Deliverable D4.1

Real-Time Traffic Rescheduling Algorithms and Perturbation Management and Hazard Prevention in Moving-Block Operations

Quaglietta, Egidio; Versluis, Nina; Goverde, Rob; Pellegrini, Paola; Nardone, Robert; Vittorini, Valeria; Manzini, Achila; Garcia, Miquel; Sanwal, Muhammad Usman

Publication date
2022

Document Version
Final published version

Citation (APA)
Quaglietta, E., Versluis, N., Goverde, R., Pellegrini, P., Nardone, R., Vittorini, V., Manzini, A., Garcia, M., & Sanwal, M. U. (2022). *Deliverable D4.1: Real-Time Traffic Rescheduling Algorithms and Perturbation Management and Hazard Prevention in Moving-Block Operations*. European Commission. https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5f3375fb6&appId=PPGMS

Deliverable D4.1

Real-Time Traffic Rescheduling Algorithms for Perturbation Management and Hazard Prevention in Moving-Block Operations

Project acronym:	PERFORMINGRAIL
Starting date:	01/12/2020
Duration (in months):	31
Call (part) identifier:	S2R-OC-IP2-01-2020
Grant agreement no:	101015416
Due date of deliverable:	Month 22
Actual submission date:	31-10-2022
Responsible/Author:	TUD
Dissemination level:	PU
Status:	Draft/~~Issued~~

Reviewed: (yes)

Document history		
Revision	*Date*	*Description*
0.1	01/10/2022	First issue
0.2	17/10/2022	Second issue
0.3	24/10/2022	Third issue
0.4	26/10/2022	Fourth issue with review and revision of contributions from all partners
0.5	28/10/2022	Fifth issue with document revision based on final peer-review process

Report contributors		
Name	Beneficiary Short Name	Details of contribution
Egidio Quaglietta	TUD	• Coordination of partner contribution and set up structure of deliverable. • Contribution to traffic management and early-warning hazard modelling and review of section content. • Writing of Executive summary, introduction, background, conclusions and finalisation of the document
Nina Versluis	TUD	• Responsible for writing TUD sections of the deliverable • Developing and describing mathematical models for traffic management and early-warning hazards • Contribution to set up deliverable structure, and draft version editing.
Rob Goverde	TUD	• Contribution to traffic management mathematical modelling • Review of traffic management modelling sections
Paola Pellegrini	UNI EIFFEL	• Mathematical model
Roberto Nardone Valeria Vittorini	CINI	• Customization of the SAN model for evaluating the impact of Signalling System Features on Safe Moving Block • Sensitivity Analysis • Writing of Section 5.3

Achila Mazini	UoB	• Writing of Section 5.1 for the definition of requirements for safe MB rail operations
Miquel Garcia	ROK	• GNSS Train Localisation and Integrity Requirements
Muhammad Usman Sanwal	MDH	• Peer-review of the final draft of the document

Funding

This project has received funding from the Shift2Rail Joint Undertaking (JU) under grant agreement No 101015416. The JU receives support from the European Union's Horizon 2020 research and innovation programme and the Shift2Rail JU members other than the Union.

Table of Contents

Executive Summary

This deliverable has the objective to define a mathematical model for an optimised real-time management of railway traffic under Moving Block (MB). The formulated real-time traffic management model contains: i) a core module for the detection and the sub-optimal resolution of track occupation conflicts under MB and ii) a non-vital module for providing early-warning predictions of potentially hazardous MB traffic situations. The proposed real-time traffic management model includes a mathematical translation of requirements and constraints identified for both MB signalling within WP2 (namely deliverables D2.1 and D2.2) and the GNSS localisation and train integrity devices within WP3 (i.e. deliverables D3.1 – D3.3). An extensive literature review on real-time traffic management models and algorithms shows that so far research efforts have mainly focused on fixed-block and distance-to-go railway operations. Significant gaps still exist in the modelling of MB train operations, despite an increasing number of research works on MB signalling technology is observed since year 2003. A modelling gap analysis is here performed which indicates the need of enhancing existing real-time traffic management algorithms to better align them to the MB concept in terms of infrastructure representation and speed-headway functional dependency. To this end, the RECIFE-MILP real-time traffic management algorithm is enhanced. On one hand a finer infrastructure discretisation is implemented to offer a more suitable track representation under moving block which no longer uses fixed block sections. On the other hand, two different speed levels (namely maximum speed and scheduled speed) are introduced enabling a speed-dependent headway computation in either nominal or delayed traffic scenarios, thereby overcoming the limitation of speed-independent headways, typical of fixed-block traffic rescheduling models.

A non-vital early-warning prediction model of hazardous MB traffic conditions is also proposed which includes a short- and a medium-term hazard identification method. In the short-term, potentially hazardous MB traffic condition are identified as violations of safety-critical threshold values of design variables relating to MB train operations (e.g. driving reaction times), the GNSS system (e.g. GNSS error or latency) and/or the GSM-R layer (e.g. MA communication delay). Safety-critical thresholds of the different design variables are identified by means of an extensive sensitivity analysis which uses a Stochastic Activity Network built for MB within WP2. In the medium-term warnings of potentially hazardous MB conditions are instead triggered whenever RECIFE-MILP detects track occupation conflicts in geographical areas with limited GNSS and/ or GSM-R signal availability, such as deep valleys or tunnels. The defined models contribute to the definition of an optimised automated Traffic Management System for Moving Block which can also support traffic dispatchers in preventively avoiding the occurrence of potentially dangerous MB traffic conditions.

Abbreviations and Acronyms

Abbreviation / Acronyms	Description
AG	Alternative Graph
ATO	Automatic Train Operations
ATP	Automatic Train Protection
CDR	Conflict Detection and Resolution
CTCS	Chinese Train Control System
ERTMS	European Rail Traffic Management System
ETCS	European Train Control System
FB	Fixed Block
GNSS	Global Navigation Satellite System
GSM-R	Global System for Mobile Communications - Railway
MB	Moving Block
MILP	Mixed Integer Linear Programming
RBC	Radio Block Centre
SAN	Stochastic Activity Network
TIM	Train Integrity Monitoring
TPR	Train Position Report
WP	Work Package

1 Introduction

The railway industry is in need of increasing capacity and competitiveness of existing railway services to accommodate the forecasted rail demand growth while aligning to European strategic goals on transport sustainability (Eureopean Comission, 2011). As extending railway tracks is a very costly solution, hardly possible in densely built-up areas, the railway industry is mainly aiming at enhancing signalling, traffic management and control technologies. Several are indeed Research and Development (R&D) programmes of National (e.g. the UK Digital Railway, 2015, the German Digitale Schiene, 2017) and International (e.g. the EC FP7 ONTIME project, 2012, the EC Shift2Rail programme, 2020) relevance, addressed to increasing railway automation levels and enable the migration to digital high-performance train operations. The transition towards advanced signalling technologies such as ERTMS/ETCS Level 2 and Level 3 (Theeg and Vlasenko, 2009) is considered as one of the main capacity enablers together with the deployment of Automatic Train Operation (Wang et al., 2022) as well as optimised traffic planning and management (Quaglietta et al., 2016). In such context, real-time traffic management plays an essential role in providing capacity-effective and seamless train service in either nominal or degraded operational conditions, especially for highly-used and dense railway networks. Extensive literature is available (e.g. Pellegrini et al., 2014, D'Ariano et al. 2008, Törnquist & Persson, 2007) on real-time traffic management tools for aiding dispatchers' decisions, however the proposed approaches mainly refer to conventional fixed-block multi-aspect signalling. Only a few works can be found on rail traffic management under advanced radio-based signalling systems (e.g. Mera et al., 2016, Janssens, 2022), which still require further investigation and validation on real case studies. The formulation of real-time traffic management models aligned to technical and functional specifications of next-generation signalling systems becomes therefore a necessary step to enable the migration to a digital capacity-effective railway. Traffic models specifically defined for next-generation signalling and especially for Moving Block (MB) will indeed support the railway community in outlining the architecture of the Advanced Traffic Management System, set as one of the main innovation goals by the Shift2Rail Multi-Annual Action Plan (Shift2Rail JU, 2020).

The work performed in this deliverable contributes to bridge the gaps in current literature by proposing a real-time rail traffic management model which complies with the latest requirements for MB signalling as well as GNSS-based train location and integrity technologies. The proposed model enhances the state-of-the-art RECIFE-MILP real-time traffic management tool (Pellegrini et al., 2014) by including moving block-specific constraints regarding train operation times, signalling as well as train location and integrity components. The proposed model also includes an early-warning moving block hazard prediction which can support dispatchers in anticipating suitable strategies to avoid the raising of safety-critical MB traffic conditions in the short and medium-term.

Section 3 provides an extensive literature review on existing real-time traffic management models together with a gap analysis with respect to advanced signalling systems and moving block in particular. Section 4 describes the mathematical formulation of the enhanced RECIFE-MILP specifically adjusted to include MB constraints regarding train operations, signalling as well as GNSS train location and integrity devices. Section 5 reports a method for early-warning prediction of hazardous moving block traffic conditions referring to both the short- and the medium-term. Conclusions and future research directions are discussed in Section 6.

1.1 Objectives and Scope

The document reports outcomes from the mathematical modelling activities performed in Task 4.1 and Task 4.2 within the scope of the EC Shift2Rail project PERFORMINGRAIL – Work Package 4 "Integrated Moving Block architecture for safe and optimised traffic operations". Specifically, activities related to Task 4.1 "*Modelling and integrating Future moving-block architectures with safe and reliable signalling and train localisation specifications*" carry out an extensive literature review of real-time traffic management algorithms and a gap analysis with respect to Moving Block operations. Furthermore, MB signalling requirements from WP1 and WP2 as well as features of GNSS-based train location and integrity from WP3 are translated into mathematical constraints to enhance the real-time traffic management tool RECIFE-MILP. Activities in Task 4.2 "Algorithms for optimised traffic management and early-warning of hazardous events under moving-block" use the MB constraints defined in Task 4.1 to provide an enhanced mathematical formulation of the RECIFE-MILP real-time traffic rescheduling tool which aligns to the specifications of MB train operations. In addition, a method for early-warning hazard prediction is defined to identify potentially dangerous MB traffic conditions which might arise in the short- or the medium-term.
Hence the main objectives of this deliverable are:

- Analysing gaps in existing real-time traffic management tools / approaches with respect to a reliable and effective modelling of Moving Block train operations.
- Define a set of mathematical constraints describing principles, signalling and GNSS-based train location and integrity requirements for safe MB train operations.
- Delineate methods for non-vital early-warning prediction of hazardous MB traffic conditions.

The objectives of this deliverable are therefore linked to Technology Demonstrator TD2.9 "TMS Evolution" in IP2 "Advanced Traffic Management and Control Systems" and Work Area 4 "Smart Mobility" of the Shift2Rail MAAP (2020).

1.2 Related Documents

This document relies on inputs provided by deliverables of other PERFORMINGRAIL WPs as well as other Shift2Rail projects. In detail the set of deliverables relative to other PERFORMINGRAIL WPs are:

- Deliverable D1.1: Baseline system specification and definition for Moving Block Systems
- Deliverable D2.2: Moving Block Specification Development
- Deliverable D3.1: Design document of the Location algorithms
- Deliverable D3.3: Multi-frequency/constellation GNSS receiver

Further inputs are provided from deliverables of the S2R project X2Rail-3, specifically:

- D4.2 Moving Block Specification
- D4.3 Future Moving Block Architecture

Outputs from this document will be instead feed the content of other tasks and WPs in the PERFORMINGRAIL project, namely:

- D4.2: Guidelines for a safe and optimised moving-block traffic management system architecture
- D5.2: Assessment report

2 Background

The present document constitutes D4.1 “Real-time traffic rescheduling algorithms for perturbation management and hazard prevention in moving-block operations” which is the first deliverable of WP4 “Integrated Moving Block architecture for safe and optimised traffic operations” of the Shift2Rail project PERFORMINGRAIL.
Referring to the Shift2Rail MAAP (2020), the work described in this document links to Technology Demonstrator TD2.9 “TMS Evolution”, tasks 2.9.3 “Framework for Traffic Management Business Service” and 2.9.6 ”Functionalities and Interfaces for Dynamic Demand and Information Management” in IP2 “Advanced Traffic Management and Control Systems”. Also, it refers to Work Area WA 4 “Smart Mobility” subtask 4.2 “Integrated mobility management” of the Shift2Rail MAAP (2020).

3 Literature Review on Rescheduling Algorithms for Fixed and Moving-Block Railway Operations

Existing literature on rail traffic rescheduling models is reviewed as a springboard for the identification of gaps in the literature regarding rail traffic rescheduling under moving-block signalling. Rail traffic rescheduling under moving-block signalling is staying behind the trend of increasing numbers of moving-block publications, as Figure 1 illustrates.

Consequently, the review does not only consider the moving-block signalling system (Section 3.3), but also the conventional fixed-block multi-aspect signalling system (Section 3.1), as well as the intermediate fixed-block distance-to-go signalling system (Section 3.2).

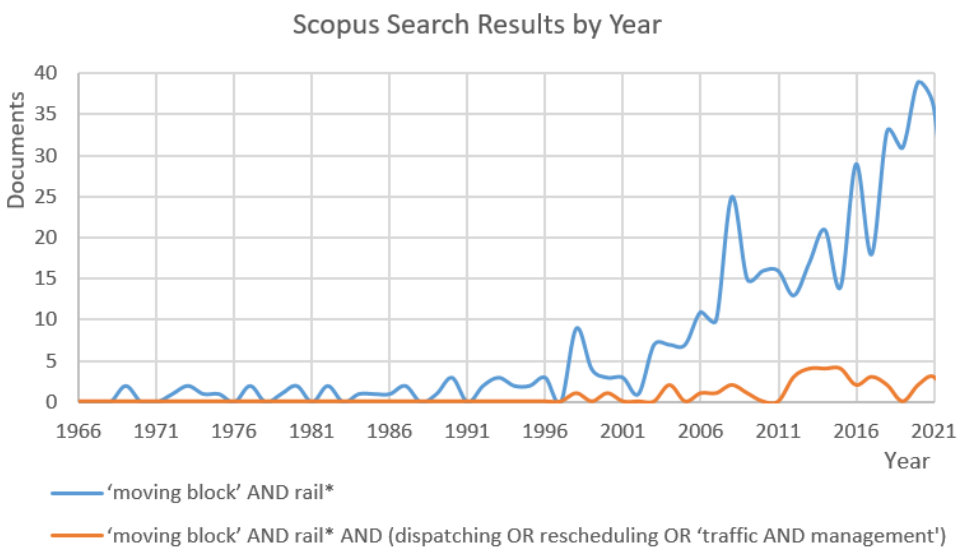

Figure 1. Trend of scientific publications on MB signalling (blue line) and traffic management models for MB (red line).

In both the modelling of rail operations and the moving-block signalling system, (minimum) train headways are a crucial feature. Important aspects of the modelling of train headways are the overall modelling approach, the representation of the infrastructure and the consideration of speed. Other relevant (rescheduling) modelling aspects are the consideration of rerouting, the solution method and the objective function.

In Table 1, representative literature is classified in terms of the presented modelling aspects.

Table 1. Overview on real-time traffic rescheduling models proposed for fixed-block, distance-to-go and MB signalling.

Publication	MAp	IR	SpM	HM	RR	SoM	OF (min .. delay)
Fixed-block multi-aspect signalling							
Corman et al. (2009)	AG	BS	Green wave	Blocking	No	Heuristic	Max secondary
D'Ariano et al. (2007)	AG	BS	Iterative	Occupation	No	B&B	Max secondary
D'Ariano et al. (2007)	AG	BS	Fixed	Occupation	No	B&B	Max secondary
D'Ariano et al. (2008)	AG	BS	Fixed	Occupation	Yes	B&B	Max secondary
Luan et al. (2018)	D-MILP	BS	Levels	Blocking time	No	Two-level	Total mean absolute
Lusby et al. (2013)	TI-MILP	TDS	Variable	Occupation	Yes	B&P	Weighted total
Mazzarello and 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 and 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 and 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							
F. Liu et al. (2021)	MILP	BS	Iterative	Default	No	Bi-level	Total final secondary
P. Xu et al. (2017)	AG-MILP	BS	Levels	Default	No	Two-step	Total final secondary
P. 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

MAp: modelling approach, IR: infrastructure representation, SpM: speed modelling, HM: headway modelling, RR: rerouting SoM: solution method, OF: objective function, 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, (M)BS: (moving) block section, TDS: track detection section, B&B: branch-and-bound, B&P: branch-and-price, Max: maximum.

3.1 Rail Traffic Rescheduling under Fixed-Block Multi-Aspect Signalling

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

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 (Rodriquez, 2007). Though some promising results of these approaches, they have not been picked up in the research field yet.

3.1.1 Alternative Graph Models

In alternative-graph based models, the (re)scheduling of rail operations is considered as a no-wait job shop scheduling problem. For this well-known scheduling problem, an alternative graph formulation was developed by (Mascis, 2002). The approach was first applied to rail operations in D'Ariano et al. (2007a) and Mazzarello & Ottaviani (2007).

An alternative graph consists of nodes, fixed arcs and pairs of alternative arcs. A rail network can be described as an alternative graph as follows, also illustrated in Figure 2. The nodes correspond to the entry (time) of a train on a block section, while the (weight of the) fixed arcs represent the train running time in the blocks and the dwell times at station platform tracks along a route. In D'Ariano et al. (2007a), the running time refers to the time it takes to traverse the block section at planned speed, while 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.

The alternative arcs represent conflicting pairs of operations that require the determination of a train passing order and a minimum train headway. This is done through selecting one of each pair of alternative arcs. In the models of D'Ariano et al. (2007a) and Mazzarello & Ottaviani (2007), the alternative arcs connect succeeding block entries of pairs of running trains (Figure 2). Hence, the running time over the block is directly included in the headway time. The rest of the headway is determined by the weight of the alternative arcs. D'Ariano (2007a) include the time between the block exit of the head and of the tail of the train, i.e., the clearing time, complemented 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 variable term depends on the train length and speed, so it includes the clearing time. Note that the time it takes 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 later version of the AG model, the approach time is included in the headway by letting alternative arcs connect two nodes corresponding to block entries that are the number of signal aspects apart (Corman et al., 2009).

Alternative graph based rescheduling models require the timetable and the current delays as input, next to the alternative graph representation of the infrastructure.

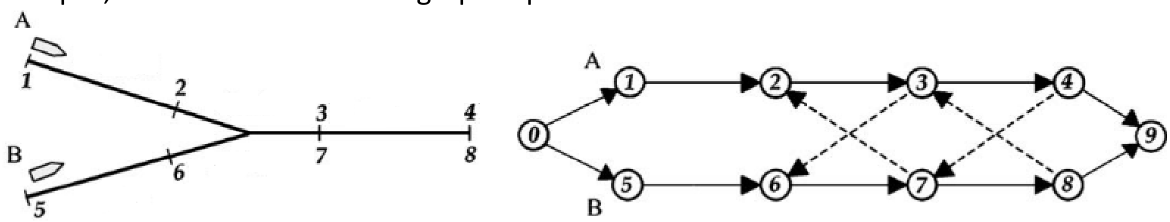

Figure 2. A simple network (left) and the associated alternative graph (right). Adapted from (Mazzarello & Ottaviani, 2007).

The input delays are included through a later earliest arrival time from the source node. The rescheduling model is solved by a heuristic algorithm, such as the truncated branch-and-bound algorithm proposed in D'Ariano et al. (2007a). The algorithm searches for a selection of arcs representing a conflict-free schedule with the shortest longest path. Herewith, the maximum secondary delay is minimised (D'Ariano et al., 2007a). That is the delay trains face due to interaction with other (delayed) trains as a result of the initial delays. This objective is the direct translation of the objective of the job shop scheduling problem, i.e., minimising the makespan.

The alternative graph approach is particularly suitable for retiming and reordering, though efforts are made to include rerouting. For example, D'Ariano et al. (2008) extend D'Ariano et al. (2007a) to include rerouting through the use of a meta-heuristic. Mazzarello & Ottaviani (2007), also address rerouting. After a feasible solution with fixed routes is found, a heuristic selection of alternative routes is considered to find the best option per train consistent with the original solution.

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 how speed is incorporated in Mazzarello & Ottaviani (2007). In this approach, an overarching traffic control algorithm iterates between the fixed-speed rescheduling 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. Hence, trains run according to scheduled speed profiles, only allowing retiming by dwell time extension at planned stops. The green wave approach is also applied in Caimi et al. (2012), in the context of a different microscopic model using model predictive control.

3.1.2 Disjunctive Mixed Integer Linear Programming Models

Alternative graph models can be extended to consider objective functions beyond maximum (secondary) delay by translating it into a MILP. For example, Samà et al. (2015) extend D'Ariano et al. (2017a) to a multi-criteria optimisation model. This approach, AG-based MILP, is in fact an example of a disjunctive MILP. In a disjunctive MILP rescheduling model, the train operations are described by decision variables and 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 a specific track part when. 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 train enters first, and routing, i.e., which route does a train take (Luan etl a., 2018; Pellegrini et al. (2014); Törnquist & Persson, 2007). The exact decision variables to be defined depend on how the operations are translated into the model. For example, whether train and infrastructure are considered separately, as in Luan et al. (2018) and Pellegrini et al. (2014), or already coupled to be considered as events, as Samà et al. (2015) and Törnquist & Persson (2007) do.

Pellegrini et al. (2014) introduce continuous variables to indicate when a train enters a track detection section and binary variables to decide on the section entering order of a pair of trains and on the route of a train. Similarly, Luan et al. (2018) use continuous variables for train arrival and departure times on block sections and binary ordering variables for trains arriving at a block section. Törnquist & Persson (2007) include continuous time variables representing event times, i.e., the start and end time of activities such as a train traversing a block section. The ordering is addressed using binary decision variables indicating whether two (rescheduled) events are occurring in the scheduled order or not. Another set of binary decision variables relate to the routes by indicating which (parallel) track is used by a specific train.

Disjunctive MILP rescheduling models rely on big-M formulations to obtain linear constraints. The 'big-M method' is a general modelling technique for the linearisation of optimisation problems. 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). Based on Törnquist & Persson (2007), Törnquist (2012) develops an effective solution method relying on a greedy depth-first algorithm. Pellegrini et al. (2015) extend Pellegrini et al. (2014) with the proposal of a heuristic algorithm further investigated in Pellegrini et al. (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. The following equations represent a generic pair of big-M constraints for the rescheduling problem:

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

with *t, t'* trains, *s* a block section and Δ the minimum separation time. Depending on whether train t goes before train t' on section s (if $order_{t,t',s} = 1$), a minimum separation time ($\Delta_{t,t',s}$) is to be respected between the two relevant operations, e.g., the occupation of section s by trains t and t', respectively.

In Pellegrini et al. (2014), operations are separated by determining a 'blocking' order. That is, the order in which a track detection section is blocked by a train according to blocking time theory (Hansen & Pachl, 2014). 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 detection section. The blocking start 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 stop unplanned. Only at planned stops the train times are strictly aligned with the scheduled departure times.

The consideration of track detection sections next to block sections is the highest level of detail to represent the infrastructure. If implemented properly, it leads to accurate results in terms of blocking times as it allows sectional release (Pellegrini et al., (2014); Hansen & Pachl, 2014). The modelling of sectional release at all track detection sections would lead to incorrect block releases where this does not coincide with clearing points. When track detection sections are used in the same way as block sections, i.e., sectional locking, then also incorrect results are obtained.

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. First, a mixed integer non-linear programming model is proposed for the integrated problem of rescheduling and speed control. Then, the non-linear model is converter into a MILP by approximating the non-linear terms with piece-wise affine functions. In this model, speed is included through predetermined speed profile options per train-block section pair.

In Törnquist & Persson (2007), both station areas and the open line between them are considered in terms of sets of parallel block sections. As 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 activities associated with different trains and one block section, however, a fixed time is used. In the case study, the minimum headway time is fixed at three minutes, independent of the specific (parallel) block section or train.

The reason to translate the 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. Pellegrini et al. (2014) restrain to the consideration of two common objectives for the rescheduling problem to minimise the maximum or the total secondary delay. Pellegrini et al. (2015) include the more general objective of minimising total weighted final delays, while Törnquist & Persson (2007), consider the total and the weighted final delay separately, and Luan et al. (2018) the total mean absolute delay.

3.1.3 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. In Lusby et al. (2013) and Reynolds et al. (2020), the rail traffic rescheduling problem is formulated as a time-indexed MILP.

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). For example, Lusby et al. (2013) consider resources that relate to combinations of a time unit, e.g., 15 seconds and a track detection section. Reynolds et al. (2020), construct a time-space graph in which the nodes correspond to a time unit and block section pair and the arcs to possible train routes, as illustrated in Figure 3. In this graph, a train route is indicated by a source-sink path. In general, the time-indexed approach is well-suited for the consideration of rerouting (Reynolds et al., 2020, Lusby et al., 2013).

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). For each resource, a set packing constraint is included, such as the following:

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

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* ($x^r_t = 1$) is restricted to at most one. With this, the occurrence of track conflicts is excluded.

Time-indexed modelling requires to express the running and headway times in fixed 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 based on blocking time theory. In the case study, 30 seconds is assumed as time unit.

The formulation also allows for various objective functions. For example, the total weighted delays (Lusby et al., 2013) or a custom utility function (Reynolds et al., 2020). In Reynolds et al. (2020), the total utility is determined by summing up the 'utility' weights of the selected arcs, which correspond to the assigned train routes.

The main drawback of the time-indexed approach is the model size (Van den Akker et al., 2000; Pellegrini et al., 2019). 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 real-time rescheduling. A validated combination is the use of a heuristic method and a model composition. For example, Lusby et al. (2013) and Reynolds et al. (2020) propose the application of branch-and-price algorithms and column generation after model decomposition, a known technique for time-indexed scheduling models (Van de Akker et al., 2000).

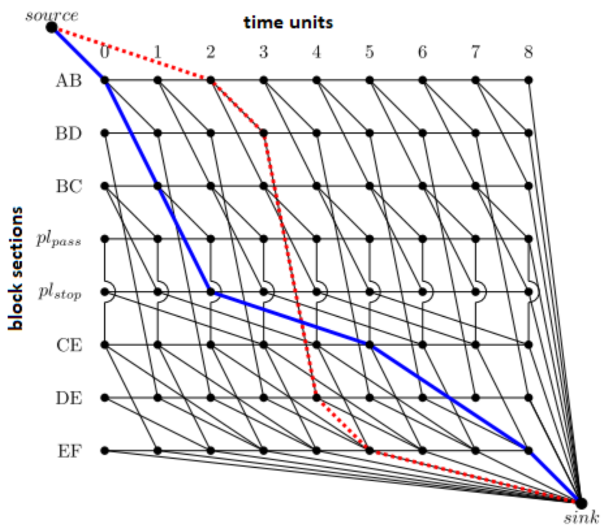

Figure 3. Example time-space graph with two train routes, i.e., source-sink paths, shown. Adapted from Reynolds et al. (2020).

The model described in Reynolds et al. (2020) and illustrated in Figure 3 is extended to a variable-speed model in Reynolds & Maher (2022). The estimation method based on historical data is extended to consider two running times per train block section pair, corresponding to the ‘near minimum' running time and the ‘accelerating/decelerating or lower speed coasting' running time. From this, sets of speed profile types are constructed on route level. A computational performance comparable to the fixed-speed model's is claimed for complete station areas (Reynolds et al., 2020). Similarly, Lusby et al. (2013) approximate speed profiles by considering the options of continuing at constant speed, or by accelerating or decelerating.

3.2 Rail Traffic Rescheduling under Fixed-Block Distance-To-Go Signalling

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 planned 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.

Gonzalez et al. (2010) and Mera et al. (2016), nicely illustrate the role of speed on microscopic level on train headways and line capacity. However, rescheduling in European distance-to-go signalling systems is not considered in literature. Not for the metro lines, but also not for the mainline equivalent, i.e., ERTMS/ETCS Level 2. Literature on the rescheduling under fixed-block distance-to-go signalling is available on the comparable Chinese signalling system, i.e., level 3 of the Chinese Train Control System (CTCS-3), also called the quasi-moving block signalling system in Chinese literature. Note that the CTCS-3 system is designed for the Chinese high-speed (up to 300 km/h) railway lines.

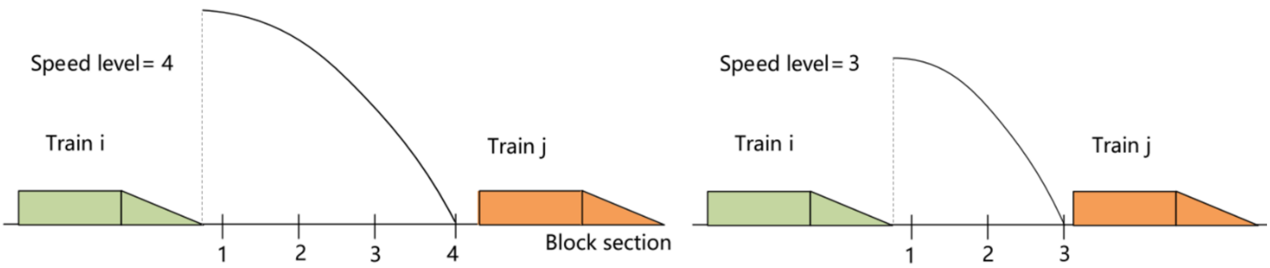

Figure 4. Example of relation between speed level and train headway in terms of block sections. Adapted from Liu et al. (2021).

In accordance with the distance-to-go system characteristics, rescheduling models for CTCS-3 consider speed-dependent running (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. This relation is illustrated in Figure 4, which shows that the speed (level) is one-on-one translated to the number of free block sections needed as minimum train 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 dispatching 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 considers 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 rescheduling model is requires. Also in 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 planned situation, so assuming no effect by the considered disturbances.

The presented rescheduling models rely on a MILP formulation in Liu et al. (2021) and 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 (Xu et al., 2017). 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) as well as in Liu et al. (2021).

3.3 Rail Traffic Rescheduling under Moving-Block Signalling

The rescheduling of rail traffic under moving-block signalling is barely addressed in literature. Janssens (2022) presents a model concept based on the fixed-block rescheduling model described in D'Ariano et al. (2007a). In the development of a moving-block rescheduling 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 interlocking 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 5, 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 in line with the 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 weight correspond to the train clearing time, i.e., the time it takes the train to traverse its length. 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.

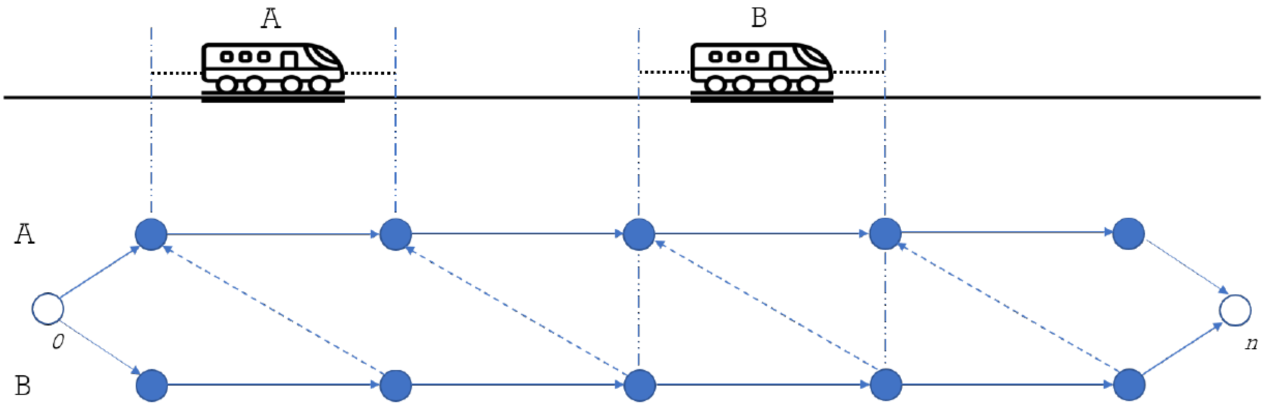

Figure 5. Alternative graph for two trains on a single track under moving-block signalling with virtual nodes in blue. Adapted from (Janssens, 2022).

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 model components. Moreover, the model adopts the fixed-speed assumption from the base model and does not consider rerouting.

In the modelling of moving-block operations in general, speed modelling is included beyond occupation time extensions in case of unplanned stops (Gao et al, 2020; Liu et al., 2016; Wang et al., 2014; Xu et al., 2014). This is in line with the moving-block feature of speed-dependent headways (ERA-UNISIG-EEIG ERTMS Users Group, 2016). Gao et al. (2020) point out that for moving block, 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 in literature, e.g. Liu et al. (2016) and Xu et al. (2014).

In Liu et al. (2016), the focus is on the speed control of following trains. Actually, the problem addressed is a dual control problem: finding the optimal speed and a safe following distance. The authors propose a dynamic system modelling approach adopted from an optimal speed car-following model. In Xu et al. (2014), the minimum headway between trains is determined dynamically based on an optimised safety envelop incorporating speed variations. The optimisation lies in finding the balance between running and headway times, for example by introducing speed limits to reduce braking distance and hence headways.

Gao et al. (2020) consider the speed-headway relation from a slightly different angle. That it, the minimum headways on a metro network are defined based on the driving strategy, of which speed is a main component.

In Wang et al. (2014), train separation is modelled in the context of train trajectory optimisation. The instantaneous braking distance is considered, supplemented with a safety margin as well as components related to the reaction time and train length. The problem is implemented using the pseudospectral method and by formulating a MILP, both solved with commercial solvers. The two methods show a similar performance with respect to the balance between solution quality and computation time, even though for the MILP necessary linearisation results in a discrete-space model (Wang et al., 2014). (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.

Better represented approaches in moving-block models are dynamic systems, e.g. Gao et al. (2020), Liu et al. (2016) and Xu et al. (2014). Dynamic system approaches are characterised by the use of differential equations (Gao et al., 2020; Lie et al., 2016). In these optimal control models, the focus is on speed profile optimisation depending on the train dynamics and constrained by the track characteristics. This train trajectory optimisation is done for single 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, 2017).

This is in line with the moving-block characteristic of viewing the infrastructure as continuous-spaced, moving away from the discretisation into block sections. Note that the fixed block sections are not fully dismissed in the moving-block system. They remain necessary to protect switches in station areas.

The dismissal of fixed block sections on the open line leads to a reduction in the blocking times as considered in blocking time theory. In the context of ERTMS/ETCS Level 3, (Büker et al.,

2019) provide enhancements of the blocking time theory developed for the conventional fixed block system. They incorporate the moving-block system characteristics, such as the bi-directional radio communication for MA and the onboard train positioning and integrity, into the blocking time components.

3.4 Gaps in Rail Traffic Rescheduling under Moving-Block Signalling

This section is dedicated to the identification and analysis of research gaps regarding the modelling of rail traffic rescheduling under moving-block signalling. In general, there is very limited literature on rail traffic rescheduling under moving-block signalling. To identify specific gaps, the focus is on main characteristics of the moving-block system, related to the train headways based on absolute braking distances.

Based on the literature review presented, the following research challenges related to gaps in rail traffic rescheduling under moving-block signalling can be identified.

1. **Rail traffic rescheduling models modelling infrastructure in continuous space.** Existing rescheduling models rely on the fixed-block based discretisation of the infrastructure, while moving-block systems view the infrastructure as continuous space.
2. **Rail traffic rescheduling models including speed dynamics.** Speed modelling in existing rescheduling models assume fixed-speed or consider predetermined speed levels, while moving-block systems are characterised by a dynamic speed-headway relation.
3. **Rail traffic rescheduling models capturing the moving-block specific headway.** The (re)formulation of moving-block headway constraints should consider the continuous infrastructure and speed dependency.

In the following, each of the gaps is illustrated and analysed. The gaps mainly originate from the fact that the existing models are developed for fixed-block signalling, either for the conventional fixed-block multi-aspect or for the more advanced fixed-block distance-to-go system. The analysis of the gaps is performed based on the three main modelling approaches considered in the literature review on rail traffic rescheduling: alternative graph (AG), disjunctive MILP, and time-indexed MILP.

3.4.1 Infrastructure Modelling

Existing (fixed-block) rescheduling models typically consider the infrastructure on the microscopic level of block or track sections, independent of the modelling approach. This `natural' infrastructure discretisation is inherent to the fixed-block system, but not to the moving-block system. At least not on the open line, where the infrastructure is considered as continuous in the modelling of moving-block operations. In station areas, fixed block sections remain in place at switch sections. Modelling the open line as continuous space, on its own and in combination with the fixed-block discretisation around switch sections in station areas, requires further investigation.

The AG formulation for fixed-block requires an infrastructure discretisation for the graph construction. The discretisation does not have to depend on block sections. A finer discretisation in terms of train lengths, as proposed in a first moving-block rescheduling model, allows for an approximation of moving-block sections corresponding to train front and rear positions.
In AG models, the graph is constructed once, thus prohibiting a dynamical update of the graph to let the nodes correspond to the real-time position of the train. Due to the inevitable update

window, the infrastructure would still not be fully continuous. A natural discretisation arises from the (moving-block) MA, which is updated at regular time intervals.

Disjunctive MILP models rely on a discretisation of the infrastructure in a similar way as AG-based models. The fixed discretisation points corresponding to section entries are used as reference points in the model. At these reference points, the minimum separation time between trains is set. Therefore, they should be fixed in space. That said, there is no need that they correspond to specific track parts such as block sections.

Time-indexed MILP models are less flexible in terms of discretisation. With the use of set packing constraints, the model does rely on a discretisation resulting in (time-space) resources with a capacity of at most one

3.4.2 Speed Modelling

Though speed modelling is not part of rescheduling, the modelling of speed profiles is crucial in the translation of an (updated) schedule to (automatic) train operation. In fixed-block multi-aspect systems, the role of speed is limited. With the minimum headway distance fully based on the fixed-block discretisation, only the blocking time components related to traversing times are influenced by speed. In signalling systems with distance-to-go ATP, i.e., fixed-block distance-to-go and moving-block signalling, the impact of speed modelling on the rescheduling decisions increases due to the speed-headway relation via the braking distance.

In trajectory optimisation literature, dynamical systems are used to describe this relation. In rescheduling literature, speed modelling is considered with limited speed levels or speed profile options. This fits the considered modelling approaches, although the models can also be coupled with an external (dynamic) speed optimisation model in an iterative way. More research is needed for the efficient incorporation of (more) dynamic speed modelling into the rescheduling model.

3.4.3 Headway Modelling

In the modelling of train headways, both infrastructure representation and speed modelling are important aspects. In rail traffic (rescheduling) models it is key to capture their role in the model formulation, which depends on the modelling approach.

In existing models, headways are included as default values, as occupation times or as blocking times. Default headways follow from input values, which can be predetermined per section, per train and/or per speed level. Depending on how the predetermined headway is expressed, e.g., in terms of sections, meters or seconds, default headways can be suitable in continuous-space models. Default values are not able to capture speed dynamics beyond predetermined speed levels and/or speed profiles options.

Models that do not consider default headways, generally apply blocking time theory for headway calculations. Through the blocking times, the minimum headway relies on the time a part of infrastructure is assigned to a train. Note that the blocking time includes the occupation time, during which a train is physically present on the considered block section, plus the other blocking time component 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.

Existing rescheduling models rely on blocking time theory as developed for the fixed-block multi-aspect signalling system. Directly related to the traversing time of the number of block sections corresponding to the number of signalling aspects, a reconsideration of blocking time theory is required for moving-block signalling. A first step is done in moving-block literature but the aspects of continuous-space infrastructure, dynamic speed dependencies as well as the influence of the overall system architecture should be further developed.

In alternative graph models, the alternative arcs are used to model train headways. The construction principle of the alternative arcs determines a split in how the different blocking time components are included. Namely, a part is included in the weights of the fixed arcs between the infrastructure points represented by the nodes incident to the alternative arcs, and a part explicitly within the alternative arc weight. For the modelling of moving-block signalling, each infrastructure point represented by the nodes incident to an alternative arc should be a minimum headway distance apart. So alternative graph models do not only need a predetermined discretisation of the infrastructure, also the possible (minimum) headways are fixed before rescheduling. The graph is constructed before rescheduling and speed decisions are taken, by which the minimum headways are already fixed and may no longer be consistent with a changed train running time. This does not mean that speed cannot be incorporated at all. By including multiple pairs of alternative arcs between trains with shared routes, speed choices can be included, depending on the space discretisation. This approach can only work with a discrete set of speed (or headway) possibilities, due to the need for a pair of alternative arcs for different values.

In disjunctive MILP models, minimum headways are modelled indirectly by big-M constraints. These disjunctive constraints illustrate the order decisions and therewith the passing times at decision points. Order (and routing) decisions are taken at junctions or switch sections in station areas, so a discretisation of the infrastructure with respect to decision points is also valid for moving-block signalling. However, these rescheduling decisions should be taken in consideration of the modelled speed, while both the rescheduling and speed choices influence the headways. As in the alternative graph models, speed modelling does not go beyond speed levels in existing disjunctive models, considering a discrete speed set and a space discretisation for the construction of binary speed choice variables. However, in theory, continuous speed choice variables can be included in disjunctive MILP models.

The application of a disjunctive MILP model to moving-block scheduling shows potential for moving-block. In this model, the big-M constraints are indeed only formulated for the decision points, so not for the open line. The open line is considered as one track part on which a range of speed profiles are possible by providing minimum and maximum running times between stops. Herewith, departure and arrival times at stops can be determined for a timetable, but for the actual modelling of the headways, the influence of the speeds should be fed back. An iterative approach with an external speed profile model can be an option. To be able to include the speed-headway relation into the model, speed levels and/or a discretisation of the open line based on, e.g., MA update times, can be included. The discretisation points on the open line can be considered for timing/speed, not for ordering or routing (similar as in alternative graph models). Alternatively, continuous speed choice variables may be included, together with constraints that define train braking distances based on the speed.

The time-indexed MILP inherent space and time discretisation limits the model's flexibility in terms of infrastructure and speed modelling. The model's set packing constraints, which ensure that at most one train is assigned to a time-space resource, requires a revision for the modelling of moving-block headways. Either the open line needs to be (dynamically) divided into short space intervals, or the restriction of one train needs be adjusted. In the first case, the problem arises that train braking distances are train and speed dependent. In the latter case, the open line could be considered as one space resource. Then the maximum capacity should be calculated in terms of the minimum headway, which in its turn is not possible to model on microscopic level.

The time discretisation could be implemented when it corresponds to (an approximation of) MA update times. In general, applying time discretisation would downplay the differences between trains leading to general approximations.

4 A Mathematical Model for Optimised Moving-Block Railway Traffic Management

4.1 Moving-Block Railway Traffic Management: Modelling Framework

The developed real-time traffic management model for moving-block is considered in connection with other PERFORMINGRAIL WPs, the main ones being WP2, WP3 and WP5. The connections are described by presenting the components of the traffic management system for moving block and its input/output relations, as illustrated in Figure 6.

Input:

- Parameter values related to moving-block signalling requirements from WP2, see Section 5.3.
- Parameter values related to GNSS train localisation and integrity requirements from WP3, see Section 5.2.
- Input data of traffic state from case study: infrastructure, rolling stock, timetable and, possibly, disturbances. Obtained through the BRASS microscopic simulator from WP5.

Model components:

- Early-warning prediction of hazardous moving-block events, see Section 5.4.
- Conflict resolution for moving-block to minimise impact of detected conflicts and hazards in terms of delay, see Section 4.3.

Output:

- Optimised real-time traffic plan based on the taken rescheduling decisions, to be communicated to the BRaSS simulator in WP5.

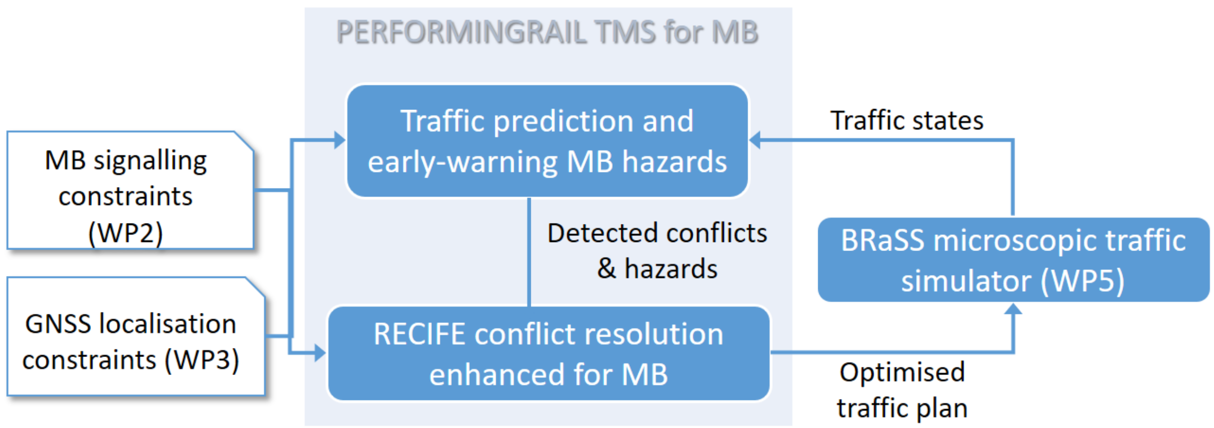

Figure 6. Modelling framework of the proposed traffic management models for MB.

4.2 The RECIFE-MILP Conflict Detection and Resolution Model

We present here the RECIFE-MILP formulation presented in the previous works of Pellegrini et al. (2014) and Pellegrini et al. (2015). It models the infrastructure at the microscopic level and implements the route-lock sectional-release interlocking system. The tracks are divided into track vacancy detection section, i.e., track segments on which the presence of a train is automatically detected. In RECIFE-MILP, the track vacancy detection sections are assumed to be *track circuits*. *Block sections* represent groups of track circuits whose access is controlled by a signal. Moreover, before a train can *occupy* a block section, all its track circuits must be reserved for the train itself.

Figure 7 depicts an example of this microscopic model of the infrastructure for a simple infrastructure. Track circuits are named tc and signals are named s, both indexed with a progressive number. Signals indicate the availability of block sections in a precise direction: for example, signal $s2$ is for block sections $s2 - s4$ including $tc1$, $tc2$ and $tc3$, in this order, and $s2 - s5$ including $tc1$, $tc2$ and $tc6$, in this order. Suppose that two trains ($t1$ and $t2$) cross the infrastructure: $t1$ going from $s2$ to $s8$ (using route $r1$ including block sections $s2 - s4$ and $s4 - s8$) and $t2$ going from $s3$ to $s9$ (using route $r2$ including block sections $s3 - s5$ and $s5 - s9$). We will see in the following how their passing through track circuits is represented in the model.

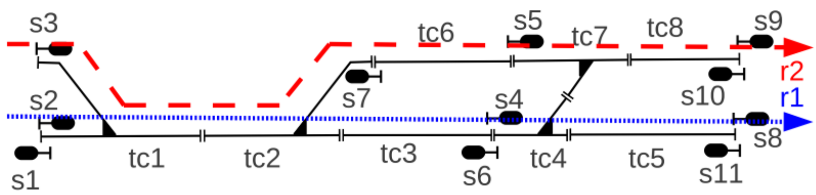

Figure 7. Example of microscopic representation of a railway infrastructure.

RECIFE-MILP uses the following sets:

- T the set of trains;
- Θ the set of train types;
- R_t the set of routes available to train $t \in T$, with $R = \cup_{t \in T} R_t$ the total set of routes;
- TC_t the set of track circuits which can be used by train $t \in T$;
- TC^r the set of track circuits belonging to route $r \in R$;
- $OTC_{ty,r,tc}$ the set of track circuits such that, if a train $t \in T$ of type $ty \in \Theta$ traverses them along $r \in R_t$ and has its head at their end, it holds that t's tail has not yet left tc. $OTC_{ty,r,tc} = tc$ if t is shorter than tc itself;
- $S_t, TCS_{t,s}$ the set of stations where $t \in T$ has a scheduled stop and set of track circuits that can be used by t for stopping at $s \in S_t$;

and parameters:

- tc_0 and tc_∞ dummy track circuits representing entry and the exit locations of the infrastructure considered;
- $sched_t$ scheduled arrival time of train $t \in T$ at destination;
- ty_t type corresponding to train $t \in T$ (train characteristics);

- $init_t, exit_t$ earliest time at which train $t \in T$ can be operated and earliest time at which it can reach its destination given $init_t$, the route assigned in the timetable and the intermediate stops;
- $i(t', t)$ indicator function equal to 1 if $t' \in T$ and $t \in T$ use the same rolling stock and t results from the turnaround, join or split of t', 0 otherwise;
- $ms_{t,t'}$ minimum separation between the arrival and the departure of trains $t, t' \in T$ using the same rolling stock;
- $rt_{r,ty,tc}, ct_{r,ty,tc}$ minimum running and clearing time of $tc \in TC^r$ along $r \in R$ for a train of type $ty \in \Theta$;
- $ref_{r,tc}$ reference track circuit for the reservation of $tc \in TC^r$ along $r \in R$, depending on block section structure and interlocking system;
- $e(tc, r)$ indicator function equal to 1 if track circuit $tc \in TC^r$ belongs to either the first or the last block section of $r \in R$, 0 otherwise;
- $bs_{r,tc}$ block section including track circuit $tc \in TC^r$ along $r \in R$;
- for_{bs}, rel_{bs} formation and release time for block section bs;
- $dw_{t,s}, a_{t,s}, d_{t,s}$ minimum dwell time, scheduled arrival and scheduled departure times for train $t \in T$ at station $s \in S_t$;
- $p_{r,tc}\ s_{r,tc}$ track circuits preceding and following $tc \in TC^r$ along $r \in R$;
- w_t weight of train $t \in T$ delay;
- M a large constant.

Note that all trains of the same type are considered to be planned to travel in the same way across a track circuit along a route. If train-specific running and clearing times need to be considered, $rt_{r,ty,tc}$ and $ct_{r,ty,tc}$ can be indexed on the train itself rather than on its type.

We also make use of the following variables, which include the binary routing and scheduling decisions as well as continuous variables used to evaluate the travel time and potential delays of the trains on each track circuit:

- $sU_{t,tc}, eU_{t,tc}$ continuous positive variables representing the time at which $tc \in TC_t$ starts and ends being utilised, i.e., blocked, by $t \in T$;
- $x_{t,r}$ binary variable equal to 1 if train $t \in T$ uses route $r \in R_t$, 0 otherwise;
- $y_{t,t',tc}$ binary variable equal to 1 if train $t \in T$ utilizes track circuit tc before train t', such that index t is smaller than index t' $(t < t')$, with $tc \in TC_t \cup TC_{t'}$ and otherwise;
- $o_{t,r,tc}$ time in which $t \in T$ starts the occupation of $tc \in TC^r$ along $r \in R$;
- $l_{t,r,tc}$ longer stay of $t \in T$'s head on $tc \in TC^r$ along $r \in R_t$, due to dwell time and scheduling decisions (delay);
- $D_{t,s}$ delay suffered by train $t \in T$ when stopping at station $s \in S_t$;
- D_t delay suffered by train $t \in T$ when exiting the infrastructure.

All these variables are imposed to be non-negative.

Depending on the objective function used, mirroring the literature, we can assess either overall or consecutive delay. The former, simply named *delay* from now on, is the delay suffered by a train when exiting the infrastructure: it is the difference between the exit arrival time and the one scheduled in the timetable. Let us remark that if a train reaches its destination within the infrastructure itself, it will still be considered as if it was exiting. The *consecutive delay* is the delay cumulated in the infrastructure due to traffic. For example, if a train enters the infrastructure with five minutes of delay and it exits five minutes later than planned not encountering traffic, then its delay will be of five minutes and its consecutive delay will be null. Instead, if within the infrastructure the train has to give precedence to others and it cumulates two further minutes, it will have a delay of seven minutes and a consecutive delay of only two.

By default, RECIFE-MILP minimizes the weighted total delay suffered by trains when exiting the infrastructure or arriving at a planned stop:

$$min \sum_{t \in T} w_t (D_t + \sum_{s \in S_t} D_{t,s}) . \tag{1}$$

The model has to respect the following sets of constraints:

$$o_{t,r,tc} \geq init_t x_{t,r} \qquad \forall t \in T, r \in R_t, tc \in TC^r, \tag{2}$$

$$o_{t,r,tc} \leq M x_{t,r} \qquad \forall t \in T, r \in R_t, tc \in TC^r, \tag{3}$$

$$o_{t,r,tc} = o_{t,r,p_{r,tc},} + l_{t,r,p_{r,tc},} + rt_{r,ty_t,p_{r,tc},} x_{t,r} \qquad \forall t \in T, r \in R_t, tc \in TC^r, \tag{4}$$

$$o_{t,r,s_{r,tc}} \geq \sum_{\substack{s \in S_t; \\ tc \in TCS_{t,s} \cap TC^r}} d_{t,s} x_{t,r} \qquad \forall t \in T, r \in R_t, tc \in \bigcup_{s \in S_t} TCS_{t,s} , \tag{5}$$

$$o_{t,r,s_{r,tc}} \geq \sum_{\substack{s \in S_t; \\ tc \in TCS_{t,s} \cap TC^r}} d_{t,s} x_{t,r} \qquad \forall t \in T, r \in R_t, tc \in \bigcup_{s \in S_t} TCS_{t,s} , \tag{6}$$

$$D_{t,s} \geq \sum_{r \in R_t} \sum_{tc \in TCS_{t,s} \cap TC^r} (o_{t,r,tc} + rt_{r,ty_t,tc} x_{t,r}) - a_{t,s} \qquad \forall t \in T, s \in S_t, \tag{7}$$

$$D_t \geq \sum_{r \in R_t,} o_{t,r,t_\infty} - sched_t \qquad \forall t \in T, \tag{8}$$

$$\sum_{r \in R_t,} x_{t,r} = 1 \qquad \forall t \in T, \tag{9}$$

$$\sum_{\substack{r \in R_t, \\ tc \in TC^r: \\ p_{r,tc} = tc_0}} o_{t,r,tc} \geq \sum_{\substack{r \in R_t, \\ tc \in TC^r: \\ s_{r,tc} = tc_\infty}} o_{t',r,tc} + (ms_{t,t'} + rt_{r,ty_{t'},tc}) x_{t',r} \qquad \forall t, t' \in T: i(t', t) = 1, \tag{10}$$

$$\sum_{\substack{r \in R_t, \\ s_{r,tc_0} = tc}} x_{t,r} = \sum_{\substack{r \in R_{t'}, \\ p_{r,tc_\infty} = tc}} x_{t',r} \qquad \forall t, t' \in T: i(t', t) = 1, tc \in TC_r: p_{r,tc} = tc_0, \tag{11}$$

$$\sum_{\substack{tc \in TC_t: \\ \exists r \in R_t: \\ p_{r,tc} = tc_0}} sU_{t,tc} \leq \sum_{\substack{tc \in TC_{t'}: \\ \exists r \in R_{t'}: \\ s_{r,tc} = tc_\infty}} eU_{t',tc_{t,t'}} \qquad \forall t, t' \in T: i(t', t) = 1, \tag{12}$$

$$sU_{t,tc} = \sum_{\substack{r \in R_t: \\ tc \in TC^r}} (o_{t,r,ref_{r,tc}} - for_{bs_{r.tc}} x_{t,r}) \qquad \forall t \in T, tc \in TC_t: \\ (\nexists t' \in T: i(t', t) = 1) \vee (\forall r \in R_t: ref_{r,tc} \neq s_{r,tc_0,}) \tag{13}$$

$$eU_{t,tc} = \sum_{\substack{r\in R_t:\\ tc\in TC^r}} o_{t,r,tc} + \sum_{\substack{tc\in TC^r:\\ tc\in OTC_{ty_t,r,tc}}} l_{t,r,tc'} + (rt_{r,ty_t,tc} + ct_{r,ty_t,tc} + rel_{bs_{r.tc}})x_{t,r} \qquad \forall t\in T, tc\in TC, (14)$$

$$eU_{t,tc} - M(1 - y_{t,t'tc}) \le sU_{t',tc} \qquad \forall t,t'\in T: t<t', tc\in TC_t\cap TC_{t'}: i(t',t)\sum_{r\in R_t} e(tc,r) = 0 \wedge i(t',t)\sum_{r\in R_{t'}} e(tc,r) = 0, \qquad (15)$$

$$eU_{t,tc} - My_{t,t'tc} \le sU_{t,tc} \qquad \forall t,t'\in T: t<t', tc\in TC_t\cap TC_{t'}: i(t,t')\sum_{r\in R_t} e(tc,r) = 0 \wedge i(t,t')\sum_{r\in R_{t'}} e(tc,r) = 0. \qquad (16)$$

Constraints (2) and (3) force train t to be operated no earlier than $init_t$ on its chosen route and set all track circuit occupations to 0 on the alternative routes.

In Constraints (4), a train starts occupying a given track circuit after spending its free-network running time in the preceding one plus the longer stay cumulated there (if the route is used).

Constraints (5) and (6) ensure that train t which stops at station s along route r does not leave track circuit $tc \in TCS_{t,s}$ before the scheduled departure time from s, and in any case spends at least its minimum dwell time on tc.

Constraints (7) and (8) quantify non-negative delay at each station when train t has a scheduled stop and at its exit from the infrastructure. Remark that t is assumed to stop at the end of the track circuit where the stop occurs. The non-negativity of the continuous variables allows to neglect the negative contribution of trains arriving in advance.

In Constraints (9), a single route is chosen for train t.

Constraints (10), (11) and (12) are used to guarantee consistency for trains using the same rolling stock, i.e., the respect of the minimum separation time between arrival and departure of such trains, the use of the same arrival and departure track circuit, and the overlapping utilisation times of this track circuit.

In Constraints (13), a train's utilisation of a track circuit starts as soon as the train starts occupying track circuit $ref_{r,tc}$ along one of the routes including it, minus the formation time. Constraints (13) are imposed as inequalities ($\le$) when they concern a track circuit of the first block sections of the route ($ref_{r,tc} = s_{r,tc_0}$ and the train t results from the turnaround, join or split of one or more other trains ($\exists t' \in T: i(t',t) = 1$). This is a consequence of the need of keeping platforms utilised. Indeed, if t results from t', Constraints (12) ensure that the track circuit where the turnaround takes place starts being reserved by t as soon as t' arrives. However, t needs to wait at least for a time ms before departing. The occupation of the track circuit by t is however starting from its actual departure, for guaranteeing the coherence of the occupation variables and the running time (Constraints (4)). Hence, t's reservation starts much earlier than its occupation.

In Constraints (14), the utilisation of a track circuit lasts till the train exits it along any route, plus the release time. If the train is long enough to keep occupying the track circuit when its head is at the end of the following ones (the ones included in set $OTC_{ty_t,r,tc}$), also the longer stay of the train on these further track circuits has to be accounted for.

Finally, Constraints (15) and (16) ensure that the track circuit utilizations by two trains do not overlap. This must hold unless they use the same rolling-stock and the track circuit is at the extreme part of their routes, where the reutilization must take place.

We refer the reader to Pellegrini el al. (2014) and Pellegrini et al. (2015) for additional details about the above formulation.

4.3 Moving-Block Conflict Detection and Resolution

The original RECIFE-MILP model formulation is altered and extended to approximate moving-block rescheduling. To this end, a block-independent discretisation of the open line, speed profile options and headways based on absolute braking distances are introduced.

The track is divided into switch areas and open line stretches, as illustrated in Figure 8. A switch area consists of one or more switches, which are modelled as track detection sections. An open line stretch connects two switch areas. Note that parallel tracks connecting two switch areas are considered as separate open line stretches. Track locations are defined to discretise open line stretches into a grid to enable the approximation of moving block by fixed virtual sections. The model considers the track locations, or the (virtual) section entry points, instead of the section as a whole. For this, track circuits (tc) are replaced by track locations (tl) in the moving-block model formulation.

In the following, the RECIFE-MILP enhancement towards moving block is described. First, the introduction of speed level options into the model is considered. Second, the redefinition of utilisation times to approximate moving-block blocking times is addressed.

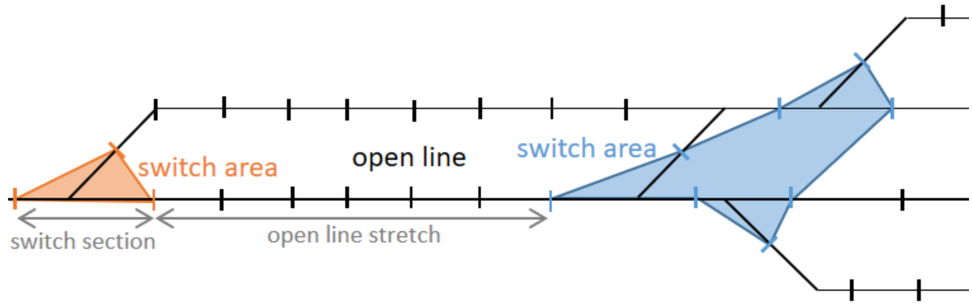

Figure 8. Modelling of the track as switch areas consisting of switch sections and discretised open line stretches.

4.3.1 Introduction of Speed

The notion of speed in the model is extended by introducing the option to run at scheduled speed next to the option to run at maximum speed. To this end, two binary variables, $v^m_{t,r,tl}$ and $v^s_{t,r,tl}$, are defined to indicate whether ($= 1$) or not ($= 0$) train $t \in T$ is running over track location $tl \ \in TL_t$ along route $r \in R_t$ according to the maximum or scheduled speed level, respectively.

The two speed levels are assumed to correspond to speed profiles. The speed profiles are provided as input to the model. The required input for each train type is the exact speed at every track location and the running time between subsequent track locations for the two speed profiles.

Maximum speed profiles refer to speed profiles obtained when the train runs at maximum power by using maximum acceleration, a target cruising speed equal to the (line) speed limit, and a maximum deceleration. Such a speed profile allows the train to reach the minimum technical running time over a given route.

Scheduled speed profiles refer to the scheduled running time which adds running time supplements to the minimum technical running time. Scheduled speed profiles are hence characterised by a lower target cruising speed and possibly a coasting phase. Specifically the coasting phase aims at using allocated running time supplements to save energy by switching off the engine and let the train being decelerated by the existing motion resistances. Figure 9 shows examples of speed profiles corresponding to maximum and scheduled speed, taking into account a lower speed limit at the beginning.

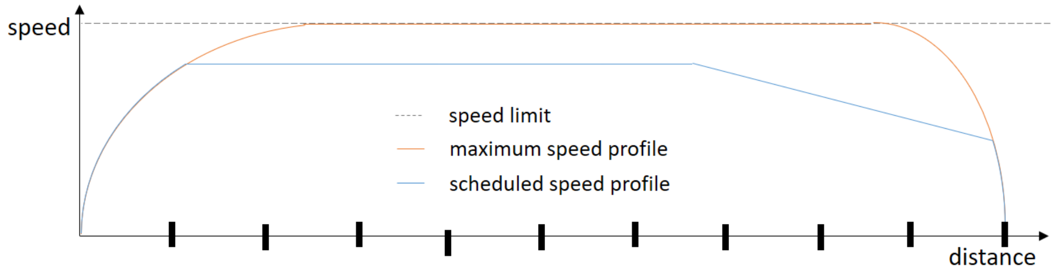

Figure 9 .Example of speed profiles corresponding to maximum speed and scheduled speed between two stops.

Constraints (5) are reformulated to include the effects of the speed options on the train running times. In the reformulation, the minimum running times, i.e., the running times corresponding to the maximum speed profile, are used as reference. Additionally, we introduce parameters $\Delta rt_{ty,r,tl}$ to represent the additional running time for trains of type $ty \in \theta$ when passing track location $tl \in TL^r$ along route $r \in R$ at scheduled speed, with respect to the minimum running time.

In reformulated Constraints (17), the difference between occupation starting times of consecutive locations along a train's route ($p_{r,tl}$ and tl) is set equal to the running time corresponding to the speed level plus the longer stay of the train. The longer stay comprises the extra occupation time due to a planned stop or an unplanned slow down (or stop) due to traffic. Note that the longer stay variables can be positive both when the train is running at scheduled or maximum speed.

$$o_{t,r,tl} = o_{t,r,p_{r,tl}} + l_{t,r,p_{r,tl}} + rt_{ty_t,r,p_{r,tl}} x_{t,r} + \Delta rt_{ty_t,r,p_{r,tl}}\, v^s_{t,r,p_{r,tl}} \quad \forall t \in T, r \in R_t, tl \in TL^r \quad (17)$$

The speed variables are introduced into the model under the following restrictions. Constraints (18) ensure that exactly one speed level is chosen for a train-location pair if the location lies on the assigned train route, i.e., $x_{t,r} = 1$. If not, i.e., $x_{t,r} = 0$, then no speed level is chosen for consistency with Constraints (Equation17).

$$v^m_{t,r,tl} + v^s_{t,r,tl} = x_{t,r} \quad \forall t \in T, r \in R_t, tl \in TL^r \quad (18)$$

4.3.2 Redefinition of Utilisation

The utilisation or blocking times are the core of the model reformulation. Moving-block principles are incorporated to approximate moving-block minimum train separation, based on absolute braking distances. To this end, Constraints (13) to (16) are reconsidered.

Constraints (13) ensure the formation and approach times. The formation time $for_{r,tl}$ comprises the setup and reaction time components, which can differ depending on the location tl and the route r, in particular whether or not the location is part of a switch section. The approach time is determined through the definition of a reference location earlier on the route such that a train's presence there triggers the reservation, i.e., the start of utilisation, of the considered track location. In other words, a train cannot pass the reference location unless the whole track between this location and track location tl is free.

To be in line with moving-block principles, we redefine the reference location of a track location to be based on the absolute braking distance depending on the approaching speed. The safe train separation distance at track location tl for speed v_{tl}, $d_{v,tl}$ (in m), is determined based on the braking distance and safety margin sm (in m) as follows:

$$d_{v,tl} \; = \; \frac{v_{tl}^2}{2b_{ty}+gG_{v,tl}} + sm,$$

with v_{tl} the speed (in m/s) at location tl , b_{ty} the braking rate (in m/s^2) of train type ty, g the gravitational acceleration (in m/s^2) and $G_{v,tl}$ the track gradient as function of the speed and track location, either positive (uphill) or negative (downhill). The braking rate is assumed to be constant per train type but can be reinterpreted to also depend on the speed. The track gradient can be approximated by taking a (weighted) average of the gradients of the track elements within the braking distance.

For a given track location, this formula provides a brake indication point per speed level for every train type, to which b is related. A reference brake location of the considered tl is defined as the last discrete track location before the brake indication point of tl itself. In this reformulation, the reference brake location does not only depend on tl and the chosen route, but also on the speed and the type of the approaching train. We define two reference locations, one for each speed level: $ref_{ty,r,tl}^{m}$ and $ref_{ty,r,tl}^{s}$, with $ty \in \theta, r \in R$ and $tl \in TL$.

In addition to the redefinition of the reference (brake) location, we enable the modelling of continuous braking curve supervision. The utilisation of a track location by an approaching train should start at the moment the train passes the brake indication point corresponding to its type, speed and route, rather than the associated reference location. The passing times, or (physical) occupation starting times, at reference locations are known in the model. For the approximation of the passing times at brake indication points, we introduce 'reservation lag' parameters $lag_{ty,r,tl}^{m}$ and $lag_{ty,r,tl}^{s}$. For a specific speed level, the reservation lag parameter indicates the time interval by which the reservation of location tl along route r for a train of type ty can be postponed, with respect to the occupation starting time of the corresponding reference location.

Figure 10 illustrates the modelling concept. A track location approached by a train is considered as reference brake location for both approaching speeds, maximum and scheduled. In case the train runs at maximum speed, the track location to be reserved lies further ahead than for scheduled speed. Also, the reservation lags for maximum and scheduled speed are not equal, as they depend on the start of the braking curve.

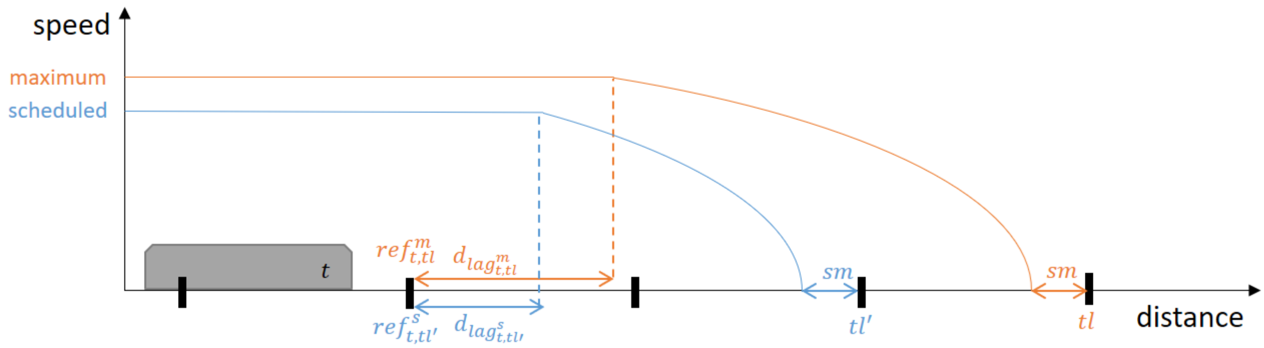

Figure 10. Moving Block modelling concept in RECIFE-MILP.

Implementing this modelling concept into Constraints (13) would result in the following non-linear constraints:

$$sU_{t,tl} = \sum_{\substack{r\in R_t:\\ tl\in TL^r}} \begin{array}{c}((o_{t,r,ref^m_{ty_t,r,tl}} + lag^m_{ty_t,r,tl})v^m_{t,r,ref_{ty_t,r,tl}} + (o_{t,r,ref^s_{ty_t,r,tl}} + lag^s_{ty_t,r,tl})v^s_{t,r,ref_{ty_t,r,tl}} \\ -for_{r,tl}x_{t,r}),\end{array}$$

with $t \in T$ and $tl \in TL_t$. Instead, two inequality constraints are formulated based on the two speed options. For trains approaching at scheduled speed, we only want to consider reference location $ref^s_{ty_t,r,tl}$ and reservation lag $lag^s_{ty_t,r,tl}$. This results in the following relation between the utilisation starting time and the occupation starting time of the `scheduled speed' reference location:

$$sU_{t,tl} = \sum_{\substack{r\in R_t:\\ tl\in TL^r}} (o_{t,r,ref^s_{ty_t,r,tl}} + (lag^s_{ty_t,r,tl} - for_{r,tl})x_{t,r}),$$

with $t \in T$ and $tl \in TL_t$. Similarly, for trains approaching at maximum speed we consider reference location $ref^m_{ty_t,r,tl}$and reservation l $lag^m_{ty_t,r,tl}$, resulting in the following relation:

$$sU_{t,tl} = \sum_{\substack{r\in R_t:\\ tl\in TL^r}} (o_{t,r,ref^m_{ty_t,r,tl}} + (lag^m_{ty_t,r,tl} - for_{r,tl})x_{t,r}),$$

with $t \in T$ and $tl \in TL_t$. The resulting inequality constraints are

$$sU_{t,tl} \le \sum_{\substack{r\in R_t:\\ tl\in TL^r}} (o_{t,r,ref^s_{ty_t,r,tl}} + (lag^s_{ty_t,r,tl} - for_{r,tl})x_{t,r}) \quad \forall t \in T, tl \in TL_t \tag{19}$$

and

$$sU_{t,tl} \le \sum_{\substack{r\in R_t:\\ tl\in TL^r}} (o_{t,r,ref^m_{ty_t,r,tl}} + (lag^m_{ty_t,r,tl} - for_{r,tl})x_{t,r}) + Mv^s_{t,r,ref_{ty_t,r,tl}}) \; \forall t \in T, tl \in TL_t. \tag{20}$$

Constraints (19) cover the case of scheduled speed. As by definition a shorter braking distance is associated with a lower speed, $ref^s_{ty_t,r,tl}$ lies closer to tl than $ref^m_{ty_t,r,tl}$. Hence, these constraints are not restrictive for trains running at maximum speed. Constraints (20) cover the case of maximum speed. If a train runs at scheduled speed, $v^s_{t,r,tl} = 1$, these constraints are trivially satisfied.

With Constraints (14), the utilisation of a track location, as entry point of a track circuit, ends with the release of the succeeding location, as exit point of the track circuit. In a moving-block system, the same principles hold for track locations in switch sections, which are modelled as fixed-block sections. The constraints are, however, to include the speed options. Speed-dependent running times are already accounted for due to reconsideration of occupation times in Constraints (17). We propose to include speed-dependent clearing times in a similar way. Besides the minimum clearing time, based on the maximum speed, an additional clearing time component can be predetermined for the case of scheduled speed.

Constraints (21) is the reformulation of Constraints (14), describing the end of utilisation of track locations that are part of a switch section, i.e., $s(tl) = 1$. For track locations that are not part of a switch section, i.e., $s(tl) = 0$, the location itself is considered instead of the succeeding one. This makes it possible to leave out the running time over the virtual section, better approximating moving-block blocking times. Constraints (23) capture the end of utilisation for moving-block locations.

Note that the last track location of a switch area is treated as an open line location, as it is the entry point of the following virtual section. Whether or not a location is the last of a switch area depends on the running direction.

$$eU_{t,tl} = \sum_{\substack{r \in R_t: \\ tl \in TL^r}} (o_{t,r,s_{r,tl}} + (ct_{ty_t,r,s_{r,tl}} + rel_{r,s_{r,tl}})x_{t,r} + \Delta ct_{ty_t,r,s_{r,tl}} v^s_{t,r,s_{r,tl}} + \sum_{\substack{tl' \in TL^r: \\ tl' \in OTL_{ty_t,r,tl} \setminus \{tl\}}} l_{t,r,tl'} \quad)$$

$$\forall t \in T, tl \in Tl: s(tl) = 1 \ \wedge \ \nexists r \in R_t: s(s_{r,tl}) = 0, \quad (21)$$

$$eU_{t,tl} = \sum_{\substack{r \in R_t: \\ tl \in TL^r}} (o_{t,r,tl} + (ct_{ty_t,r,tl} + rel_{r,tl})x_{t,r} + \Delta ct_{ty_t,r,tl} v^s_{t,r,tl} + \sum_{\substack{tl' \in TL^r: \\ tl' \in OTL_{ty_t,r,tl}}} l_{t,r,tl'} \quad)$$

$$\forall t \in T, tl \in Tl: s(tl) = 0 \ \vee \exists r \in R_t: s(s_{r,tl}) = 0,. \qquad (22)$$

4.4 Parameters for Modelling Moving-Block Operations

An important concept in the modelling of railway operations, is blocking time theory. With blocking time theory, the minimum headway is based on the time a track part is assigned to a train and hence, blocked for other trains. The blocking time starts when a train requests the track part for its route and ends after the route has cleared the track part after traversing it. Blocking time theory can also be used to detect track conflicts by overlapping blocking times (Hansen & Pachl, 2014).

In general, a blocking time can be constituted by the following components:

- Setup time: request, set and, in an interlocking area, lock the route;
- Reaction time: perceive and react to speed indication;
- Approach time: run from speed indication to considered track part;
- Running time: run over track part, including possible dwell time;
- Clearing time: run until rear of the train has cleared track part;
- Release time: release the route.

4.4.1 Blocking Time Components: Fixed-Block vs Moving-Block

As blocking time theory is developed for traditional multi-aspect fixed-block signalling system, a reinterpretation of the components is needed to apply it to the moving-block system. Here, we shortly go over the different components to point out the differences between fixed and moving block, of which the key features are summarised in Table 2.

Table 2. Blocking time parameter values for Fixed and Moving block operations.

Component	Fixed Block	Moving Block
Setup time	Route + signals	Route + RBC communication
Reaction time	Perception and control application (by human)	Perception and control application (by human or ATO)
Approach time	Fixed number of blocks	Braking distance + safety margin
Running time	Block section	Infinitesimal
Clearing time	Train length	Train length
Release time	Route + signals	Route + RBC communication

The difference between fixed and moving-block in terms of **setup time**, relates to the shift from trackside to onboard signalling. In the fixed-block system, trackside signals are set to communicate to trains about whether or not to proceed, while in the moving-block system, the MA is communicated to the trains by the RBC.

The **reaction time** component remains similar to fixed-blocks, as the MA information, whether received through trackside signal or onboard, needs to be translated into an action. However, in moving-block there is a higher incentive to use ATO, which would replace the human reaction time by a shorter time for the system to react.

The **approach time** significantly changes in moving-block systems as it relates to the time to traverse the minimum separation distance. So, it is no longer derived from a fixed number of blocks, but by the absolute braking distance supplemented with a safety margin.

The **running time** component is the other main change. In fixed-block, the track is considered in terms of block sections, which take a significant amount of time to travers. In moving-block, the track is considered in terms of points, which take infinitesimal time to travers.

Probably the most similar component for the fixed and moving-block systems, is the **clearing time**. The time it takes to run with a train length over a point on the track, whether the end of block section or not, remains the same.

With the introduction of onboard train positioning and train integrity monitoring, the **release time** no longer considers trackside train detection. Whereas in fixed-block systems, the signals need to be released, i.e., set to the default value, this subcomponent can be omitted.
As result of the above reported considerations, the blocking time stairways of a train moving under MB on an open track assumes the shape of a bandwidth rather than a rectangular block (as the running time component becomes infinitesimal as no block section exists), as it is illustrated in Figure 11. For interlocking areas with switches, the distance-time diagram of the track will instead be rectangular as movable infrastructure elements still require the presence of fixed-blocks to prevent safety-critical events in case of switch failures.

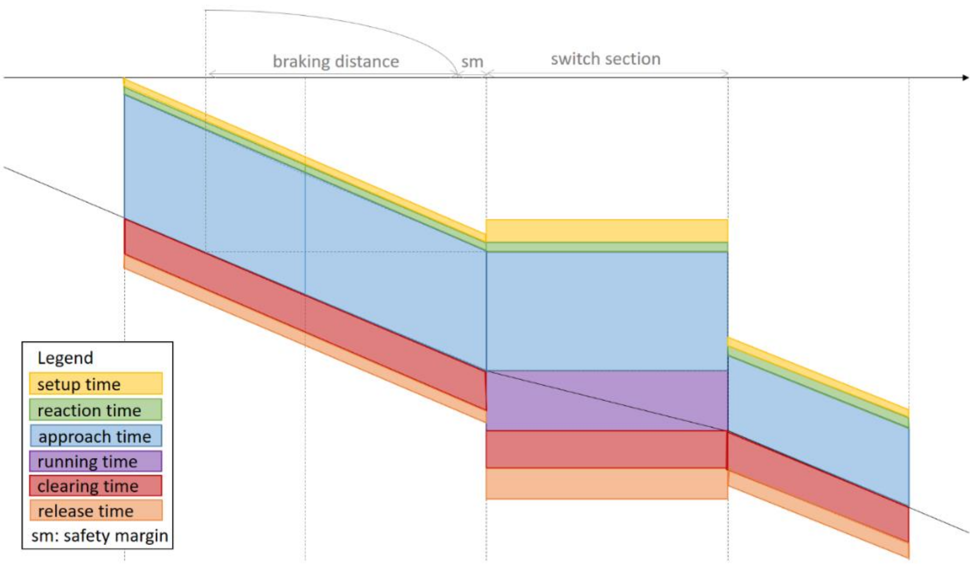

Figure 11. Blocking time 'stairways' in moving-block system in which a switch section is considered as a fixed block.

4.4.2 Blocking Time Parameters for Moving Block

Here, we list the relevant parameters for the definition of moving-block blocking times.

Setup time
- Route setting time
- MA update time

Reaction time
- Human reaction time
- ATO reaction time

Approach time
- Train braking rate
- Track gradients

Clearing time
- Speed measurement error

Release time
- TPR reporting time
- GNSS update frequency
- GNSS positioning error
- TIM update frequency
- TIM reporting time

5 Early-Warning Prediction of Hazardous Moving-Block Conditions

5.1 Requirements for Safe Moving-Block Railway Operations

The "requirements for safe moving block operation" was addressed through the initial phase of WP1 (T1.1 and T1.2) by analysing, verifying, and enhancing principles and system specifications defined academic and industrial projects on moving block systems such as X2Rail-1, X2Rail-2, X2Rail-3, MOVINGRAIL, ASTRAIL, and NGTC to enable safety and standard performance levels of moving-block railway operations.

A review of the relevant academic and industrial projects was undertaken to define the moving block system in terms of system architecture, existing approaches and variants, and the main system functionalities. Subsequently, a high-level architecture of ETCS-L3 and the related subsystems were defined, as well as how ETCS L3 would interface with external systems and actors using an explanatory use case for normal train movement. Then, the two main moving block approaches (Full Moving Blocks and Fixed Virtual Blocks) have been reviewed alongside the four induced system variants. Finally, the main moving block system functionalities (e.g., train integrity modelling, train localisation, and determination of track status) were specified.

After reviewing the moving block systems including ETCS L3 and virtual coupling, the outcomes represent a rigorous baseline specification, including system requirements, engineering and operational rules, and hazard analysis for the subsequent modelling and Verification and Validation (V&V) activities of the project. The specification has been carried out considering:

- Moving Block system architecture, variants, existing approaches, and main functionalities
- Moving Block system requirements including time constraints
- Moving Block operational and engineering Rules
- Moving Block hazard analysis including GNSS-specific hazards
- Moving Block use cases detailing external actors, related internal functions, expected system behaviour, system parameters, and safety hazards.

Table 3 synthesizes result data of the X2Rail-3 project related to ETCS L3 MB system and provide an overview of the number of defined requirements and rules.

Table 3. Number of defined requirements, operational and engineering rules in (D4.2 X2Rail-3, 2020).

Topics		Requirements	Operational Rules	Engineering Rules
GENERIC	Train Location	14	1	2
	Track Status	19	6	1
	Reserved Status	6	0	1
	Fixed Virtual Blocks	1	0	4
	Trackside Train Detection	7	0	2
OPERATIONAL SCENARIOS AND DEGRADED MODES	Points Control	4	0	2
	Movement Authorities	12	0	0
	EoA Exclusion Area	2	0	0
	Start of Train	15	6	0
	On Sight movement	0	5	1
	Staff Responsible (SR) movement	5	0	1
	First MA	2	0	0
	Loss of Communication	6	4	1
	Recovery manag. after loss of com	3	1	0
	Radio hole	8	0	4
	Reverse movement	5	0	0
	End of Mission	4	1	2
	Loss of Train Integrity	10	2	6
	Level Transition	2	2	1
	Trackside Initialisation	5	5	3
	Handover	3	0	1
	Shunting movement	4	3	1
	Joining	3	0	0
	Splitting	1	0	0
	Traffic Management System interface	1	0	0
	Total	142	36	33

Besides the synthesized data, SysML requirements diagrams were defined for both the main and timed requirements of the moving block system considering the system time constraints. Furthermore, hazards analysis introduced in X2Rail-3 were extended to focus on GNSS-specific hazards used on terrestrial transportation as well as the GNSS-based VBTS hazards as the most mature GNSS-based solution.

Finally, ETCS L3 MB use cases are defined as a specific context, involving L3 on-board, trackside and external actors and interfaces. Use cases are characterized by specific features, system hazards or applicable operational/engineering rules and related to functionalities of the ETCS L3 MB system (e.g., handover, train splitting), operation on trackside entities (e.g., track status initialization), coexistence of ETCS and non-ETCS traffic (e.g., ghost trains) and to a degraded mode (e.g., march in SR).

A detailed description of 16 use cases taking into consideration external actors, related internal functions, expected system behaviour, system parameters, and safety hazards were presented. A criterion has been set to select the most common use cases to represent their behaviour using SysML sequence diagrams. The criteria used to make this choice are the following:

- At least one modelled UC should be related to Virtual Coupling;
- Maximize the number of UCs related to train positioning;
- Maximize the number of impacted Oss according to the OS-UC mapping table.

According to all these inputs, the following use cases were described: Normal Train Movement, Supervising Distance in normal driving, Loss/Restore of Communications and Loss of Train Integrity. As the first phase of WP1 aimed to define a baseline for the ETCS MB/VC without defining a formal model, the use cases were modelled following a semi-formal approach. Using the SysML language, the behaviours were specified using Activity Diagrams and/or Sequence to highlight the interactions between components and using State Machine Diagrams for specifying components' inner behaviours.

The system definition and specifications provided by WP1 represents a baseline for the subsequent modelling activities that are part of PERFORMINGRAIL WP2.

5.2 GNSS Train Localisation and Integrity Requirements

From the developments performed during WP3, the relevant parameters, along with expected ranges and its corresponding rationales, for its consideration in train operations, are listed in Table 4.

Table 4. GNSS relevant parameters for train location and integrity.

Parameter	Range (with justification)
GNSS update rate	[0.01:1] (in seconds) Rationale: Most commercial grade GNSS receiver allow for GNSS update rates from 1Hz (latencies of 1s) to a maximum of 10Hz (latencies of 0.1s). Rate can be reduced to shorter values of 100Hz (0.01s) by using inertial measurement units (IMU) and a loosely coupled GNSS/INS integration. In any case this can be usually configured in the receiver and the best value will be dependent on the train speed. If the train moves at a maximum train speed of 250km/h, it will displace around 70m in 1 second. Therefore, for that maximum speed, the latency should be reduced to 0.01s for an uncertainty of ~0.7m. For more strict requirements on the train position error, faster rates might be necessary, depending on the expected maximum speed of the train.
GNSS positioning error	[0.5:2] (in 2D meters, at 2 σ) Rationale: For the railways use case, the relevant metric for the GNSS positioning error is the 2D (horizontal) error, or even 1D

	(along the railway). In any case, for standalone multi-frequency and multi-constellation, 1D/2D accuracies of 0.5m to 1m at 99.5% of the time (i.e., 2 σ) should be attainable. Depending on the processing strategy and assistance data, the positioning error could be reduced. Using differential GNSS techniques such as Real Time Kinematics (RTK), accuracies better than 0.5m could be achieved in real time. However, these require the deployment of additional infrastructure (GNSS base stations, connectivity between the base station and the GNSS receiver in the train, …). The usage of Precise Point Positioning (PPP) strategy with corrections obtained from a third-party provider (e.g., Galileo High Accuracy Service) might allow the reduction of the positioning error to few decimetres (<50cm).
GNSS Time-to-first-fix	<10 seconds Rationale: The Time-To-First-Fix (TTFF) is the time elapsed from the moment in which the GNSS is turned on and the first position fix under the target accuracy. For standalone systems, the specified TTFF for the level of accuracy mentioned above (2m) should be sufficient. Lower TTFF could be obtained with processing strategies such as Real Time Kinematics, but these require the deployment of additional infrastructure as mentioned above.
TIM update rate	[1:5] (in seconds) Rationale: The Train Integrity Monitoring update rate might in principle depend on the operations requirements, however this rate is limited to the update rate of the location system (established above). In any case, for operations, longer latencies (than GNSS) such as the ones specified here are deemed sufficient for operations.
TIM reporting: positioning and integrity	[0.01:1] (in seconds) Rationale: The reporting of the position is limited to the GNSS update latency. Integrity reports within the position report can be frozen between 2 consecutive TIM updates.

5.3 Impact of Signalling System Features on Safe Moving Block

A model of the MB system has been developed in WP2 using Stochastic Activity Networks (SANs) [SAN]. This model is introduced and discussed in Deliverables D2.2 and D2.3; it represents the movement of a train fleet on a track under the control of the trackside according to the *Normal Train Movement Use Case* and it aims to quantitatively evaluate the impact of some parameters (e.g., the maximum speed, the safety margin, the trackside reaction time, etc.) on the service offered by the system.

In addition, the developed model also allows to evaluate the effect of possible failures and analyse possible hazardous situations. Specifically, in the current version the following issues are considered for the performability analysis: integrity not confirmed by TIMS and loss of messages (Train Position Report and Movement Authority).

This model has been extended to enable a sensitivity analysis suitable to the objective of WP4. The goal is to provide indications about the impact of the signalling system features on the safe march of the trains. Specifically, we want to study the impact of fluctuations of signalling system features (i.e., RBC processing time, delays in the communication network, period between subsequent TPRs, driver reaction time) on the behaviour of the trains with respect to brake events, given as input the set of system configuration parameters (e.g., train headway, safety margin, train mechanical features).

5.3.1 Braking Curves

ETCS continuously supervises the position and speed of the trains that must remain within the allowed speed and distance limits. At this aim, the ETCS on-board unit must be able to predict the decrease of the train speed against a distance that must not be exceeded according to a mathematical model. This prediction (used by the ETCS on-board unit to calculate in real time the braking distance) is called "braking curve".
It is the responsibility of the driver to monitor the speed of the train and maintain it within the permitted speed but, if necessary, an emergency brake command is issued by ETCS on-board unit to avoid that the train exceed the allowed limits (e.g., the end of the MA). This emergency brake command overrides the driver and stops the train.
Several curves are drawn whose objective is to establish specific ***target locations*** (corresponding either to a speed reduction or to a stop location) before an emergency brake. "*Therefore, the ETCS on-board calculates in real time other supervision limits. They consist of locations that, when crossed by the train, will trigger some information to be given to the driver*" (ERA, 2020)
The analysis carried out in WP4 considers three of those curves, the description is taken from (ERA, 2020):

- *Indication (I)*: *"the I supervision limit leaves the driver enough time to act on the service brake so that the train does not overpass the permitted speed."*
- *Permitted (P)*: *"the P supervision limit in case of overspeed, leaves the driver an additional time to act on the service brake so that the train will not overpass the point beyond which ETCS will trigger the command of the brakes."*
- *Warning (W): "the W supervision limit provides an additional audible warning after the permitted speed has been overpassed."*
- *Emergency (EBD): "the braking curve related to the speed decrease due to the emergency brake is called Emergency Brake Deceleration (EBD) curve."*

Several input parameters are needed to enable the calculation of the braking curves and the on-board supervision: the train instantaneous position, speed and acceleration; the train data providing the necessary information about the vehicle's braking dynamics, and trackside data.

5.3.2 Modelled Behaviour

The trains movements are modelled *at a system level*, focusing on the exchange of messages between the on-board units and the trackside. *The movement of the train is discretized*, that is it is modelled as a sequence of steps (the duration of each step is constant and it is a parameter of the simulation).

After entering the line, each train periodically computes its position and speed and moves on the track based on the information about the End of Authority (EoA), which is dynamically updated by the trackside on the basis of the "known" position of the next danger point (which can be e.g. a switch or the tail of a preceding train), i.e. the Supervised Location minus a safety margin.

Each train periodically sends the Train Position Report to the trackside and, on the reception of a Train Position Report from a specific train, the trackside updates the extent of the Track Status Area associated with that train depending upon the communicated position and integrity status. Then the trackside updates and sends the Movement Authority for that train, such that the train is authorised to proceed safely up to the indicated EoA.

The on-board system makes sure that the train never overshoots the provided EoA by dynamically supervising that that the speed of the train aligns to the speed indication imposed by the ETCS braking curves.

Whenever a train is approaching an EoA, the on-board system asks the driver to brake in order to meet the speed limit imposed by the Indication Braking curve. After a certain time is elapsed (driver intervention time), if the speed limit of the Indication braking curve has not been respected, the on-board system informs the driver in order to meet the Permitted supervision speed limit, and if this latter is also disregarded it will provide the limit imposed by the Warning Braking curve. In the case that also the speed indication relative to the Warning braking curve is infringed then the trains will be automatically braked by the intervention of the Emergency brakes.

The model here defined assumes that whenever the speed imposed by one of the ETCS braking curves is overshot, the train will brake until its speed will align again to that imposed by the "Indication" braking curve profile.

Otherwise, if there is no need for braking (EoA far enough away, greater than the sum of the current position plus the indication braking distance), the on-board checks the possibility of accelerating (since the train is below the maximum speed) or continues at constant (maximum) speed.

5.3.3 Model Structure

The model is built according to a modular and compositional approach. The components of the model are the On-Board, the Communication Network and the Trackside sub-models.

The current model consists of one Trackside sub-model and *n* modules, each containing one On-Board and its associated Communication Network sub-model for the communication with the trackside.

Figure 12 shows the structure of the model. The sub-models may share global variables and they have their own local status. The figure shows how the sub-models are composed by superposing places (the circles in the figure) which store the global variables representing the Train Position Reports (*TPRmsg, TPRmessages*) and Movement Authorities (*MAmsg, MAmessages*) messages.

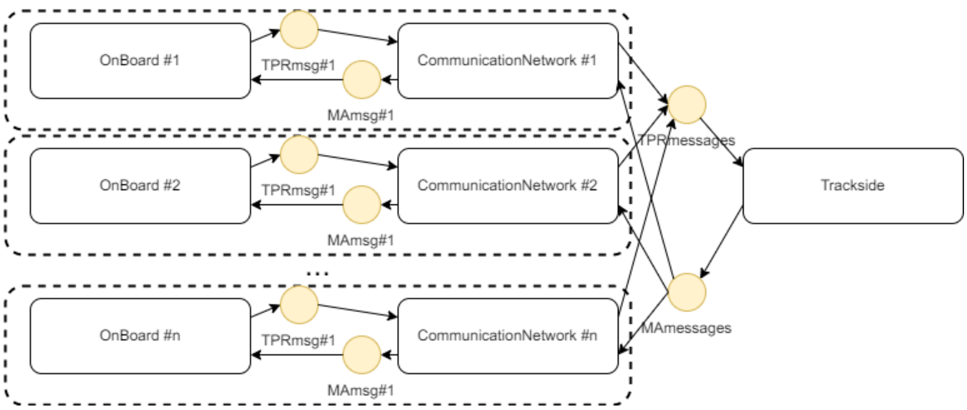

Figure 12. SAN Model Structure.

5.3.4 Model Parameters

Sensitivity Analysis Parameters

Table 5 reports the set of parameters used for sensitivity analysis.

Table 5. Moving Block signalling parameters and ranges investigated in the SAN analysis.

Computation Times				
Parameter	*Description*	*Range/Value*	*Model Name*	*Type*
RBC processing time	Computation by the Radio Block Center (RBC)	[0.5 : 5] (in seconds)	RBCprocessingTime	double
EVC processing time of the MA	This time includes the onboard translation of received MA into speed profile, as well as the speed indication computation by the European Vital Computer (EVC). Moreover, the given range leaves some room for more case of system functioning.	[1.5 : 4] (in seconds)	EVCProcessingTime	double
Communication Times				
Parameter	*Description*	*Range/Value*	*Model Name*	*Type*
Train to RBC communication time (TPR)	Communication time from the train to the RBC.	[0.5 : 2.5] (in seconds)	TPRnetDelay	double
RBC to Train communication time (MA)	MA broadcasting time	[0.5 : 2.5] (in seconds)	MAnetDelay	double
MA update time	Includes the Movement authority (MA) computation by the Radio Block Center (RBC) and broadcasting it to the train.	[1.5 : 10] (in seconds)	RBCprocessingTime + MAnetDelay + TPRnetDelay	double
Train Position and Integrity Report update time	The position report includes position and integrity. The position reporting time	[1 : 5] (in seconds)	TPRUpdatePeriod	double

	includes the communication time from the train to the RBC. This range considers GSM-R and GNSS.			
Train and Human Braking Reaction Time				
Parameter	*Description*	*Range/Value*	*Model Name*	*Type*
Indication Braking	Train deceleration	[0.4 : 0.8] (in m^2 /s)	indicationBraking	double
Service (Permitted) Braking	Train deceleration	[0.6 : 1.0] (in m^2 /s)	serviceBraking	double
Warning Braking	Train deceleration	[0.75 :1.25] (in m^2 /s)	warningBraking	double
Emergency Braking	Train deceleration	[0.9 : 1.5] (in m^2 /s)	emergencyBraking	double
Human (Driver) Reaction Time	Reaction time includes the onboard translation of received MA into speed indication, i.e. visualization by the EVC. That is besides the time for the driver to interpret and react to the indication.	[4 : 12] (in seconds)	driverReactionTime	double

<u>Other Input Parameters</u>

According to the layout of the Melton RIDC network and the reference scenario simulated in the context of WP5, the analysis considers the set of train service input parameters as reported in Table 6.

Table 6. Train service input parameters.

Parameter	*Description*	*Range/Value*	*Model Name*	*Type*
Train scheduled headway	time between two successive trains	150 (in seconds)	TrainScheduling (150 s)	short
Number of trains	Number of trains moving	20	nTrains	short

	on the line			
Train length	Length of the train for the specific scenario	120 (in meters)	trainLength	int
Line length	Length of the track for the specific scenario	24000 (in meters)	lineLength	int
Maximum Train Speed	Maximum speed	80 (in m/s)	maxTrainSpeed	short
Maximum Train Acceleration	Maximum acceleration	0.5 (in m^2 /s)	maxTrainAcceleration	double
Step movement	Duration of a step	0.15 (in seconds)	positionUpdatePeriod	double
Safety Margin	safe braking distance	50 (in meters)	safetyMargin	Short
Integrity not confirmed (probability)	Probability that the train integrity is not conformed by TIMS	[0,1)	integrityNotConfirmedProb	Double
Message not delivered (probability)	Probability that a message (MA or TPR) is lost by the communication network	[0,1)	netNotDeliveryProb	Double

5.3.5 Metrics and Simulation Results

The sensitivity analysis aims at evaluating the fluctuations in parameters of the SAN model on the behaviour of the train with respect to brake events. In particular a One-at-a-Time (OAT) sensitivity analysis is conducted which fixes the value of all the considered parameter but on which is instead varied across its domain. For each analysed combination of input parameters, a total of 1000 simulation runs, are performed to make results independent from stochastic variations in the train behaviour. This means that for a given input parameter combination the corresponding value computed for a specific measure of performance is obtained as the average over all the 1000 simulation runs executed for that parameter combination.

The sensitivity analysis is performed on a case study consisting of a double-track railway line having a total length of 24 Km with a maximum speed of 80m/s (i.e. 288 km/h). A total of 20 trains are analysed which operate along one single direction of movement, having a minimum technical running time of 300 s (24.000 m / 80 m/s) from origin to destination. Trains run at a scheduled headway of 150 s, which results in a distancing of 12 km (80 m/s ·150 s). When running in undisturbed nominal conditions all trains (20 trains) are able to reach their destination at 3150 s (300 s + 19 · 150 s). This means that the nominal timetable cycle is set to 3150 s

The conducted analysis evaluates the following metrics:

- **#Trains**: number of trains exiting the line in the time interval [0 : 3200] s;
- ***N****overspeed*: number of times that trains exceed the speed imposed by a specific ETCS braking curve, namely "indication", "permitted", "warning" and "emergency" (as explained in Section 5.3.1) within the total simulation time [0 : 3200 s]. Such a metric is specifically computed as the sum across all trains of the times they overshoot the speed limit of a given ETCS braking curve and then averaged over the total number of simulation runs performed for a single parameter configuration.

The baseline operational scenario defined in this analysis has the following characteristics:

- maximum decelerations for indication, service, warning and emergency braking (respectively, 0.8, 1.0, 1.25 and 1.5 m/s^2);
- minimum TPRUpdatePeriod (1 s);
- minimum latencies for delivering and processing messages, that are RBCProcessingTime, TPRnetDelay, MAnetDelay, EoAUpdateTime (respectively 0.5, 0.5, 0.5, 1.5 s);
- minimum driverReactionTime (4 s).

The baseline scenario refers to nominal undisturbed traffic, where trains run in free-flow conditions without any track occupation conflicts with other trains. In this baseline scenario, train speeds never overshoot the ETCS speed limit relative to the "indication" braking curve and each train exits the line at the scheduled time, as represented in Figure 13, which plots the number of circulating trains versus the simulated time.

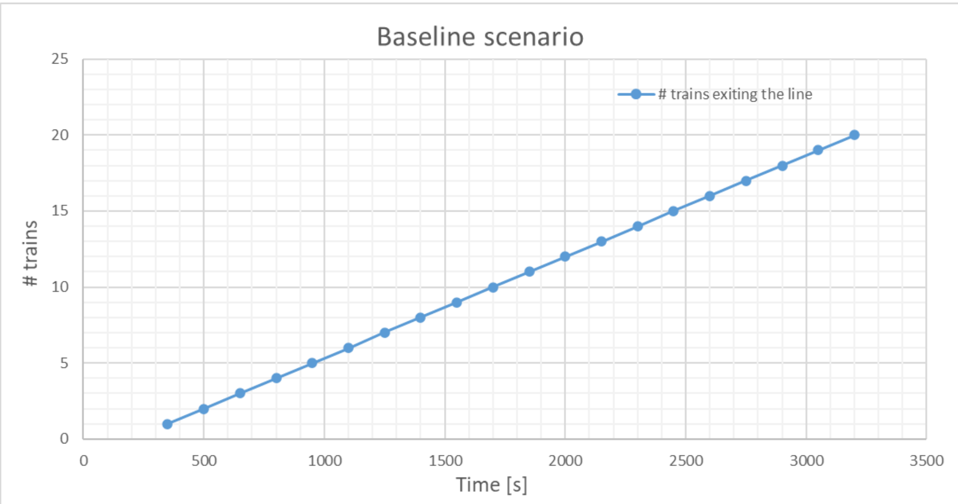

Figure 13. No. Trains versus simulated time for Baseline scenario

Starting from this baseline scenario, we varied the described parameters within the ranges reported in Table 7:

Table 7. Parameter variation range used in the SAN analysis

Parameter	Range	Step
RBCProcessingTime	[0.5 : 4] s	0.5 s
TPRnetDelay	[0.5 : 2.5] s	0.5 s
MAnetDelay	[0.5 : 2.5] s	0.5 s
DriverReactionTime	[4 : 12] s	0.5 s
TPRUpdatePeriod	[1 : 5] s	0.5 s

The analysis was oriented to measure the effects of these variations on the metrics defined before. Mainly, the results are useful to identify the values acting as breaking points where dynamic supervision of the MB signalling system might lead to potentially unsafe train running. The results are described in the following paragraphs.

RBCProcessingTime variation. The impact of the variation of the RBCProcessingTime on the system in the considered scenario is significant. The graph illustrated in Figure 14 shows that if the value of this parameter is higher than 0.5 s, then the number of trains able to exit the line in the considered time interval dramatically decreases. In the considered time interval, a mean of only 2.27 trains exit the line with RBCProcessingTime = 1 s, while an average of 3.25 trains exit for a RBCProcessingTime = 1.5 s.

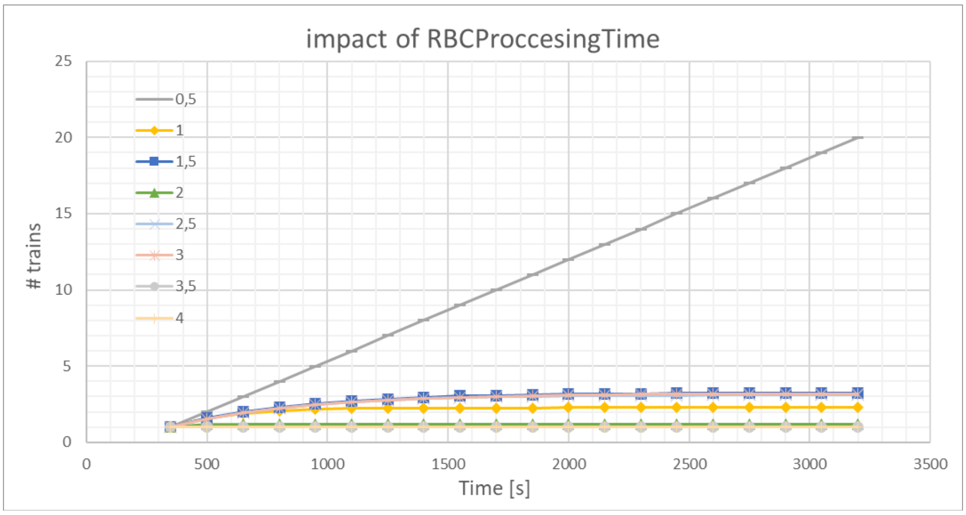

Figure 14. No. of trains versus simulated time for different values of the RBC processing time

The considered effects on the train management capacity are a consequence of the number of times that train speeds in the simulation exceed the thresholds imposed by the different ETCS braking curves, with RBCProcessingTime > 0.5 s. Table 8, specifically reports the values that the indicator $N_{overspeed}$ assumes for the different configuration of the RBC processing time. The results illustrated in the table hence provide the number of times across all the 20 trains that train speeds overshoot the speed limit imposed by the corresponding ETCS braking curves averaged over the number of simulation runs for a specific RBCProcessingTime. For instance, when the

RBCProcessingTime is set to a value of 1s (see second column, fifth row), the total number of times that trains overshoot the speed indicated by the "Indication" Braking curve is 5621.15 on average across all the simulation runs. This number is intended to refer to all the 20 trains across the entire simulation period, meaning an average of 281 times that a single train overshoots the speed limit imposed by the "Indication" Braking curve. It is important to highlight that the number of activations of the emergency braking is very low in all the scenarios compared to the other types of braking.

Table 8. Number of ETCS speed limit overshooting for different values of RBC processing time

$N_{overspeed}$	RBCProcessingTime							
	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
Emergency Braking	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.01
Warning Braking	0.00	17.75	17.07	19.71	17.23	19.11	19.90	17.12
Service Braking	0.00	47.89	44.94	51.14	54.64	60.45	83.27	85.02
Indication Braking	0.00	5621.15	5145.64	6283.85	5202.43	5156.44	6343.70	6287.71

TPRnetDelay variation. Analogously to the previous case, also the variation of the TPRnetDelay has a great impact on the system performance and safety, as it can be seen in Figure 15. With a TPRnetDelay = 1s, the mean number of trains able to exit the line reduces from 20 to only 3.01, while it further drops below 2 for a value of TPRnetDelay > 1 s.

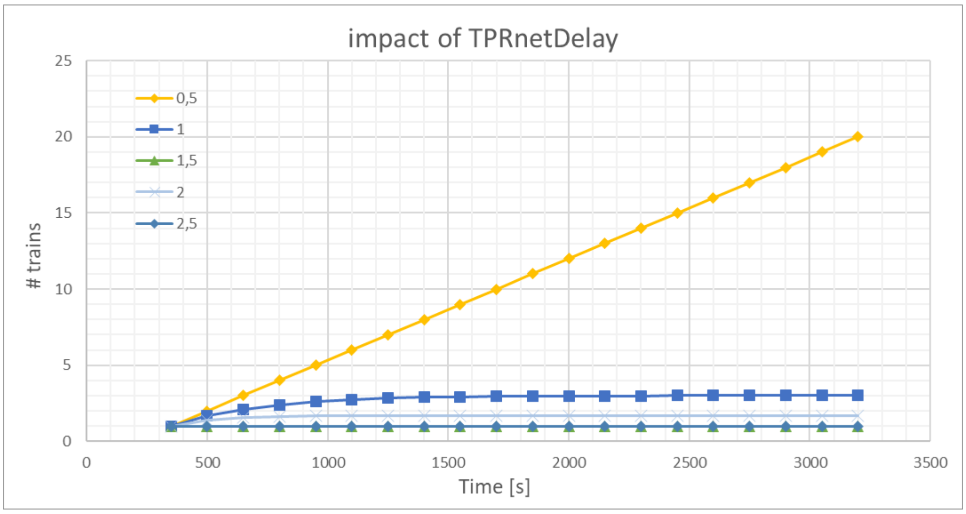

Figure 15. No. of trains versus simulated time for different values of TRPnetDelay

The overall simulated number of speed overshooting of the thresholds imposed by each of the ETCS braking curves (i.e. $N_{overspeed}$) is reported in Table 9. As previously explained the $N_{overspeed}$ indicator reports the number of times across the 20 trains that train speeds overshoot the speed limit imposed by the corresponding ETCS braking curves, averaged over the total number of simulation runs performed for a given value of TPRnetDelay. Despite the large number of times that simulated train speeds overshoot the limit imposed by the "indication" braking curve, the activation of an emergency braking is a rare event under the considered hypothesis.

Table 9. Number of ETCS speed limit overshooting for different values of TPRnetDelay

$N_{overspeed}$	TPRnetDelay				
	0.5	1.0	1.5	2.0	2.5
Emergency Braking	0.00	0.00	0.00	0.00	0.00
Warning Braking	0.00	16.99	19.00	18.32	19.00
Service Braking	0.00	45.24	51.00	48.99	51.00
Indication Braking	0.00	5208.36	6408.00	5970.79	6408.00

MAnetDelay variation. The variation on MAnetDelay in the interval [0.5 : 2.5] s has no significant impact in terms of number of ETCS speed limit overshooting. The results obtained by varying the MAnetDelay has shown the same results as in the baseline scenario, i.e. 20 trains are able to exit the line without any substantial overshooting of the speed limit imposed by the Indication braking curve. Such a result might also be in part due to the level of abstraction considered in the SAN model, which does not represent into detail the ETCS parameter relative to the communication timer (e.g., T_NVCONTACT). However, the model ensures that trains can safely run, with a known EoA, even if the MA is received with a delay of 2.5 s.

DriverReactionTime variation. The increment of the DriverReactionTime has no impact in the baseline scenario. In fact, since no braking is appreciable in the baseline scenario and no driver intervention is required, a delay in her/his reaction would not impact the safe running of trains. This parameter, instead, has a great impact when the driver is required to brake, resulting in a higher probability of an enforced braking application by the EVC when the value increases.

TPRUpdatePeriod variation. At last, the variation of the TPRUpdatePeriod does not introduce perturbation in the system. It means that in the considered case study it is possible to increase this value up to 5s, without appreciating a significant number of infringements of the ETCS speed limits, thereby not substantially affecting the safe train movement.

Discussion. With the considered system configuration parameters, a dramatic impact on performance and safety is obtained when either the RBC processing time is higher than 0.5 s, or the TPR delay is larger than 0.5 s. A delay in the delivery of the MA which is below 2.5 s still allows an effective supervision of the train speed. Similarly, an increment in the period between two subsequent TPR messages up to 5.0 s has no significant impact on the capability of the MB signalling system to effectively supervise safe train movements, even if that would result in a reduction of the number of messages exchanged through the communication network.

Of course, an increase in the headway or a reduction in the maximum running speed, allows tolerating higher TPR delivery delays and RBC processing times. Results of the reported analysis shall indeed be considered valid to the context of the chosen case study as well as to the defined range of input parameters. As described in more detail in deliverable D2.1, the SAN model represents a means to assess a given set of parameters and evaluate the mutual impact in specific scenarios.
The behaviour of the MB represented by the SAN model discussed in this section is a high-level

description of the system, it provides indications about the impact that some variations of the considered parameters may have. Hence, what has been observed in this study depends on the specific implementation of the model and the set of used inputs, the observed results should be considered valid for this specific case and not used to draw a general conclusion on the analysed signalling system.

5.4 Early-Warning Prediction Framework of Hazardous Moving-Block Conditions

In Deliverable D1.1, the moving-block hazards identified in X2Rail-3 are complemented with GNSS hazards. The moving-block hazards mostly refer to errors in train localisation and erroneously clearance of track. The GNSS hazards relate to conditions in which the GNSS cannot perform as desired. In the following, two approaches for providing early-warnings of hazardous moving-block events are presented; one based on short-term prediction and one based on medium-term prediction.

5.4.1 Short-Term Hazard Prediction

The short-term prediction of moving-block hazards focuses on real-time values of parameters, checking whether or not they violate the safety thresholds identified in Section 5.3. In case one or more parameter values are predicted to violate the safety thresholds, a warning indicating the parameter and possible impact should be communicated to traffic management.

Figure 16 illustrates the proposed framework for the short-term hazard prediction. The threshold values for safe MB operations resulting from the SAN analysis are compared against operational data dynamically simulated by the GNSS and the BraSS simulators, respectively. Specifically, the short-term hazard prediction model will trigger an early-warning message any time simulated values of GNSS-, GSM-R- and/or train-related parameters violate corresponding safety-critical thresholds identified by the SAN analysis. For instance, if the RBC processing time simulated by the BraSS platform assumes values exceeding the safety-critical threshold of 0.5 s (which has been identified by the SAN analysis for the parameter RBCProcessingTime), the short-term hazard prediction will issue an early-warning message which is readily communicated to the dispatcher. Such a warning notification contains information about the potential origin of the hazard (in this case a larger RBC processing time), as well as the ID of trains or routes which can be potentially affected in order to help the dispatcher to become alert and get ready to promptly intervene (by e.g. stopping affected trains or blocking dangerous routes), so to prevent the raising of hazardous MB traffic conditions.

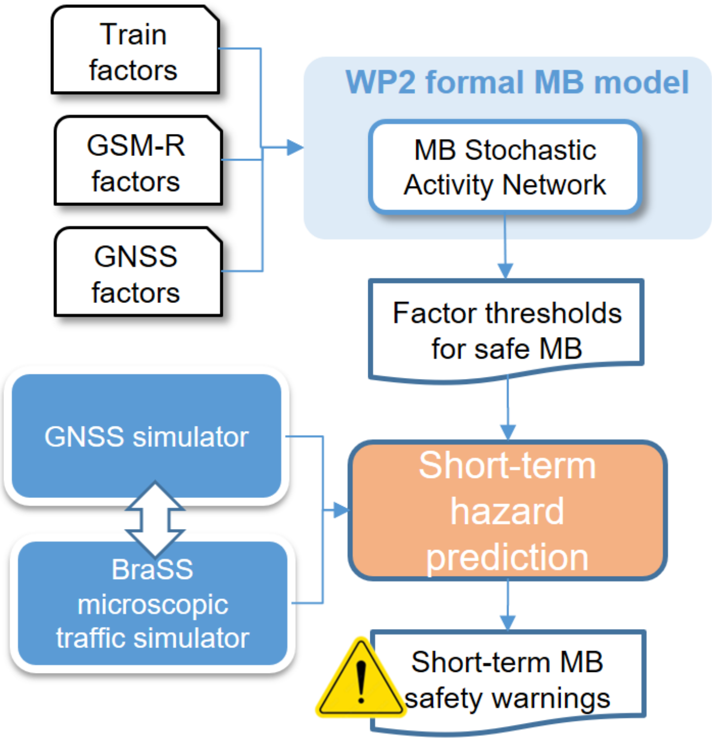

Figure 16. Framework for short-term hazard prediction: violation of threshold values for safe MB operations.

5.4.2 Medium-Term Hazard Prediction

The medium-term prediction of moving-block hazards consists in checking whether track occupation conflicts are predicted to arise in areas where either GNSS visibility or GSM-R communication coverage is limited or compromised. These areas are geographically identified based on static GSM-R and/or GNSS-related data indicating limited / compromised availability of satellite and radio communication signals due for instance to narrow canyons, metal bridges or even GNSS spoofing. Those areas can be considered as 'potentially dangerous zones' as the limited availability of GNSS or GSM-R coverage might increase the likelihood of safety-critical train movements, especially when trains run at a relatively short headway, due to scheduling requirements or disturbed/delayed traffic conditions. The raising of potentially dangerous MB traffic conditions can hence be anticipated in the medium-term when track occupation conflicts are detected in a potentially dangerous zone with limited GNSS or GSM-R signal availability. A medium-term hazard prediction model can be therefore setup as illustrated in Figure 17. As it can be seen, geographical data regarding GNSS and GSM-R coverage / availability is combined with information about predicted location of track occupation conflicts. An early-warning message is then triggered by the medium-term prediction model any time a track occupation conflict is detected within a zone with limited / compromised GNSS or GSM-R availability. Such early-warning message provides the dispatcher with details about the cause of the warning (e.g. limited GNSS visibility) as well as the ID of the trains and/or routes which might be affected and/or involved in a potentially hazardous MB traffic condition. The dispatcher can hence be preventively alerted so to promptly intervene to avoid the raising of MB safety hazards.

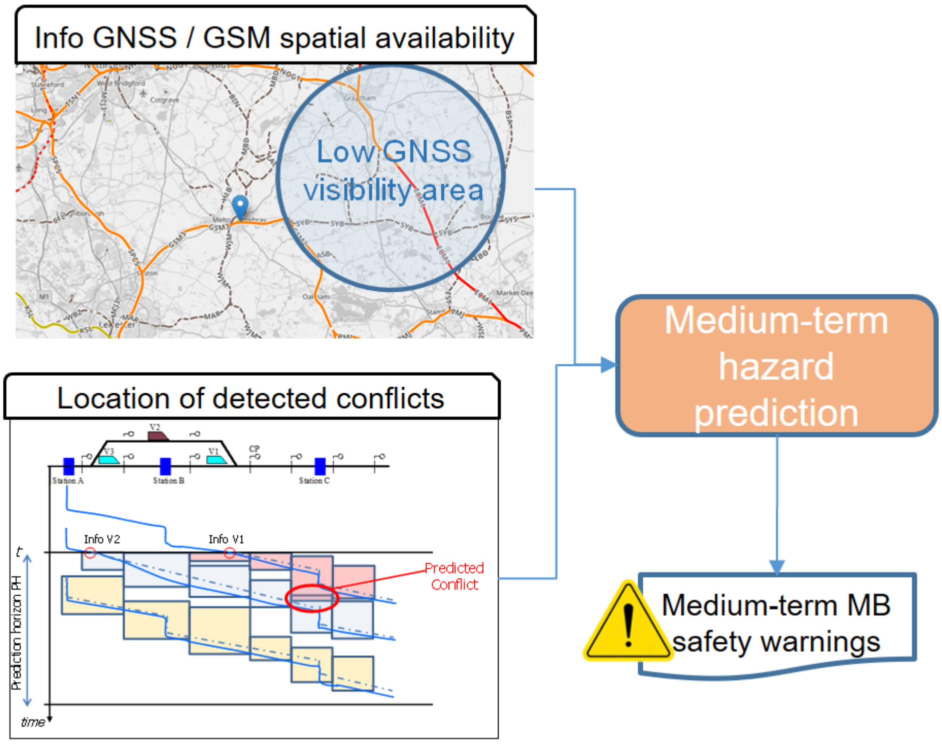

Figure 17

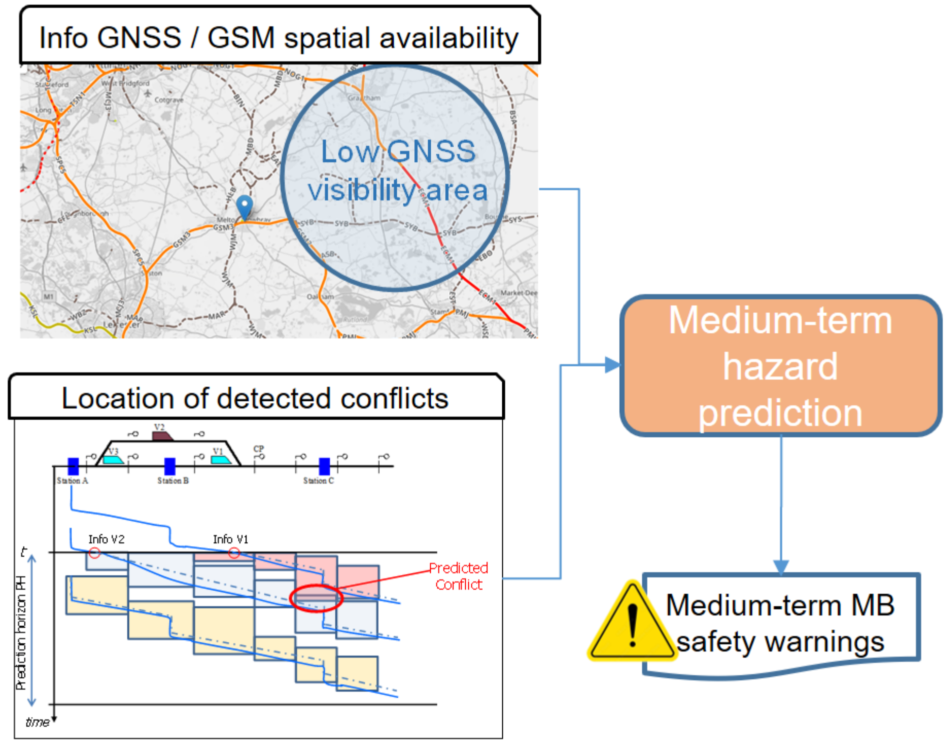

Figure 17. Framework for medium-term hazard prediction: conflicts detected in GNSS or GSM-R critical areas.

6 Conclusions

The work described in this deliverable has the objective of contributing to bridge literature gaps in existing real-time rail traffic management decision support tools with respect to reliable and effective modelling of MB train operations which align to up-to-date signalling requirements. An extensive literature review has revealed that most of the existing tools for optimised real-time rail traffic management refer to fixed-block multi-aspect signalling where headway is considered as a function of the number of block sections (and the number of signal aspects) and not as a function of speed as it is instead required for MB operations. Also, the type of track discretisation used in those approaches does not align with the concept of moving block where fixed-blocks are only found in interlocking areas (to ensure safety around switches) but no longer present on open lines. Only a few works (e.g. Mera et al. (2016) and Janssens (2022)) are proposed for rail traffic management under distance-to-go or MB signalling, however deeper investigation is still required to verify the alignment with the latest requirements defined by the railway industry.

In order to contribute to fill those literature gaps, a model for an optimised real-time rail traffic management under MB is mathematically formulated by building on the latest research advances as well as on up-to-date MB signalling requirements and specifications. In particular, the mathematical formulation of the state-of-the-art tool RECIFE-MILP is here enhanced to align with MB train operations. The proposed RECIFE-MILP formulation includes a set of constraints which translate MB signalling specifications provided in WP2 as well as GNSS-based train localisation and integrity characteristics delineated in WP3. The infrastructure representation previously based on multi-aspect block sections is adjusted to a finer discretisation on open lines which is more suitable to describe MB operations. At the same time, two speed levels are added to the RECIFE-MILP model, namely a maximum and a scheduled speed so that MB speed-dependent headways can be computed for both on-time trains (running at scheduled speed) and delayed ones (using instead the max allowed speed).

In addition, a method is defined for early-warning prediction of potentially hazardous MB traffic conditions. The proposed method can be seen as a non-vital function of the TMS which warns dispatchers about possible critical MB operational conditions which might compromise safety and overall service availability (e.g. punctuality, regularity). Those early-warnings are useful to identify and implement preventive measures to avoid the raising of safety-critical MB traffic conditions in either the short- or the medium-term. In the short-term warnings are triggered anytime real-life and/or simulated railway operations violate safety-critical thresholds identified via a SAN analysis for train operation features (e.g. driving reaction times), signalling- (e.g. MA communication delay) as well as GNSS-related (e.g. train position error) design variables. In the medium-term hazard warnings are instead provided whenever the real-time traffic rescheduling tool detects track occupation conflicts in areas with limited GNSS and/or GSM-R availability, such as tunnels or narrow canyons.

Future activities will be addressed to implement the proposed RECIFE-MILP enhanced formulation for MB and the early-warning MB hazard methods. Both the enhance RECIFE-MILP model and the early-warning prediction method will be interfaced with simulated traffic operations in the UoB's BraSS platform and tested on a real case study. Also, modelling specifications reported in this deliverable will be used together with the outcomes from the simulation-based model testing to delineate a set of guidelines to support the railway industry in defining requirements of an integrated and Advanced Traffic Management Architecture for MB.

References

Shift2Rail JU. (2020). *Shift2Rail Multi-Annual Action Plan.* Luxembourg: Publication Office of the European Union.

Büker, T. G., 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)*, (pp. 219-240).

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

D'Ariano, A., Corman, F., Pacciarelli, D., & Pranzo, M. (2008). Reordering and local rerouting. *Transportation Science, 42*(4), 405-419.

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

Deutsche Bahn. (2017). *Digital Schiene*. Retrieved from https://digitale-schiene-deutschland.de/en

ERA. (2020). *Technical Report: Introduction to ETCS braking curves.*

ERA-UNISIG-EEIG ERTMS Users Group. (2016). *ERTMS/ETCS System Requirements Specification (Subset-026).* European Union Agency for Railways.

European Commission. (2011). *White paper on Transport: Roadmap to a single European transport area – towards a competitive and resource-efficient transport system.* Luxembourg: Publication Office of the European Union.

European Commission-FP7. (2014). *ONTIME project*. Retrieved from https://cordis.europa.eu/project/id/285243

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)*, (pp. 482-486).

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.

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

Jansses, M. L. (2022). *Multi machine approaches for conflict resolution under moving block signalling.* The Netherlands: Delft University of Technology.

Liu, F., Xun, J., Dong, H., Zhang, Z., & Li, Y. (2021). A bi-level real-time rescheduling approach for train operation in high-speed railways. *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, (pp. 2417-2422).

Liu, R. (2016). Simulation model of speed control for the moving-block systems under ERTMS Level 3. *2016 IEEE International Conference on Intelligent Rail Transportation (ICIRT)*, (pp. 322-327).

Lövétei, I. F., Kövári, 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., G, L., & Corman, F. (2018). Integration of real-time traffic management and train control for rail networks - part 1: Optimization problems and solution approaches. *Transportation Reserach 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.

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

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, J., & 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, 47*(5), 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.

Network Rail. (2015). *Digital Railway*. Retrieved from https://www.networkrail.co.uk/running-the-railway/railway-upgrade-plan/digital-railway/

Pachl, J. (2008). Timetable design principles. In I. Hansen, & J. Pachl (Eds.), *Railway Timetable & Traffic* (pp. 9-42). Hamburg, Germany: Eurailpress - DVV Rail Media.

Pellegrini, P., Marlière, G., & Rodriguez, J. (2014). *Transportation Research Part B: Methodological, 59*, 58-80.

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. (2019). Efficient train re-routing and rescheduling: Valid inequalities and reformulation of RECIFE-MILP. *Transportation Research Part B: Methodological, 120*, 33-48.

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.

Quaglietta, E., Pellegrini, P., Goverde, R., Albrecht, T., B., J., Rodriguez, J., . . . Carcasole, D. (2016). The ON-TIME real-time railway traffic management framework: a proof-of-concept using a scalable standardize data communication architecture. *Transportation Research Part C: Emerging Technologies, Vol. 63*, pp. 23-50.

Reynolds, E., & Maher, S. J. (2015). A multi-criteria decision support methodology for real-time train scheduling. *Journal of Rail Transport Planning & Management, 5*(3), 146-162.

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

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

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.

Sanders, W., & Meyer, J. (2001). *Stochastic activity networks: formal definitions and concepts.* Heidelberg: Springer Berlin.

Schlechte, T., and Bornörfer, R., Denissen, J., Heller, S., Klug, T., Küpper, M., . . . Steadman, W. (2022). Timetable Optimization for a Moving Block System. *Journal of Rail Transport Planning & Management, 22*, 100315.

Shift2Rail JU. (2020). Retrieved from https://projects.shift2rail.org/

Theeg, G., & Vlasenko, S. (2009). *Railway Signalling and Interlocking International Compendium.* Eurail Press.

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. (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., & Savelsbergh, M. W. (2000). Time-indexed formulations for machine scheduling problems: Column generation. *INFORMS Journal on Computing, 12*(2), 111-124.

Wang, P., & Goverde, R. M. (2017). Multi-train trajectory optimization for energy efficiency and delay recovery on single-track railway lines. *Transportation Research Part B: Methodological, 105*, 340-361.

Wang, Y., De Schutter, B., Van Den Boom, T. J., & Ning, B. (2014). Optimal trajectory planning for trains under fixed and moving signalling systems using mixed integer linear programming. *Control Engineering Practice, 22*(1), 44-56.

Wang, Z., Quaglietta, E., Bartholomeus, M., & Goverde, R. (2022). Assessment of architectures for Automatic Train Operation driving functions. *Journal of Rail Transport Planning & Management, Vol. 24*.

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)*, (pp. 931-935).

Xu, P., Corman, F., Peng, Q., & Luan, X. (2017). A train rescheduling 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*.

Zwaneveld, P. J., Kroon, L. G., & Van Hoesel, S. P. (2001). Routing trains through a railway station based on a node packing model. *European Journal of Operational Research, 128*(1), 14-33.