequently without increasing delays. This reduction in the use of maximum speed profiles could be beneficial for railway operations, particularly in terms of energy consumption.

Specifically related to the modelling, there are implications for the model complexity and computational performance. Due to the addition of speed profile options in the enhancement of the original model, the size of the conflict detection and resolution model, specifically the number of binary variables, has increased. This has an effect on the real-time performance of the enhanced model version with an increase in mean computation time of 42% (junction) and 832% (corridor). However, after the considered run time of one hour, the maximum optimality gap of the nine scenarios that were not solved to optimality is only 0.43%. Moreover, after the real-time computation time of three minutes, the mean optimality gap is still only 1.45%. Hence, the real-time performance of the enhanced model is acceptable.

4. What is the impact of track discretisation granularity on conflict detection and resolution under virtual-block signalling?

In Chapter 4, an impact assessment of the track discretisation granularity on conflict detection and resolution is conducted using the RECIFE-MILP model enhanced for ETCS Level 2 Virtual Block. The analysis is based on two distinct case studies, i.e., a complex, densely operated railway junction and a railway corridor with lower traffic density. Each case study includes 25 random delay scenarios, where 20% of the trains enter the area with a delay of 5 to 15 minutes. The track discretisations considered are derived from the original discretisation in terms of TTD sections and are characterised by maximum section lengths that incrementally double from 50 to 800 metres. Sections whose length exceed the respective maximum length are divided into subsections of equal length. To ensure compliance with safety regulation, sections that contain a switch are excluded.

Model dimensions increase linearly with the number of sections. However, the most significant impact on model complexity arises from the introduction of route-dependent train separation at switches. As a result, the model is rarely solved to optimality within the time limit of one hour for finer discretisations and the junction case study specifically. Despite the more frequent suboptimality for the finer discretisations, the corridor case study features mostly stable or improved objective values (up to 4.17%). In this case, the improvements are solely due to shorter train separation distances enabled by shorter sections, with no apparent impact of the track discretisation on rescheduling solutions. For the junction case study, effects of the track discretisation on conflict detection and resolution can be observed. While most improvement in the objective function is due to shorter train separation resulting from the shorter sections, different discretisations lead to distinct rescheduling decisions in some delay scenarios. These differences primarily involve a limited number of route assignments, though occasionally a difference in train ordering occur.

Overall, maximum section lengths of 400 or 200 metres can offer a good balance between solution quality and computational complexity for virtual-block signalling. Focusing on the rescheduling decisions, the 800-metre discretisation already performs well, providing stable and reliable solutions, particularly for the corridor case study. In general, the appropriate discretisation threshold strongly depends on the specific case study and track layout.

5. How can railway traffic management be supported in mitigating hazardous traffic conditions under radio-based distance-to-go signalling with onboard train integrity monitoring?

In Chapter 5, an early-warning hazard prediction framework is proposed to support railway traffic management in mitigating hazardous traffic conditions under radio-based distance-to-go signalling with onboard train positioning and train integrity monitoring through Global Navigation Satellite Systems (GNSS). The early-warning framework extends core traffic management functionalities, i.e., traffic state monitoring and prediction, and conflict detection and resolution, by integrating an early-warning mechanism that predicts potentially hazardous traffic conditions, such as conflicting operations that may lead to enforced braking by the signalling system.

To enable timely warnings, the framework relies on the monitoring and prediction of the values of signalling parameters relevant to onboard train positioning and train integrity monitoring, including GNSS positioning errors and radio communication delays. These parameters directly influence the train blocking time, i.e., the period during which a track section is assigned to a specific train and hence blocked for other trains. When the parameter values exceed predefined critical thresholds, they are forwarded to the conflict detection and resolution module. If a conflict is detected, the conflict resolution outcome is combined with the parameter values to generate an early warning for traffic management operators.

In the application of the framework to a use case involving atmospheric disturbances leading to degraded GNSS and radio signals, the early detection of potentially conflicting operations due to increased positioning errors and communication delays enabled traffic management to respond proactively, thereby preventing further compromise of operational efficiency.

6.2 Recommendations for practice

This thesis has led to several recommendations that can be applied in railway traffic management practice and beyond. The main recommendations are as follows:

- **Consideration of speed-dependent train separation in rescheduling distance-to-go operations.** The comparative analysis between the conventional fixed-block and fixed-block distance-to-go variants of a conflict detection and resolution model, as presented in Chapter 3, showed the added value of modelling speed-dependent train separation under distance-to-go signalling featuring radio-based cab signalling and continuous braking curve supervision in terms of rescheduling decisions. In addition, the conflict detection and resolution model developed in Chapters 3 and 4 can be implemented within a decision support system for dispatchers managing railway operations in various distance-to-go signalling systems featuring radio-based cab signalling and continuous braking curve supervision. Already applied to ETCS Level 2 with Trackside Train Detection and ETCS Level 2 Virtual Block, the model's flexibility in terms of track discretisation also enables it to approximate moving-block operations and be extended to accommodate the intermediate solution of ETCS Hybrid Train Detection. Due to the current lack of such decision support systems, initial implementation steps should focus on simulation environments to enable testing and further improvement of the model.

- **Identification of track discretisation granularity balancing rescheduling quality and efficiency.** The impact assessment of track discretisation granularity on conflict detection and resolution, as presented in Chapter 4, provides practical insights for the design and implementation of virtual-block as well as hybrid train detection signalling systems. Using the conflict detection and resolution model for virtual-block signalling from Chapter 4, it is possible to compare different virtual discretisation schemes and evaluate their effects on the performance of railway traffic management in terms of rescheduling decisions. Based on such a comparison, appropriate track discretisation granularities can be selected for developing models to support railway traffic management offering the best trade-off between solution quality and efficiency.

- **Inclusion of early-warning hazard prediction in traffic management support.** The early-warning hazard prediction framework presented in Chapter 5 can be implemented to support dispatchers in mitigating hazardous traffic conditions under radio-based distance-to-go railway signalling. When integrated into the broader railway traffic management framework, the early-warning hazard prediction module can help maintain operational efficiency and safety by enabling dispatchers to proactively respond to the early identification of hazardous situations. While originally proposed for radio-based distance-to-go signalling systems with GNSS-based train integrity monitoring, the framework is also applicable to other radio-based distance-to-go signalling systems, as it accounts for both positioning errors and communication delays in general. Given the current state-of-practice in railway signalling, another important focus should be the predictive modelling of these parameters in the context of early-warning hazard prediction for both current and future signalling system variants.

6.3 Recommendations for future research

Based on the findings and limitations of the work presented in this thesis, the following directions are recommended for future research on models to support for railway traffic management under radio-based distance-to-go signalling:

- **Comparative analysis of discrete and continuous modelling approaches.** Several points of the research agenda proposed in Chapter 2 remain open. Notably, this includes the development and evaluation of modelling approaches that represent the infrastructure and/or speed in a continuous, rather than a discrete, manner. As a next step, future research should focus on comparing discrete and continuous modelling approaches to gain further insights for the development of effective conflict detection and resolution models for moving-block signalling.

- **Investigation of decomposition approaches.** Future research should explore the use of decomposition methods in both the problem formulation and the solution approach for conflict detection and resolution for radio-based distance-to-go signalling. Decomposition approaches that may be considered include spatial decomposition, e.g., geographical, temporal decomposition, e.g., rolling horizon, and hierarchical decomposition, e.g., decentralised. The effectiveness of a decomposition approach is likely to depend on case-specific factors such as network layout, traffic density and operational priorities. A trade-

off can be expected between solution quality and computational efficiency, especially in the context of real-time applications.

- **Exploration of alternative track discretisation methods.** In future research into the impact of track discretisation on conflict detection and resolution for railway operations under virtual-block signalling, various track discretisation approaches should be evaluated. In this thesis, discretisation was based on the existing trackside train detection sections, which may not provide an optimal balance between solution quality and computational efficiency. Future work could explore alternative discretisation methods such as a uniform discretisation, which divides the track into (virtual) sections of equal length, or a criticality-based discretisation, where a finer resolution is applied near operationally significant points, i.e., locations such as switches and stations where conflicts are most likely to occur and rescheduling measures most frequently implemented.

- **Validation of early-warning hazard prediction framework.** Next research steps for the early-warning hazard prediction framework should focus on its validation against real-life data. To support a wider range of applications in railway traffic management under radio-based signalling with train integrity monitoring, this validation should include integrating methods to predict the availability and accuracy of GNSS and radio signals under degraded conditions.

Bibliography

Akinwole, B. H., & Yussuf, A. I. O. (2023). Investigating the Impacts of Tropospheric Parameters on Received Signal Strength of the Mobile Communication System. *Annals of Science and Technology*, 8, 63–73.

Aoun, J., Goverde, R. M. P., Nardone, R., Quaglietta, E., & Vittorini, V. (2024). Analysis of safe and effective next-generation rail signalling systems. *Transportation Research Part C: Emerging Technologies*, 162, 104573.

Bazaraa, M. S., Jarvis, J. J., & Sherali, H. D. (2008). *Linear programming and network flows*. John Wiley & Sons.

Benders, J. F. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4, 238-252.

Bettinelli, A., Santini, A., & Vigo, D. (2017). A real-time conflict solution algorithm for the train rescheduling problem. *Transportation Research Part B: Methodological*, 106, 237-265.

Blatnik, A., & Batagelj, B. (2025). Evaluating GNSS Receiver Resilience: A Study on Simulation Environment Repeatability. *Electronics*, 14(9), 1797.

Brünger, O., & Dahlhaus, E. (2014). Running Time Estimation. In: Hansen, I. A., & Pachl, J., eds., *Railway Timetabling and Operations: Analysis, Modelling, Optimisation, Simulation, Performance Evaluation*, Eurailpress.

Büker, T., Graffagnino, T., Hennig, E., & Kuckelberg, A. (2019). Enhancement of Blocking-time Theory to Represent Future Interlocking Architectures. *RailNorrköping 2019, 8th International Conference on Railway Operations Modelling and Analysis (ICROMA), Norrköpping, Sweden*.

Cacchiani, V., Huisman, D., Kidd, M., Kroon, L., Toth, P., Veelenturf, L., & Wagenaar, J. (2014). An Overview of Recovery Models and Algorithms for Real-time Railway Rescheduling. *Transportation Research Part B: Methodological*, 63, 15–37.

Caimi, G., Fuchsberger, M., Laumanns, M., & Lüthi, M. (2012). A model predictive control approach for discrete-time rescheduling in complex central railway station areas. *Computers & Operations Research*, 39(11), 2578–2593.

Corman, F., D'Ariano, A., Pacciarelli, D., & Pranzo, M. (2009). Evaluation of green wave policy in real-time railway traffic management. *Transportation Research Part C: Emerging Technologies*, 17(6), 607–616.

Corman, F., & Meng, L. (2015). A Review of Online Dynamic Models and Algorithms for Railway Traffic Management. *IEEE Transactions on Intelligent Transportation Systems*, 16(3), 1274–1284.

Cuppi, F., Vignali, V., Lantieri, C., Rapagnà, L., Dimola, N., & Galasso, T. (2021). High density European Rail Traffic Management System (HD-ERTMS) for urban railway nodes: The case study of Rome. *Journal of Rail Transport Planning & Management*, 17, 100232.

D'Ariano, A., Corman, F., Pacciarelli, D., & Pranzo, M. (2008). Reordering and Local Rerouting Strategies to Manage Train Traffic in Real Time. *Transportation Science*, 42(4), 405–419.

D'Ariano, A., Pacciarelli, D., & Pranzo, M. (2007a). A branch and bound algorithm for scheduling trains in a railway network. *European Journal of Operational Research*, 183(2), 643–657.

D'Ariano, A., Pranzo, M., & Hansen, I. A. (2007b). Conflict Resolution and Train Speed Coordination for Solving Real-Time Timetable Perturbations. *IEEE Transactions on Intelligent Transportation Systems*, 8(2), 208–222.

Davis, W. J. (1926). The tractive resistance of electric locomotives and cars. *General Electric Review*, 29, 685–707.

EEIG ERTMS Users Group (2024). *ERTMS/ETCS Hybrid Train Detection: Principles*. https://ertms.be/wp-content/uploads/2023/06/16E0421F_HTD.pdf.

European Environment Agency (2021). *Passenger transport demand*. https://www.eea.europa.eu/data-and-maps/indicators/passenger-transport-demand-version-2/assessment-9.

European Rail Supply Industry Association (2022). *ERTMS*. https://www.ertms.net/.

European Rail Supply Industry Association (2024). *ERTMS/ETCS Levels*. https://www.ertms.net/wp-content/uploads/2024/03/ERTMSETCS-Levels-Updated-2024-Edition.pdf.

European Rail Supply Industry Association (2025). *ERTMS Signaling levels*. https://www.ertms.net/ertms-signaling-levels/.

European Union Agency for Cybersecurity (2024). *Telecom Security Incidents 2022: Annual Report*. https://www.enisa.europa.eu/sites/default/files/2024-11/ENISA_Telecom%20Incident%20Reporting_en_1.pdf.

European Union Agency for Railways (2016). *ERTMS/ETCS System Requirements Specification (Subset-026)*. https://www.era.europa.eu/era-folder/archived-set-specifications-3-etcs-b3-r2-gsm-r-b1.

European Union Agency for Railways (2020). *Introduction to ETCS Braking Curves*. https://www.era.europa.eu/system/files/2022-11/Introduction%20to%20ETCS%20braking%20curves.pdf.

Furness, N., Van Houten, H., Arenas, L., & Bartholomeus, M. (2017). ERTMS Level 3: the Game-Changer. *IRSE News*, 232, 2–9.

Gao, H., Zhang, Y., & Guo, J. (2020). Calculation and Optimization of Minimum Headway in Moving Block System. *2020 IEEE 5th International Conference on Intelligent Transportation Engineering (ICITE), Beijing, China*.

Gonzalez, J., Rodriguez, C., Blanquer, J., Mera, J. M., Castellote, E., & Santos, R. (2010). Increase of metro line capacity by optimisation of track circuit length and location: In a distance to go system. *Journal of Advanced Transportation*, 44(2), 53–71.

GSM-R Operators Group (2015). *European Integrated Railway Radio Enhanced Network System Requirements Specification Version 16.0.0*. https://uic.org/IMG/pdf/srs-16.0.0_uic_951-0.0.2_final.pdf.

Hansen, I. A., & Pachl, J., eds. (2014). *Railway Timetabling & Operations: Analysis, Modelling, Optimisation, Simulation, Performance Evaluation*. Eurailpress.

Himrane, O., Beugin, J., & Ghazel, M. (2023). Implementation of a Model-Oriented Approach for Supporting Safe Integration of GNSS-Based Virtual Balises in ERTMS/ETCS Level 3. *IEEE Open Journal of Intelligent Transportation Systems*, 4, 294–310.

Institute of Electrical and Electronics Engineers (2025). *IEEE Standard for Communications-Based Train Control (CBTC) Performance and Functional Requirements*. https://standards.ieee.org/ieee/1474.1/6959/.

International Union of Railways (2024). *FRMCS: Future Railway Mobile Communication System*. https://uic.org/rail-system/telecoms-signalling/frmcs.

Jansen, J. M. (2019). *ERTMS/ETCS Hybrid Level 3: A simulation-based impact assessment for the Dutch railway network*, Master's thesis, Delft University of Technology, the Netherlands.

Janssens, M. L. (2022). *Multi machine approaches for conflict resolution under moving block signalling*, Master's thesis, Delft University of Technology, the Netherlands.

Jeffrey, C., & Munro, R. (2023). Chapter 4: GNSS error sources. In: *An Introduction to GNSS: A Primer in Using Global Navigation Satellite Systems for Positioning and Autonomy (3rd edition)*, Hexagon.

Knutsen, O., Olsson, N. O. E., & Fu, J. (2024). Capacity evaluation of ERTMS/ETCS hybrid level 3 using simulation methods. *Journal of Rail Transport Planning & Management*, 30, 100444.

Lamorgese, L., & Mannino, C. (2015). An Exact Decomposition Approach for the Real-Time Train Dispatching Problem. *Operations Research*, 63(1), 48-64.

Lamorgese, L., & Mannino, C. (2019). A Noncompact Formulation for Job-Shop Scheduling Problems in Traffic Management. *Operations Research*, 67(5), 1586–1609.

Lamorgese, L., Mannino, C., Pacciarelli, D., & Tönquist Krasemann, J. (2018). Train dispatching. In: Borndörfer, R., Klug, T., Lamorgese, L., Maninno, C., Reuther, M., & Schlechte, T., eds., *Handbook of Optimization in the Railway Industry*, Springer.

Lamorgese, L., Mannino, C., & Piacentini, M. (2016). Optimal Train Dispatching by Benders'-Like Reformulation. *Transportation Science*, 50(3), 910–925.

Lazarescu, M. T., & Poolad, P. (2021). Asynchronous resilient wireless sensor network for train integrity monitoring. *IEEE Internet of Things Journal*, 8(5), 3939–3954.

Leutwiler, F., & Corman, F. (2022). A logic-based Benders decomposition for microscopic railway timetable planning. *European Journal of Operational Research*, 303(2), 525–540.

Leutwiler, F., & Corman, F. (2023). A review of principles and methods to decompose largescale railway scheduling problem. *EURO Journal on Transportation and Logistics*, 12, 100107.

Lippes, S. J. H. (2024). *Distributed Rail Traffic Management under Moving-Block Signalling*, Master's thesis, Delft University of Technology, the Netherlands.

Liu, F., Xun, J., Liu, R., Yin, J., & Dong, H. (2021). A Real-Time Rescheduling Approach by Using Loop Iteration for High-Speed Railway Traffic. *IEEE Intelligent Transportation Systems Magazine*, 15(1), 318–332.

Liu, R. (2016). Simulation Model of Speed Control for the Moving-Block Systems under ERTMS Level 3. *2016 IEEE International Conference on Intelligent Rail Transport (ICIRT), Birmingham, UK*.

Lövétei, I. F., Kővári, B., & Bécsi, T. (2021). MCTS Based Approach for Solving Real-time Railway Rescheduling Problem. *Periodica Polytechnica Transportation Engineering*, 49(3), 283–291.

Luan, X., Wang, Y., De Schutter, B., Meng, L., Lodewijks, G., & Corman, F. (2018). Integration of real-time traffic management and train control for rail networks - Part 1: Optimization problems and solution approaches. *Transportation Research Part B: Methodological*, 115, 41–71.

Lusby, R. M., Larsen, J., Ehrgott, M., & Ryan, D. M. (2013). A set packing inspired method for real-time junction train routing. *Computers & Operations Research*, 40(3), 713–724.

Mannino, C. (2011). Real-time traffic control in railway systems. *11th Symposium on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2011), Saarbrücken, Germany*.

Mannino, C., & Mascis, A. (2009). Optimal real-time traffic control in metro stations. *Operations Research*, 57(4), 1026–1039.

Marcelli, E., & Pellegrini, P. (2021). Literature review toward decentralized railway traffic management. *IEEE Intelligent Transportation Systems Magazine*, 3(3), 234–252.

Marlière, G., Sobieraj Richard, S., Pellegrini, P., & Rodriguez, J. (2023). A conditional time-intervals formulation of the real-time Railway Traffic Management Problem. *Control Engineering Practice*, 54(2), 187–194.

Martinez, L., & Martin, U. (2020). Terminology, Differences, and Challenges of Communications-based Train Control and European Train Control Systems. *WIT Transactions on The Built Environment*, 199, 15–26.

Mascis, A., & Pacciarelli, D. (2002). Job-shop scheduling with blocking and no-wait constraints. *European Journal of Operational Research*, 143(3), 498–517.

Mayolle, Q., Marais, J., Fasquelle, M., Tardif, V., & Chéneau-Grehalle, E. (2025). GNSS-Based Solutions Testing in an ERTMS Context: A Framework for Statistical Performance Analysis. *2025 IEEE/ION Position, Location and Navigation Symposium (PLANS), Salt Lake City, UT, USA*.

Mazini, A., Chen, L., Kirkwood, D., Lyu, X., & Garcia-Fernandez, M. (2022). Co-simulation platform for demonstration and testing of moving block systems. *34th European Modelling and Simulation Symposium (EMSS 2022), Rome, Italy*.

Mazzarello, M., & Ottaviani, E. (2007). A traffic management system for real-time traffic optimisation in railways. *Transportation Research Part B: Methodological*, 41(2), 246–274.

Meng, L., & Zhou, X. (2011). Robust single-track train dispatching model under a dynamic and stochastic environment: A scenario-based rolling horizon solution approach. *Transportation Research Part B: Methodological*, 45(7), 1080–1102.

Mera, J. M., Carabano, E., Soler, M., & Castellote, E. (2016). Increasing metro line capacity by optimisation of track circuit in a speed code Automatic Train Protection system. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 230(1), 165–180.

Meunier, H., Baro, S., Borodin, V., Dauzére-Pérès, S., & Pochet, J. (2021). A Microscopic Modeling Approach for Real-Time Train Retiming under Disturbances for a CBTC Suburban Railway Line. *RailBeijing 2021, 9th International Conference on Railway Operations Modelling and Analysis (ICROMA), Beijing, China*.

NetworkRail (2015). *RIDC Melton Site Layout*. https://www.networkrail.co.uk/wp-content/uploads/2019/05/RIDC-Melton-Site-Layout.pdf.

Pellegrini, P., Marlière, G., Pesenti, R., & Rodriguez, J. (2015). RECIFE-MILP: An Effective MILP-Based Heuristic for the Real-Time Railway Traffic Management Problem. *IEEE Transactions on Intelligent Transportation Systems*, 16(5), 2609–2619.

Pellegrini, P., Marlière, G., & Rodriguez, J. (2014). Optimal train routing and scheduling for managing traffic perturbations in complex junctions. *Transportation Research Part B: Methodological*, 59, 58–80.

Pellegrini, P., Marlière, G., & Rodriguez, J. (2016). A detailed analysis of the actual impact of real-time railway traffic management optimization. *Journal of Rail Transport Planning & Management*, 6(1), 13–31.

Pellegrini, P., Pesenti, R., & Rodriguez, J. (2019). Efficient train re-routing and rescheduling: Valid inequalities and reformulation of RECIFE-MILP. *Transportation Research Part B: Methodological*, 120, 33–48.

PERFORMINGRAIL (2021). *Deliverable D3.1: Design Document of the Location Algorithms*. https://projects.shift2rail.org/s2r_ip2_n.aspx?p=S2R_PERFORMINGRAIL.

PERFORMINGRAIL (2022a). *Deliverable D1.1: Baseline System Specification and Definition for Moving Block Systems*. https://projects.shift2rail.org/s2r_ip2_n.aspx?p=S2R_PERFORMINGRAIL.

PERFORMINGRAIL (2022b). *Deliverable D4.1: Real-Time Traffic Rescheduling Algorithms for Perturbation Management and Hazard Prevention in Moving-Block Operations*. https://projects.shift2rail.org/s2r_ip2_n.aspx?p=S2R_PERFORMINGRAIL.

PERFORMINGRAIL (2023). *Deliverable D3.4: Location Algorithm Validation Report*. https://projects.shift2rail.org/s2r_ip2_n.aspx?p=S2R_PERFORMINGRAIL.

Pochet, J., Baro, S., & Sandou, G. (2016). Supervision and rescheduling of a mixed CBTC traffic on a suburban railway line. *2016 IEEE International Conference on Intelligent Rail Transportation (ICIRT), Birmingham, UK*.

Pochet, J., Baro, S., & Sandou, G. (2017). Automatic train supervision for a CBTC suburban railway line using multiobjective optimization. *2017 IEEE International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan*.

Quaglietta, E., Pellegrini, P., Goverde, R. M. P., Albrecht, T., Jaekel, B., Marlière, G., Rodriguez, J., Dollevoet, T., Ambrogio, B., Carcasole, D., Giaroli, M., & Nicholson, G. (2016). The ON-TIME real-time railway traffic management framework: A proof-of-concept using a scalable standardised data communication architecture. *Transportation Research Part C: Emerging Technologies*, 63, 23–50.

Ranjbar, V., Olsson, N. O. E., & Sipilä, H. (2022). Impact of signalling system on capacity – Comparing legacy ATC, ETCS Level 2 and ETCS Hybrid Level 3 systems. *Journal of Rail Transport Planning & Management*, 23, 100322.

Reynolds, E., Ehrgott, M., Maher, S. J., Patman, A., & Wang, J. Y. T. (2020). A multicommodity flow model for rerouting and retiming trains in real-time to reduce reactionary delay in complex station areas. *Optimization Online*.

Reynolds, E., & Maher, S. J. (2022). A data-driven, variable-speed model for the train timetable rescheduling problem. *Computers & Operations Research*, 142, 105719.

Rodriguez, J. (2007). A constraint programming model for real-time train scheduling at junctions. *Transportation Research Part B: Methodological*, 41(2), 231–245.

Saastamoinen, J. (1972). Atmospheric Correction for the Troposphere and Stratosphere in Radio Ranging Satellites. *The Use of Artificial Satellites for Geodesy*, 215, 247–251.

Samà, M., Meloni, C., D'Ariano, A., & Corman, F. (2015). A multi-criteria decision support methodology for real-time train scheduling. *Journal of Rail Transport Planning & Management*, 5(3), 146–162.

Schlechte, T., Borndörfer, R., Denißen, J., Heller, S., Klug, T., Küpper, M., Lindner, N., Reuther, M., Söhlke, A., & Steadman, W. (2022). Timetable Optimization for a Moving Block System. *Journal of Rail Transport Planning & Management*, 22, 100315.

Schön, W., Larraufie, G., Moëns, G., & Poré, J., eds. (2013). *Signalisation et automatismes ferroviaires - Railway signalling and automation*. La Vie du Rail.

Scopus (2023). *Scopus*. https://www.scopus.com.

Shift2Rail (2023). *PERFORMINGRAIL*. https://projects.shift2rail.org/s2r_ip2_n.aspx?p=S2R_PERFORMINGRAIL.

Shift2Rail (2024). *IP2 projects*. https://projects.shift2rail.org/s2r_ip.aspx?ip=2.

Sizun, H. (2005). *Radio Wave Propogation for Telecommunication Applications*. Springer.

Space Weather Prediction Center (2025). *Ionospheric Scintillation*. https://www.swpc.noaa.gov/phenomena/ionospheric-scintillation.

Theeg, G., & Vlasenko, S., eds. (2020). *Railway Signalling and Interlocking*. PMC Media.

Törnquist, J. (2012). Design of an effective algorithm for fast response to the rescheduling of railway traffic during disturbances. *Transportation Research Part C: Emerging Technologies*, 20, 62–78.

Törnquist, J., & Persson, J. A. (2007). N-tracked railway traffic re-scheduling during disturbances. *Transportation Research Part B: Methodological*, 41(3), 342–362.

Van den Akker, J. M., Hurkens, C. A. J., & Savelsbergh, M. W. P. (2000). Time-Indexed Formulations for Machine Scheduling Problems: Column Generation. *INFORMS Journal on Computing*, 12(2), 111–124.

Vergroesen, R. (2020). *ERTMS/ETCS Hybrid Level 3 and ATO: A simulation-based capacity impact study for the Dutch railway network*, Master's thesis, Delft University of Technology, the Netherlands.

Versluis, N. D., Pellegrini, P., Quaglietta, E., Goverde, R. M. P., & Rodriguez, J. (2023). An Approximate Conflict Detection and Resolution Model for Moving-Block Signalling by Enhancing RECIFE-MILP. *RailBelgrade 2023, 10th International Conference on Railway Operations Modelling and Analysis (ICROMA), Belgrade, Serbia*.

Versluis, N. D., Pellegrini, P., Quaglietta, E., Goverde, R. M. P., & Rodriguez, J. (2025a). Conflict detection and resolution for distance-to-go railway signalling. *Transportmetrica A: Transport Science*, in press.

Versluis, N. D., Pellegrini, P., Quaglietta, E., Goverde, R. M. P., & Rodriguez, J. (2025b). Impact of track discretisation on conflict detection and resolution under ETCS with onboard train integrity monitoring. *Journal of Rail Transport Planning & Management*, 35, 100533.

Versluis, N. D., Quaglietta, E., Goverde, R. M. P., Pellegrini, P., & Rodriguez, J. (2024). Real-time railway traffic management under moving-block signalling: A literature review and research agenda. *Transportation Research Part C: Emerging Technologies*, 158, 104438.

Wang, P., & Goverde, R. M. P. (2019). Multi-train trajectory optimization for energy-efficient timetabling. *European Journal of Operational Research*, 272, 621–635.

Wang, Y., De Schutter, B., Van Den Boom, T. J. J., & Ning, B. (2014). Optimal trajectory planning for trains under fixed and moving signaling systems using mixed integer linear programming. *Control Engineering Practice*, 22(1), 44–56.

X2Rail-2 (2020). *Deliverable D4.2: Functional architecture & Interfaces specifications & Candidate technologies selection*. https://projects.shift2rail.org/s2r_ip2_n.aspx?p=X2RAIL-2.

X2Rail-3 (2020). *Deliverable D4.2: Moving Block Specifications*. https://projects.shift2rail.org/s2r_ip2_n.aspx?p=X2RAIL-3.

X2Rail-4 (2022). *Deliverable D7.3: Standardisation Proposal*. https://projects.shift2rail.org/s2r_ip2_n.aspx?p=X2RAIL-4.

Xu, G., & Xu, Y. (2016). Physical Influences of GPS Surveying. In: *GPS: Theory, Algorithms and Applications*, Springer.

Xu, L., Zhao, X., Tao, Y., Zhang, Q., & Liu, X. (2014). Optimization of Train Headway in Moving Block Based on a Particle Swarm Optimization Algorithm. *13th International Conference on Control, Automation, Robotics and Vision (ICARCV 2014), Singapore*.

Xu, P., Corman, F., Peng, Q., & Luan, X. (2017). A train rescheduling model integrating speed management during disruptions of high-speed traffic under a quasi-moving block system. *Transportation Research Part B: Methodological*, 104, 638–666.

Xu, P., Zhang, D., Guo, J., Liu, D., & Peng, H. (2021). Integrated Train Rescheduling and Rerouting during Multidisturbances under a Quasi-Moving Block System. *Journal of Advanced Transportation*, 6652531.

Yi, X., Marlière, G., Pellegrini, P., Rodriguez, J., & Pesenti, R. (2023). Coordinated train rerouting and rescheduling in large infrastructures. *Transportation Research Procedia*, 72, 319–326.

Zhan, S., Wong, S. C., Shang, P., & Lo, S. M. (2022). Train rescheduling in a major disruption on a high-speed railway network with seat reservation. *Transportmetrica A: Transport Science*, 18(13), 532–567.

Zhang, S., Cheng, Y., Chen, K., Ma, C., Wei, J., & Hu, X. (2024). A general metro timetable rescheduling approach for the minimisation of the capacity loss after random line disruption. *Transportmetrica A: Transport Science*, 20(3), 2204965.

ZhG, R. (2021). Entanglement difference of GNSS carrier phase for vehicle attitude determination. *International Journal of Transportation Science and Technology*, 10(1), 69–82.

Zhu, N., Marais, J., Bétaille, D., & Berbineau, M. (2018). GNSS position integrity in urban environments: A review of literature. *IEEE Transactions on Intelligent Transportation Systems*, 19(9), 2762–2778.

Zwaneveld, P. J., Kroon, L. G., & Van Hoesel, S. P. M. (2001). Routing trains through a railway station based on a node packing model. *European Journal of Operational Research*, 128(1), 14–33.

English summary

Railways are a safe and efficient mode of transport, supported by railway signalling and railway traffic management. Responsible for safe route setting and train separation, the railway signalling system directly influences the capacity of the railway network. Conventional fixed-block multi-aspect systems, in which trains are separated by a number of blocks with fixed lengths based on worst-case braking distances, cannot fulfil the forecasted increase in railway demand. Radio-based distance-to-go signalling systems such as the European Train Control System (ETCS) Level 2 are being developed to increase the capacity of the existing railway network. In these systems, train separation is dynamically determined based on train- and speed-dependent braking distances. This is enabled by radio-based cab signalling and continuous braking curve supervision. In state-of-practice ETCS Level 2 variants, train integrity is monitored by the track through sections in which the absence of trains is automatically detected. Next-generation ETCS Level 2 variants instead feature onboard train integrity monitoring (TIM), for example based on satellite navigation. TIM complements the existing onboard train positioning, enabling the replacement of trackside train detection. The transition to radio-based distance-to-go signalling affects the operational principles and rules for railway traffic.

To maintain efficient operations, railway traffic management is responsible for implementing rescheduling measures in case of delays, for example by updating arrival and departure times, and orders and/or routes of trains. For the case that delays lead to track conflicts, i.e., when more than one train requires the same part of the track during overlapping time intervals, conflict detection and resolution models have been developed to support human dispatchers in taking mathematically optimised rescheduling decisions. However, the vast majority of available conflict detection and resolution models refers to conventional signalling systems rather than to radio-based distance-to-go signalling systems. To address this gap, this thesis proposes models for the effective management of real-time railway operations under radio-based distance-to-go signalling.

First, research steps are proposed to bridge existing gaps and challenges in the modelling of real-time railway traffic management for radio-based distance-to-go signalling. Specifically, conflict detection and resolution models available in literature are reviewed in light of their applicability to radio-based distance-to-go systems, with and without onboard TIM. The main gap is identified to be the lack of conflict detection and resolution models that accurately describe railway operations considering the characteristic train separation based on speed-dependent braking distances. Therefore, a key challenge is the integration of speed dynamics into these models. This necessitates a more continuous representation of the railway track, which depends on whether onboard TIM is implemented or not. Based on the presented analysis, various options for track and speed modelling are proposed to address the identified gaps, followed by an assessment of modelling approaches for the different options. Consistent

with the proposal, future research should aim at the further investigation of the possibilities of existing approaches in terms of track discretisation and speed alternatives. A comparison of the different modelling options is needed to determine the extent to which continuous modelling of track and speed brings advantages over discretised models in terms of solution quality and computational efficiency.

Second, an approach is proposed to enhance existing conflict detection and resolution models to describe radio-based distance-to-go railway operations. The main enhancements are: 1) a distinction between switches and open track, allowing for the additional safety protection required at switches while representing the open line more continuously, 2) alternative speed profiles based on the scheduled and the maximum allowable speed, where the latter can be assigned to support delay recovery, and 3) train separation based on speed-dependent braking distances, with brake indications points located anywhere along the track. The enhancement approach is demonstrated by applying it to a state-of-the-art model, originally developed for conventional signalling. For two case studies, representing a railway corridor and a complex railway junction with mixed traffic in France, it is established that the enhanced model outperforms the original model. The improvement is reflected in both a lower total train delay and fewer assignments of speed profiles corresponding to the maximum allowable speed. In general, the reduction in total delay is limited but for several case study instances it is significant with improvements up to 7%. Most of the reduction originates from the differences in the modelled signalling systems, predominately the shorter train separation. However, additional gain is achieved by proposing different rescheduling decisions. Due to the inclusion of a penalty for the assignment of maximum speed profiles, trains are less often required to run at maximum speed. With that, a secondary benefit in terms of energy consumption is obtained. Given the insights in the added value of modelling speed-dependent train separation for radio-based distance-to-go signalling system, it is recommended to implement the enhanced model in decision support systems for railway traffic management under radio-based distance-to-go signalling.

Third, an impact assessment is carried out of the track discretisation on the conflict detection and resolution model for radio-based distance-to-go signalling. First, the earlier introduced model is updated to capture virtual-block signalling by considering onboard TIM parameters and introducing a procedure to discretise the track into virtual sections derived from the original track partitioning into trackside train detection sections. The considered discretisation granularities are characterised by maximum section lengths in the range of 50 to 800 metres. The model dimension grows linearly with the number of sections. The main increase is, however, due to a general update of the model to reformulate the train separation at switches. The increasing model size leads to frequent suboptimality of the obtained solutions. Nevertheless, for the corridor case study, non-increasing objective values are obtained (leaving out the finest discretisation of 50 metres). There are no apparent effects of the discretisation on the rescheduling decisions; the observed improvements are solely due to shorter train separation as a result of the shorter sections. For the junction case study, however, some impact of the discretisation on the rescheduling decisions is observed. While most benefit in objective values do originate from the shorter train separation, different rescheduling decisions proved to be better for the various discretisations in certain delay scenarios. Mostly, these different decisions concern a limited number of (re)routing decisions, but occasionally also a (re)ordering decision. The reported results indicate maximum section lengths of 400 or 200 metres as possible thresholds in terms of the balance between solution quality and computational complexity, with the note that it strongly depends on the case study and its track layout. The results regarding the balance

between solution quality and computational efficiency for different track discretisations indicate the importance of identifying a suitable discretisation for the intended application of the developed model.

Fourth, an early-warning hazard prediction framework is proposed for the effective management of railway operations in radio-based distance-to-go signalling systems featuring satellite-based train positioning and TIM. The framework extends core railway traffic management functionalities with an early-warning module that monitors and predicts values of signalling parameters, such as radio communication delays and satellite-based positioning errors. In case a parameter value potentially leads to hazardous traffic conditions, a warning is issued to the dispatcher. Through the application of the early-warning framework to a use case of atmospheric disturbances leading to degraded satellite and radio signals, it is shown that the early detection of potentially hazardous traffic conditions enables dispatchers to act proactively to significantly reduce the impact of increased positioning errors and communication delays on the operations. The presented framework could be implemented in a decision support system for railway traffic management, once radio-based distance-to-go signalling systems with satellite-based TIM are deployed.

In summary, this dissertation provides insights into the support of the real-time management of railway traffic under radio-based distance-to-go signalling systems, both for the scientific community and for the railway industry. With that, it supports the future implementation of radio-based distance-to-go signalling systems such as ETCS Level 2, with or without onboard TIM, contributing to maintaining and further improving the railways as a safe and efficient mode of transport.

Nederlandse samenvatting

Spoorvervoer is een veilige en efficiënte vorm van transport, mede dankzij treinbeveiliging en treinverkeersleiding. Het treinbeveiligingssysteem bepaalt de veilige rijwegen en de onderlinge afstand tussen treinen, en beïnvloedt daarmee de capaciteit van het spoornetwerk. In conventionele vaste-bloksystemen met meerdere seinbeelden wordt de treinafstand bepaald door een aantal blokken met vaste lengtes gebaseerd op de ongunstigste remafstanden. Hierdoor kunnen deze systemen niet voldoen aan de verwachte toename in spoorvervoer. Om de capaciteit van het bestaande netwerk te verhogen, worden radio-gebaseerde distance-to-go (letterlijk: resterende afstand) treinbeveiligingssystemen ontwikkeld, zoals het European Train Control System (ETCS) Niveau 2. In deze systemen worden treinafstanden dynamisch bepaald op basis van trein- en snelheidsafhankelijke remafstanden. ETCS Niveau 2 beschikt over seingeving in de cabine via radiocommunicatie en continue remcurvebewaking. In de huidige implementaties van ETCS Niveau 2 wordt de treinintegriteit gemonitord door baanapparatuur, via secties waarin treinen automatisch worden gedetecteerd. Nieuwe-generatie ETCS Niveau 2 varianten maken daarentegen gebruik van treinintegriteitsmonitoring (TIM) aan boord, bijvoorbeeld op basis van satellietnavigatie. Als aanvulling op de bestaande positiebepaling aan boord, kan met TIM de baangebonden treindetectie worden vervangen. De overgang naar radio-gebaseerde distance-to-go treinbeveiligingssystemen heeft gevolgen voor de operationele principes en regels voor het treinverkeer.

Om het treinverkeer efficiënt te laten verlopen, is de treinverkeersleiding verantwoordelijk voor het aanpassen van de dienstregeling als er vertragingen optreden, bijvoorbeeld door het wijzigen van vertrek- of aankomsttijden en de volgorde en/of de rijwegen van treinen. Door vertragingen kunnen meerdere treinen tegelijkertijd hetzelfde spoor willen gebruiken, waardoor conflicten ontstaan. Om deze conflicten te detecteren en op te lossen, zijn wiskundige modellen ontwikkeld. Deze optimalisatiemodellen kunnen treinverkeersleiders ondersteunen bij het nemen van beslissingen over het aanpassen van de dienstregeling. De meeste van deze conflictdetectie en -oplossingsmodellen zijn echter gebaseerd op conventionele treinbeveiligingssystemen en niet op radio-gebaseerde distance-to-go systemen. Om op deze tekortkoming in te spelen, draagt dit proefschrift modellen aan voor het effectief bijsturen van treinverkeer onder radio-gebaseerde distance-to-go treinbeveiligingssystemen.

Ten eerste zijn onderzoeksstappen opgesteld om bestaande hiaten in het modelleren van treinbijsturing onder radio-gebaseerde distance-to-go treinbeveiliging te overbruggen. De toepasbaarheid van conflictdetectie en -oplossingsmodellen in de literatuur op radio-gebaseerde distance-to-go treinbeveiligingssystemen, met en zonder TIM aan boord, is onderzocht. Het grootste hiaat betreft het ontbreken van modellen die de karakteristieke treinafstanden gebaseerd op snelheidsafhankelijke remafstanden beschrijven. De grootste uitdagingen vormen daarom het modelleren van de snelheidsdynamiek en de vereiste continue representatie van het spoor,

wat afhangt van het al dan niet implementeren van TIM aan boord. Op basis van een analyse zijn modelleringsopties voorgesteld, gevolgd door een beoordeling van modelleringsaanpakken voor deze opties. Gebaseerd op de onderzoeksagenda zou toekomstig onderzoek zich moeten richten op het verder verkennen van de mogelijkheden van bestaande aanpakken met betrekking tot de discretisatie van het spoor en de snelheid. Een vergelijking van deze opties is noodzakelijk om te bepalen in hoeverre continue modellering van spoor en snelheid voordelen biedt ten opzichte van discrete modellen, zowel wat betreft kwaliteit van de oplossingen als rekenkundige efficiëntie.

Ten tweede is een uitbreidingsaanpak gepresenteerd voor bestaande conflictdetectie en -oplossingsmodellen om treinverkeer onder radio-gebaseerde distance-to-go treinbeveiliging correct te beschrijven. De belangrijkste uitbreidingen zijn: 1) een onderscheid tussen wissels en de vrije baan, waardoor de veiligheidseisen voor wissels meegenomen kunnen worden terwijl de vrije baan continu gemodelleerd kan worden, 2) snelheidsprofielvarianten afgeleid van de geplande en de maximaal toegestane snelheid, waarbij de laatste gebruikt kan worden om vertragingen in te halen, en 3) treinafstanden gebaseerd op snelheidsafhankelijke remafstanden, met remindicaties die overal op het spoor gegeven kunnen worden. De uitbreidingen zijn gedemonstreerd door integratie in een conflictdetectie en -oplossingsmodel oorspronkelijk ontworpen voor conventionele treinbeveiliging. Het uitgebreide model is toegepast op twee casussen: een spoorcorridor en een complex spoorwegknooppunt met gemengd verkeer in Frankrijk. In beide gevallen geeft het uitgebreide model betere resultaten dan het originele model, met zowel een lagere totale vertraging als een lager aantal toegewezen snelheidsprofielen voor de maximaal toegestane snelheid. Hoewel de afname in totale vertraging over het algemeen beperkt is, loopt het voor sommige scenario's op tot 7%. Het grootste deel van de afname is toe te schrijven aan de verschillen in de gemodelleerde beveiligingssystemen, voornamelijk vanwege de kortere treinafstanden. Desalniettemin wordt er extra afname gerealiseerd door verschillen in dienstregelingsaanpassingen. Door de invoering van een boete voor het toewijzen van maximale-snelheidsprofielen, hoeven treinen minder vaak met maximale snelheid te rijden, wat een positief effect heeft op het energieverbruik. Daarmee biedt het modelleren van snelheidsafhankelijke treinafstanden toegevoegde waarde voor radio-gebaseerde distance-to-go treinbeveiligingssystemen. Met dit inzicht is het aan te bevelen om het uitgebreide conflictdetectie en -oplossingsmodel te implementeren in beslissingsondersteunende systemen voor het real-time bijsturen van radio-gebaseerd distance-to-go treinverkeer.

Ten derde is het effect geanalyseerd van de discretisatie van het spoor op het conflictdetectie en -oplossingsmodel voor radio-gebaseerde distance-to-go treinbeveiliging. Hiervoor is het eerder geïntroduceerde model verder uitgebreid naar virtuele-blokbeveiliging, waarbij TIM-parameters zijn meegenomen en een procedure is geïntroduceerd om het spoor op te delen in virtuele secties afgeleid van de oorspronkelijke opdeling in treindetectiesecties. De onderzochte discretisaties zijn gekarakteriseerd door maximale sectielengtes tussen de 50 en de 800 meter. De dimensie van het model groeit lineair met het aantal secties, waarbij de meeste groei wordt veroorzaakt door de algemene modelaanpassing waarbij de treinafstand rondom wissels is herzien. De toenemende modelgrootte kan leiden tot suboptimale oplossingen. Desondanks worden er voor de spoorcorridor niet-stijgende doelfunctiewaarden verkregen (de 50-meter discretisatie buiten beschouwing gelaten), terwijl er geen zichtbare effecten van de discretisatie op de dienstregelingsaanpassingen zijn. De geobserveerde verbeteringen zijn uitsluitend toe te schrijven aan de kortere treinafstand door kortere secties. Voor het complexe spoorwegknooppunt met intensief verkeer is daarentegen wel een effect te zien van de discretisatie op

de dienstregelingsaanpassingen. Hoewel de meeste verbetering in doelfunctiewaarden voortkomt uit de kortere treinafstanden, leveren in sommige scenario's andere bijstuurmaatregelen betere resultaten voor de verschillende discretisaties. Deze verschillen zijn veelal een beperkt aantal rijwegaanpassingen, maar ook een enkele treinvolgorde is anders. De 400- en 200-meter discretisaties zijn een mogelijke grens voor de balans tussen de kwaliteit van de oplossing en de rekencomplexiteit, met de kanttekening dat dit sterk afhangt van de casus en de spoorlayout. De uitkomsten wat betreft de balans tussen oplossingskwaliteit en rekenkundige efficiëntie voor discretisatie van het spoor geven aan dat het belangrijk is om de discretisatie van de modellen te optimaliseren.

Ten vierde is een waarschuwings- en risicovoorspellingsframework geïntroduceerd voor de effectieve bijsturing van treinverkeer onder radio-gebaseerde distance-to-go treinbeveilingssystemen met treinpositionering en TIM op basis van satellietnavigatie. Het framework breidt de basisfunctionaliteit van de treinverkeersleiding uit met een module die treinbeveiligingsparameters monitort en voorspelt, zoals vertragingen in radiocommunicatie en positioneringsfouten van satellietnavigatie. Wanneer een parameterwaarde mogelijk tot een risicovolle verkeerssituatie leidt, genereert het system een waarschuwing voor de treinverkeersleider. Aan de hand van een gebruiksscenario met atmosferische verstoringen die leiden tot verzwakte satelliet- en radiosignalen, is aangetoond dat vroegtijdige detectie van potentieel risicovolle verkeerssituaties de treinverkeersleider in staat stelt proactief te handelen. Op deze manier kan de impact van toegenomen positioneringsfouten en vertaagde communicatie op het treinverkeer aanzienlijk worden beperkt. Het gepresenteerde framework zou geïmplementeerd kunnen worden ter ondersteuning van de treinverkeersleiding, als radio-gebaseerde distance-to-go treinbeveilingssystemen met satelliet-gebaseerde TIM worden uitgerold.

Samenvattend biedt dit proefschrift inzichten in de ondersteuning van het bijsturen van treinverkeer onder radio-gebaseerde distance-to-go treinbeveiligingssystemen. Deze inzichten zijn relevant voor zowel de wetenschappelijke gemeenschap als de spoorwegsector, en ondersteunen de toekomstige implementatie van radio-gebaseerde distance-to-go beveiligingssystemen zoals ETCS Niveau 2, met en zonder TIM aan boord. Daarmee dragen ze bij aan het behoud en de verdere verbetering van het spoorvervoer als een veilige en efficiënte vorm van transport.

About the author

Nina Versluis (1995) was born and raised in Amstelveen, the Netherlands. She received both her bachelor's (2017) and master's degree (2020) in Applied Mathematics from Delft University of Technology, with a specialisation in optimisation. During her studies, Nina worked as a teaching assistant and did internships at the Nederlandse Spoorwegen (NS) and TBA Group.

In 2021, Nina joined the Department of Transport and Planning at Delft University of Technology to start her PhD research on 'optimal real-time traffic management of next-generation railway operations'. Her research was funded by the Shift2Rail project PERFORMINGRAIL, to which she contributed through several project tasks and deliverables. As part of her PhD, she spent ten months in Lille, France, as a visiting researcher to the COSYS-ESTAS department of Gustave Eiffel University.

Throughout her PhD, Nina was actively involved in department activities. She co-organised monthly department and PhD meetings (2021-2022) and acted as the PhD representative in the department management team and the well-being task force (2023-2025). She also contributed to education by supervising bachelor's and master's graduation projects.

After her PhD, Nina continued working at the Department of Transport and Planning, as a post-doctoral researcher involved in the Europe's Rail project Pods4Rail.

Publications

Journal articles

Versluis, N. D., Quaglietta, E., Garcia-Fernandez, M., & Goverde, R. M. P. (in press). Early-warning hazard prediction for railway traffic management under radio-based signalling with satellite-based train integrity monitoring. *International Journal of Transportation Science and Technology*.

Versluis, N. D., Pellegrini, P., Quaglietta, E., Goverde, R. M. P., & Rodriguez, J. (in press). Conflict detection and resolution for distance-to-go railway signalling. *Transportmetrica A: Transport Science*.

Versluis, N. D., Pellegrini, P., Quaglietta, E., Goverde, R. M. P., & Rodriguez, J. (2025). Impact of track discretisation on conflict detection and resolution under ETCS with onboard train integrity monitoring. *Journal of Rail Transport Planning & Management*, 35, 100533.

Versluis, N. D., Quaglietta, E., Goverde, R. M. P., Pellegrini, P., & Rodriguez, J. (2024). Real-time railway traffic management under moving-block signalling: A literature review and research agenda. *Transportation Research Part C: Emerging Technologies*, 158, 104438.

Conference contributions

Versluis, N. D., Pellegrini, P., Quaglietta, E., Goverde, R. M. P., & Rodriguez, J. (2025). Impact of track discretisation on conflict detection and resolution under ETCS with onboard train integrity monitoring. *RailDresden 2025, 11th International Conference on Railway Operations Modelling and Analysis (ICROMA), Dresden, Germany*.

Versluis, N. D., Pellegrini, P., Quaglietta, E., Goverde, R. M. P., & Rodriguez, J. (2024). The impact of track discretisation in rescheduling models for advanced distance-to-go railway signalling. *ODS 2024, International Conference on Optimization and Decision Science, Badesi - Sardinia, Italy*.

Versluis, N. D., Pellegrini, P., Quaglietta, E., Goverde, R. M. P., & Rodriguez, J. (2024). Track discretisation in railway traffic rescheduling models for next-generation distance-to-go signalling. *EURO 2024, 33rd European Conference on Operational Research, Copenhagen, Denmark*.

Versluis, N. D., Pellegrini, P., Quaglietta, E., Goverde, R. M. P., & Rodriguez, J. (2023). A MILP-based railway conflict detection and resolution model approximating moving-block signalling. *ODS 2023, International Conference on Optimization and Decision Science, Ischia, Italy*.

Versluis, N. D., Pellegrini, P., Quaglietta, E., Goverde, R. M. P., & Rodriguez, J. (2023). An approximate conflict detection and resolution model for moving-block signalling by enhancing RECIFE-MILP. *RailBelgrade 2023, 10th International Conference on Railway Operations Modelling and Analysis (ICROMA), Belgrade, Serbia*.

PERFORMINGRAIL project deliverables

Quaglietta, E., Versluis, N. D., Beugin, J., Ghazel, M., Kirkwood, D., & Garcia-Fernandez, M. (2023). *Deliverable D1.2: Best Practices, Recommendations and Standardisation to Definition of the Railway Minimum Operations Performance Standards.*

Quaglietta, E., Versluis, N. D., Goverde, R. M. P., Pellegrini, P., Manzini, A., & Garcia-Fernandez, M. (2023). *Deliverable D4.2: Guidelines for a Safe and Optimised Moving-Block Traffic Management System Architecture.*

Quaglietta, E., Versluis, N. D., Goverde, R. M. P., Pellegrini, P., Nardone, N, Vittorini, V., Manzini, A., Garcia-Fernandez, M., & Sanwal, M. U. (2022). *Deliverable D4.1: Real-Time Traffic Rescheduling Algorithms and Perturbation Management and Hazard Prevention in Moving-Block Operations.*

Seceleanu, C., Flammini, F., Marrone, S., Mogavero, F., Nardone, R., Starace, L. L. L., Vittorini, V., Beugin, J., Ghazel, M., Saddem, R., Goverde, R. M. P., & Versluis, N. D. (2021). *Deliverable D2.1: Modelling Guidelines and Moving Block Use Cases Characterization.*

TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series. For a complete overview of more than 400 titles see the TRAIL website: www.rsTRAIL.nl.

The TRAIL Thesis Series is a series of the Netherlands TRAIL Research School for TRAnsport, Infrastructure and Logistics.

Versluis, N. D., *Optimising Railway Traffic Management under Radio-Based Distance-to-Go Signalling*, T2026/2, January 2026, TRAIL Thesis Series, the Netherlands.

Jiao, Y., *Proactive Collision Risk Quantification in Multi-directional Traffic Interactions*, T2026/1, January 2026, TRAIL Thesis Series, the Netherlands.

Asadi, M., *Accessibility and Road Safety: Integration Road Safety in Accessibility Evaluation*, T2025/20, November 2025, TRAIL Thesis Series, the Netherlands.

Akse, R., *Understanding and untangling the uncertainty knot: How to catalyse decision-making in mobility innovations*, T2025/19, November 2025, TRAIL Thesis Series, the Netherlands.

Führer, K., *Participatory Decision-making under Deep Uncertainty: Modelling Mobility Transitions*, T2025/18, November 2025, TRAIL Thesis Series, the Netherlands.

Picco, A., *Monitoring and Feedback in Driving*, T2025/17, October 2025, TRAIL Thesis Series, the Netherlands.

Cebecci, M. S., *Behaviour of Prosumers in Last-mile Logistics: The Case of Crowdshipping*, T2025/16, September 2025, TRAIL Thesis Series, the Netherlands.

Kuijpers, A., *Enabling Inter-Organizational Collaboration Through Platforms: The Role of Trust*, T2025/15, September 2025, TRAIL Thesis Series, the Netherlands.

Song, R., *Human-MASS Interaction in Decision-Making for Safety and Efficiency in Mixed Waterborne Transport Systems*, T2025/14, June 2025, TRAIL Thesis Series, the Netherlands.

Destyanto, A. R., *A Method for Evaluating Port Resilience in an Archipelago*, T2025/13, June 2025, TRAIL Thesis Series, the Netherlands.

Karademir, C., *Synchronized Two-echelon Routing Problems: Exact and Approximate Methods for Multimodal City Logistics*, T2025/12, May 2025, TRAIL Thesis Series, the Netherlands.

Vial, A., *Eyes in Motion: A New Traffic Sensing Paradigm for Pedestrians and Cyclists*, T2025/11, May 2025, TRAIL Thesis Series, the Netherlands.

Chen, Q., *Towards Mechanical Intelligence in Soft Robotics: Model-based Design of Mechanically Intelligent Structures*, T2025/10, April 2025, TRAIL Thesis Series, the Netherlands.

Eftekhar, Z., *Exploring the Spatial and Temporal Patterns in Travel Demand: A Data-Driven Approach*, T2025/9, June 2025, TRAIL Thesis Series, the Netherlands.

Reddy, N., *Human Driving Behavior when Interacting with Automated Vehicles and the Implications on Traffic Efficiency*, T2025/8, May 2025, TRAIL Thesis Series, the Netherlands.

Durand, A., *Lost in Digitalisation? Navigating public transport in the digital era*, T2025/7, May 2025, TRAIL Thesis Series, the Netherlands.

Dong, Y., *Safe, Efficient, and Socially Compliant Automated Driving in Mixed Traffic: Sensing, Anomaly Detection, Planning and Control*, T2025/6, May 2025, TRAIL Thesis Series, the Netherlands.

Droffelaar, I. S. van, *Simulation-optimization for Fugitive Interception*, T2025/5, May 2025, TRAIL Thesis Series, the Netherlands.

Fan, Q., *Fleet Management Optimisation for Ride-hailing Services: From Mixed Traffic to Fully Automated Environments*, T2025/4, April 2025, TRAIL Thesis Series, the Netherlands.

Hagen, L. van der, *Machine Learning for Time Slot Management in Grocery Delivery*, T2025/3, March 2025, TRAIL Thesis Series, the Netherlands.

Schilt, I. M. van, *Reconstructing Illicit Supply Chains with Sparse Data: A Simulation Approach*, T2025/2, January 2025, TRAIL Thesis Series, the Netherlands.

Ruijter, A. J. F. de, *Two-Sided Dynamics in Ridesourcing Markets*, T2025/1, January 2025, TRAIL Thesis Series, the Netherlands.

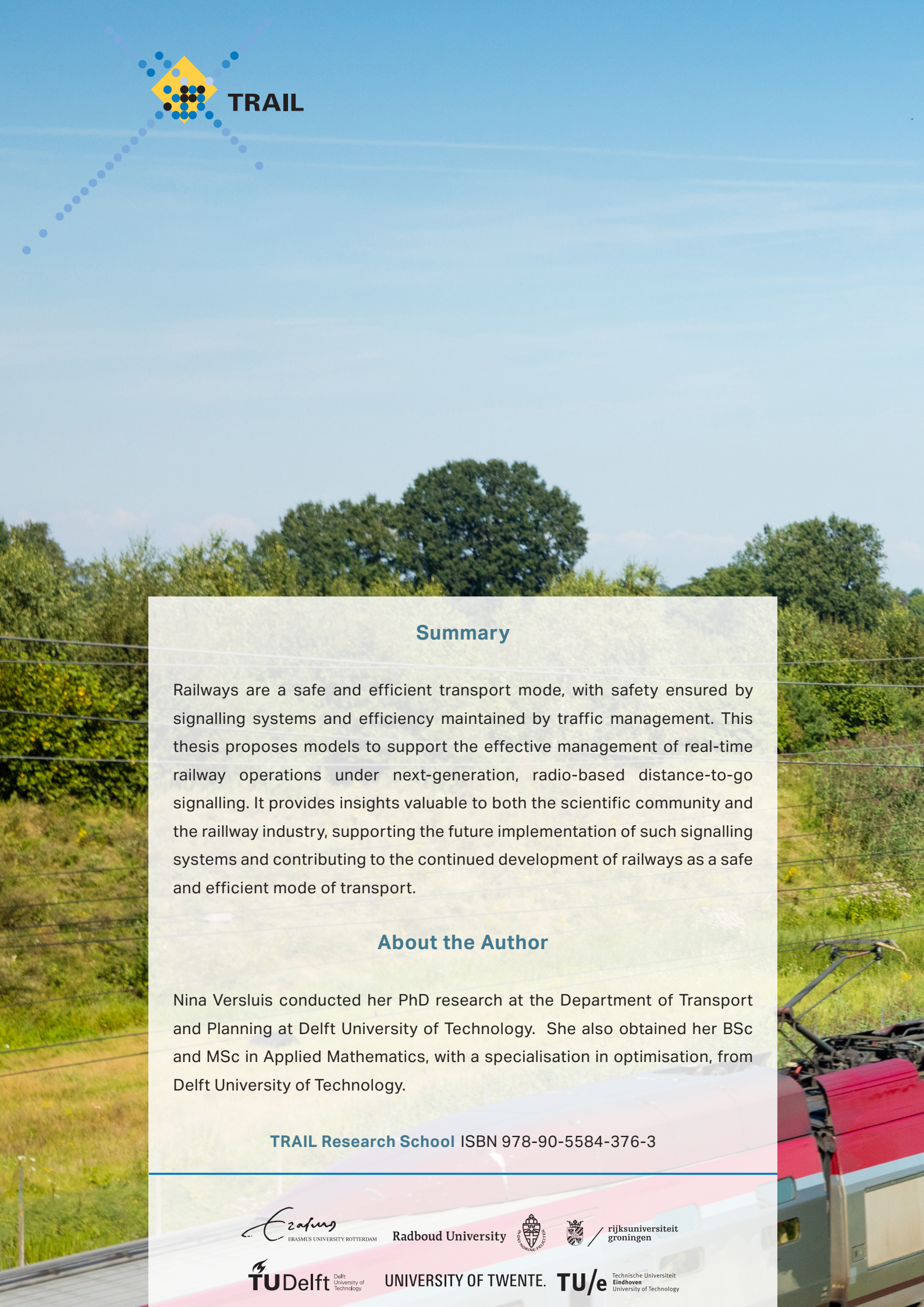
Fang, P., *Development of an Effective Modelling Method for the Local Mechanical Analysis of Submarine Power Cables*, T2024/17, December 2024, TRAIL Thesis Series, the Netherlands.

Zattoni Scroccaro, P., *Inverse Optimization Theory and Applications to Routing Problems*, T2024/16, October 2024, TRAIL Thesis Series, the Netherlands.

Kapousizis, G., *Smart Connected Bicycles: User Acceptance and Experience, Willingness to Pay and Road Safety Implications*, T2024/15, November 2024, TRAIL Thesis Series, the Netherlands.

Lyu, X., *Collaboration for Resilient and Decarbonized Maritime and Port Operations*, T2024/14, November 2024, TRAIL Thesis Series, the Netherlands.

Nicolet, A., *Choice-Driven Methods for Decision-Making in Intermodal Transport: Behavioral heterogeneity and supply-demand interactions*, T2024/13, November 2024, TRAIL Thesis Series, the Netherlands.



TRAIL

Summary

Railways are a safe and efficient transport mode, with safety ensured by signalling systems and efficiency maintained by traffic management. This thesis proposes models to support the effective management of real-time railway operations under next-generation, radio-based distance-to-go signalling. It provides insights valuable to both the scientific community and the raillway industry, supporting the future implementation of such signalling systems and contributing to the continued development of railways as a safe and efficient mode of transport.

About the Author

Nina Versluis conducted her PhD research at the Department of Transport and Planning at Delft University of Technology. She also obtained her BSc and MSc in Applied Mathematics, with a specialisation in optimisation, from Delft University of Technology.

TRAIL Research School ISBN 978-90-5584-376-3

ERASMUS UNIVERSITY ROTTERDAM · Radboud University · rijksuniversiteit groningen · TU Delft Delft University of Technology · UNIVERSITY OF TWENTE. · TU/e Technische Universiteit Eindhoven University of Technology