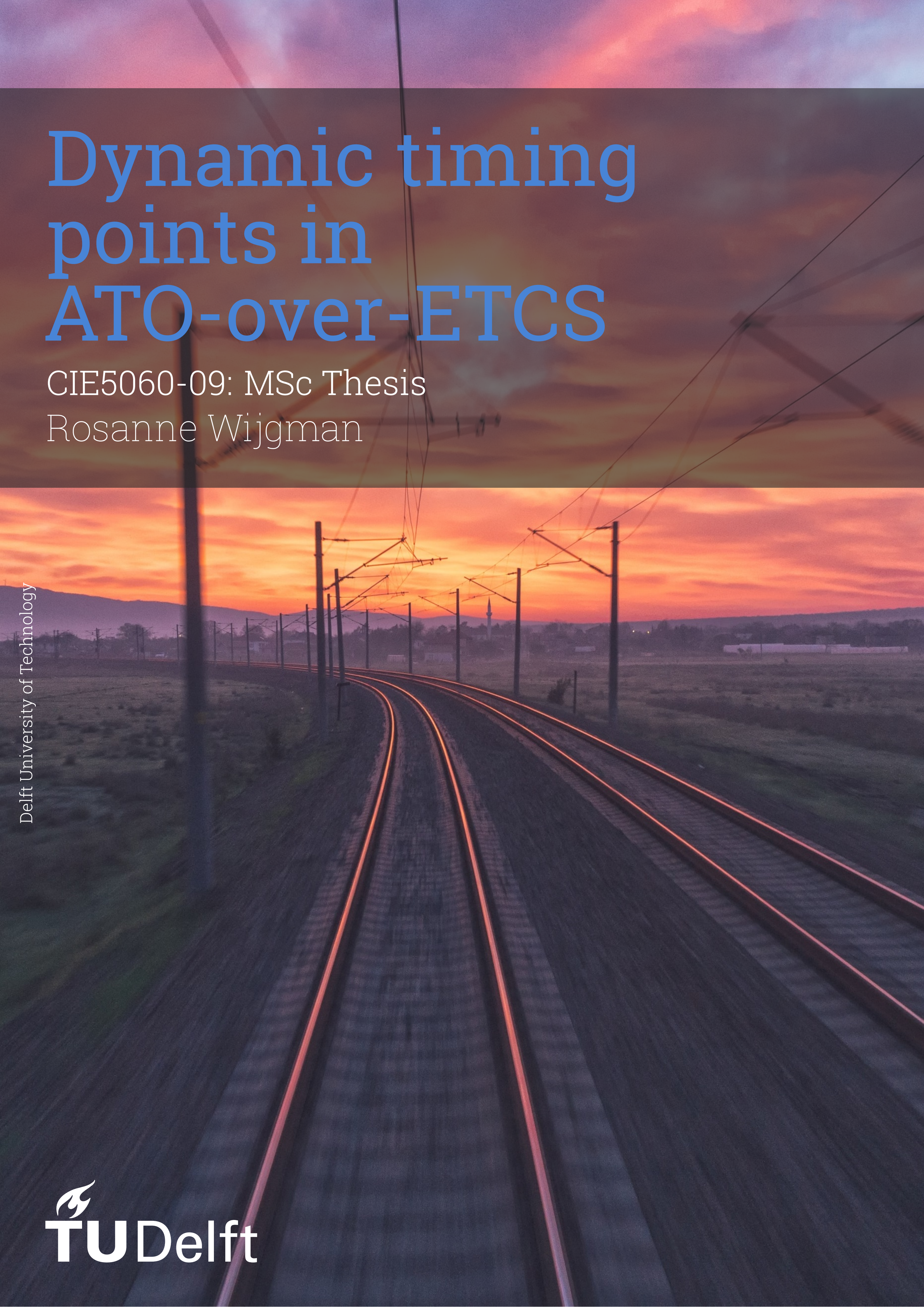


Dynamic timing points in ATO-over-ETCS

CIE5060-09: MSc Thesis

Rosanne Wijgman

Delft University of Technology

Dynamic timing points in ATO-over-ETCS

by

Rosanne Wijgman

to obtain the degree of Master of Science in Civil Engineering
at the Delft University of Technology,
to be defended publicly on Thursday 17 July 2025, 2025 at 14:30.

Student number:	4601939	
Project duration:	February 18, 2024 – July 17, 2025	
Thesis committee:	Prof. dr. R.M.P. Goverde,	TU Delft, supervisor
	Dr. ir. A. Hegyi	TU Delft
	Ir. Z. Wang	TU Delft

TU Delft

Preface

This thesis marks the culmination of my master's of Transport and Planning journey and reflects both my academic interest in and personal fascination with the operational intricacies of railway systems. Specifically, this work explores the concept of dynamic timing points to manage disruptions within railway networks which is increasingly relevant as modern railways seek more adaptive and efficient operational frameworks.

The process of developing this thesis has been intellectually stimulating and, at times, deeply challenging because I choose to combine studying with a full-time career in railway consulting. Throughout, I have benefitted immensely from the guidance, insight and encouragement of those around me.

First and foremost, I would like to express my heartfelt gratitude to Professor Rob Goverde, my main academic supervisor. His expert knowledge, strategic questioning, and unwavering support have been pivotal in shaping the research direction and ensuring academic rigour. His positivity and kind constructive criticism contributed significantly to my enjoyment of the process.

I am equally indebted to my daily supervisor, Dr. Ziyulong Wang, for his attentive guidance, constructive feedback, and for being a constant sounding board during every phase of this work. His commitment and insight in the topic made a big impact on both the technical depth and clarity of this thesis. Ziyulong really went above and beyond as a daily supervisor, always willing to help and being remarkably adept in bringing positive notes to the hours of coding and writing.

My sincere thanks also go to Dr. Andreas Hegyi, whose perspective from outside the railway domain offered a refreshing and invaluable lens. His insightful questions during the progress meetings, helping me to create clarity and sharp focus.

Finally, I am profoundly grateful to my family and friends for their unwavering support, patience, and encouragement. Thank you Estelle for being an incredible support and guiding light in the art of railway theses and the Pareto principle. Thank you dad for believing in me out loud. Thank you Vera and David for listening as I learned to explain my research in layman's terms. Thank you Femke, Evi and Sharon for reminding me to take breaks and enjoy the process.

I hope that this thesis contributes meaningfully to ongoing discussions on dynamic control in railway operations and provides a stepping stone for both the academic and operational world.

Rosanne Wijgman
Delft, July 2025

Summary

This thesis explores the potential of dynamic timing points as a method to enhance punctuality and energy efficiency in mainline railway operations under Automatic Train Operation over the European Train Control System (ATO-over-ETCS). The work begins by situating the study within the structure of European railway systems, where operations are typically divided between infrastructure managers (IM) and railway undertakings (RU). The focus is placed squarely on the operational level, where trains are controlled in real time, and where the interplay between traffic management systems (TMS) and Automatic Train Operation (ATO) becomes most critical.

This thesis focuses on mainline rail networks with ETCS, especially Level 2, which enables continuous train monitoring and removes reliance on physical signals. It also introduces the concept of ATO and its various Grades of Automation, with particular attention to Grade of Operation 2, where train driving itself is automated but the train driver remains responsible for the driving process (including starting the automated driving algorithm) and its safety. The thesis outlines the existing structure of TMS, noting its largely reactive nature and the current reliance on manual interventions for conflict detection and resolution. Although technological advances such as Driver Advisory Systems (DAS) and Connected DAS are increasingly adopted in mainline railway networks, these systems often fall short in managing small-scale disturbances without the computational cost of full rescheduling.

The research problem is defined by the observation that small disturbances, such as brief delays at intermediate stations, can disrupt traffic flow by forcing faster following trains to brake unnecessarily. The TMS is responsible for updates to the Real-Time Traffic Plan (RTTP), but these updates are not typically triggered by minor deviations because of the resource intensity of the calculations. As a result, the ATO algorithm on the train continues to follow the trajectory corresponding to the RTTP which now includes conflict(s), leading to inefficient energy use. The thesis proposes the concept of dynamic timing points (flexibly inserted timing constraints that are calculated during operations) in the context of ATO-over-ETCS as a lightweight alternative to full timetable recalculations.

To ground the research, a comprehensive literature review is presented. It surveys advances in real-time railway traffic management, highlighting existing algorithmic frameworks such as ROMA and RECIFE for conflict detection and resolution. It also examines the field of energy-efficient train operations, with particular attention to eco-driving strategies and train trajectory optimisation using methods such as the Pontryagin Maximum Principle, mixed-integer linear programming, and pseudospectral optimal control. The review identifies a clear research gap: while trajectory optimisation and fixed timing points are well studied, the dynamic insertion of timing points in response to real-time disturbances remains underexplored.

The problem studied in this thesis is the response to disturbances under ATO-over-ETCS by adding dynamic timing points to avoid braking and re-acceleration of the following train. The impact of this approach depends on whether there is enough warning time and warning distance (time and distance between the delay occurring and the time resp. location that

delay information is received by the following train). Warning time and distance enables the following train to reduce its speed to avoid the conflict with the delayed train. Dynamic timing points can also be used to regulate speeding up to prevent conflicts during re-acceleration after forced braking.

The methodological core of the thesis is developed around the use of blocking time theory to detect emerging conflicts between a delayed train and a faster following train. When a conflict is anticipated, the proposed system calculates the optimal location for a timing point at the beginning of the block where the greatest overlap occurs and assigns it a permissible time window. This dynamic timing point is then added to the Train Path Envelope (TPE) of the following train, enabling the onboard ATO subsystem (ATO-OB) to re-optimise its speed profile in a way that avoids unnecessary braking. The process is iterative and may include multiple timing points, depending on the extent of the conflict. The trackside intelligence responsible for generating these updates is envisioned to reside within the ATO Trackside subsystem (ATO-TS), rather than the TMS, to minimise communication latency and computational burden.

To evaluate the proposed methodology, the thesis applies it to a case study of the Utrecht–'s Hertogenbosch corridor in the Netherlands, one of the busiest mainline railway routes in the country. This corridor features both Sprinter (regional) and Intercity (express) services and a mix of four-track and double-track sections, offering a realistic setting for modelling interactions between trains of varying speed profiles and stopping patterns. Various delay scenarios are modelled, introducing delays of 1 to 5 minutes of the preceding Sprinter train at different stations. For each scenario, the resulting behaviour of the following Intercity is evaluated under two conditions: an uncontrolled case, where the Intercity train reacts only when forced by the signalling system, and a dynamic timing point case, where the ATO-OB receives a real-time update and adapts the Intercity train trajectory proactively.

The results demonstrate significant operational and environmental benefits. In multiple scenarios, the dynamic timing point approach enables the Intercity train to avoid up to three forced braking regimes, resulting in energy savings of up to 70.9 percentage point compared to the uncontrolled case. In scenarios where the delay is small and communicated early (yielding longer warning times and distances) the energy and punctuality benefits are most pronounced. Even in cases where the Intercity is forced to stop behind the stationary delayed Sprinter, the dynamic timing point strategy facilitates a more efficient re-acceleration process. Additional experiments show that intermediate updates, where delay information is sent in increments as it accumulates, further enhance performance. A sensitivity analysis of the system's processing time reveals that increases in latency can reduce the benefits, underscoring the importance of timely communication between trackside and onboard systems.

In conclusion, the thesis finds that dynamic timing points offer a promising alternative to full RTTP updates for managing small disturbances in high-density, mixed-traffic railway corridors. They enable significant improvements in energy efficiency and punctuality while preserving the modular separation between IM and RU responsibilities. The work paves the way for future research into broader network-level impacts, optimal trajectory re-optimisation strategies, and the integration of this method within environments where human factors come into play (notably Connected DAS and ATO Grade of Automation 2).

Nomenclature

Abbreviation	Definition
ATO	Automatic Train Operations
ATO-OB	ATO-Onboard subsystem
ATO-TS	ATO-Trackside subsystem
ATO-over-ETCS	ATO operations on track with ETCS
DAS	Driver Advisory System
EETC	Energy Efficient Train Control
EoA	End of Authority
ERTMS	European Rail Traffic Management System
ETCS	European Train Control System
ETCS-OB	ETCS-Onboard subsystem
ETCS-TS	ETCS-Trackside subsystem
IM	Infrastructure Manager
JP	Journey Profile
MA	Movement Authority
MTTC	Minimum Time Train Control
RTTP	Real-Time Traffic Plan
RU	Railway Undertaking
SP	Segment Profile
TMS	Traffic Management System
TPE	Train Path Envelope
UNISIG	Union of Signalling Industry

Contents

1
Introduction

This thesis focuses on improving punctuality and energy efficiency of Automatic Train Operation (ATO) in mainline railways under situations with disturbances (i.e., delays without requiring resource modifications as opposed to disruptions that do require this). This chapter starts by introducing the railway system as a whole and then zooms in on railway traffic management and real-time operation. The operational processes during disturbances are characterised, leading to the formulation of challenges. The chapter will furthermore introduce the research objectives and questions, followed by an overview of the structure of the rest of this thesis.

1.1. Railway system context

A railway system refers to the integrated and coordinated infrastructure, rolling stock, operational processes, and supporting elements designed to enable the efficient, safe, and reliable movement of passengers and/or freight over rail tracks. One way to structure the railway system at a high level is to separate it into the physical infrastructure and the operations running on this railway infrastructure.
In European mainline railways, these are organisationally separated into railway Infrastructure Managers (IM) and Railway Undertakings (RU) — a structure known as vertical separation (Ait Ali and Eliasson, 2022). In contrast, urban railway systems such as metros and many mainline railway systems in other regions outside Europe tend to integrate the operation of infrastructure and traffic into a single company.

There are many interactions between infrastructure and operations along all phases of the operational value chain — strategic, tactical and operational:

1. At the **strategic level**, the transport network is designed based on expected transportation demand. The network design forecasts years ahead and includes routes and frequencies adapted to the expected travel patterns and available infrastructure. There are two kinds of networks at the strategic level: the physical network with stations and tracks (infrastructure side) and the line plan with routes, stop patterns and frequencies (operations side). In general, the IM leads the strategic planning process.
2. At the **tactical level**, the timetable is designed by the IM and one or more RU. The timetabling phase begins a few years prior to the implementation of a timetable and assigns capacity in the form of train paths across the infrastructure. A train path is a time-space allocation on the railway infrastructure that ensures that a particular train can operate at the defined time on a specific route in the network. In vertically separated railway systems, the timetabling needs to balance capacity requests from all relevant RU and resolve any conflicts, a role typically performed by the IM.
3. The **operational level** covers the day of operations, including traffic management. There is significant interaction between IM and RU at this stage, especially for disruption management

This thesis focuses on the operational level, specifically centring on the interaction between traffic management and control of the train during operations.

1.2. Railway traffic management

Railway traffic management is the process of monitoring, controlling and preferably optimising the movement of trains across a railway network to ensure safe, efficient and reliable operations.

The railway traffic management functions are performed by the Traffic Management System (TMS). The TMS receives information about the locations of the trains on the railway network and coordinates the train movements through traffic control, which sets safe routes over the railway network and communicates with the train via signals or cab interfaces. The TMS generally does not perform a safety function, which is covered by the signalling system. Advanced traffic management depends on the implementation of advanced signalling systems.
The advanced TMS mechanisms described in this section correspond to European Train Control System (ETCS) Level 2, which is described in subsection 1.3.1 below.

Railway traffic management is generally characterised by the closed-loop framework presented in Figure 1.1.

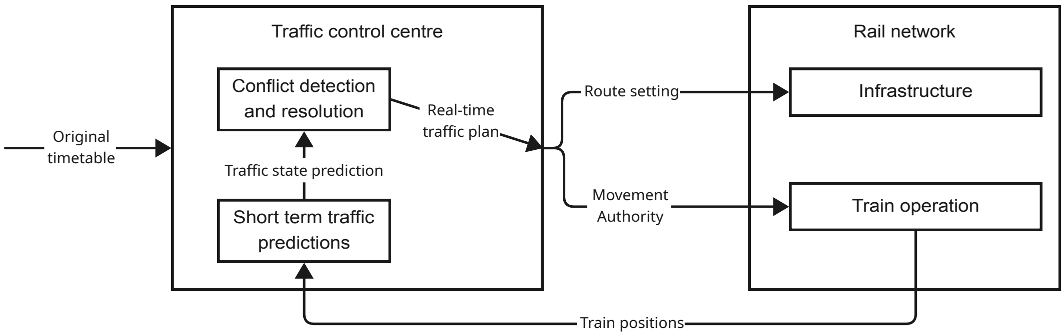


Figure 1.1: Schematic overview of the closed-loop framework for railway traffic management.

The railway traffic management process ideally starts with a conflict-free timetable for the day of operations.
In the Netherlands, the TMS can set routes automatically when all trains run according to the timetable, requiring minimal active management in this state. However, when conflicts arise, the traffic control centre needs to respond and actively control the infrastructure and trains. In most European railways, the automatic route setting is currently lacking completely.

The current railway traffic management is reactive and is not able to optimise the response to deviations from the timetable. Future advanced railway traffic management seeks to optimise its response to such deviations, relying on real-time detailed positioning of trains. This contrasts with the current system, where the TMS can only know which sections of the track are occupied. For an optimised response to deviations, the locations of the front and end of the train are ideally known at all times.
In European railways, the future real-time positioning under ETCS Level 2 (see subsection 1.3.1) is envisioned as continuously communicating the train front positions, supplemented by a train integrity confirmation. Combined with the length of the train, this data provides an accurate up-to-date picture of the full train positioning on the network.

Deviations from the originally planned schedule are detected through train positioning and solved using train management measures such as re-routing, retiming and reordering (Quaglietta et al., 2016). The choice of which measures to apply is either made through predefined scenarios or based on the experience of the dispatcher at the traffic control centre. These dispatchers are responsible for monitoring and updating train paths in response to deviations.
In future advanced TMS, the process of conflict resolution starts with feeding the actively monitored full train positions into a short-term traffic prediction algorithm with automatic conflict detection. Advanced

TMS will then use optimisation algorithms to choose appropriate measures, delivering them either as advice or automatically applied without human intervention.
With these adaptations, the updated timetable becomes the real-time traffic plan, which is used by the advanced TMS to set routes, send departure and arrival time targets to trains, and update passenger information systems.

1.3. Train operation

Train operation is kept safe with signalling systems consisting of components located both in the infrastructure and onboard the train. Advanced signalling systems enable increasing automation of the train driving role, reducing the influence of human performance and therefore increasing reliability.

1.3.1. Safety through signalling

The signalling system is responsible for safety and relies on interlocking to ensure that a train can safely follow its train path. The signalling system provides real-time instructions to the train, regulates train speeds, and prevents collisions by maintaining safe distances between trains. It consists of both trackside and onboard elements. An important subsystem is the Automatic Train Protection that enforces speed limits and prevents unauthorised movements. This ensures that the train does not leave its approved train path and does not exceed the speed limits of the track.

European countries operate with many legacy mainline signalling systems with national Automatic Train Protection systems (for example, ATB-EG and ATB-NG in the Netherlands; PZB and LZB in Germany for conventional and high-speed lines, respectively; KVB and TVM in France for conventional and high-speed lines, respectively; SCMT in Italy; and ASFA in Spain). Legacy signalling systems have various degrees of protection, and among these, high-speed signalling systems such as LZB and TVM are the most advanced.

The European Train Control System (ETCS) is a modern train protection system that is currently being implemented across the continent in an effort to develop a Single European Railway Area. ETCS involves continuous supervision of train movement and has two main variants: Level 1 and Level 2. In ETCS L1 (illustrated in Figure 1.2), there are signals along the track and the communication between the trackside and the train occurs through Eurobalises, which are mounted between the track in order for the train to read them. If a train path is safely set and locked in place, the signalling system communicates the Movement Authority (MA) for this path to the train via a balise. Balises also report their position to the train, which enables the onboard computer to calculate where the train is. Balises in front of the signals report the signal state to the train, thus communicating whether the train will need to brake for the next signal. The onboard computer calculates the braking curve to which the train is permitted to proceed (End of Movement Authority, abbreviated as EoA) and continuously supervises that both this curve and the maximum speed are not exceeded.

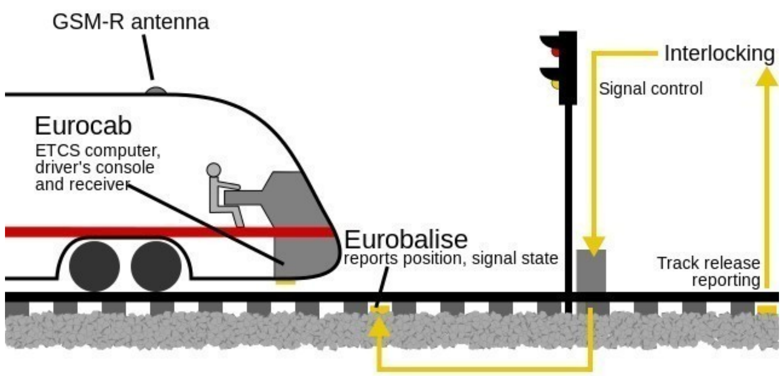


Figure 1.2: Schematic overview of ETCS Level 1 (Mobility and Transport, 2023).

In ETCS Level 2 (see Figure 1.3), there are no lineside signals any more: the MAs are communicated via radio connection, which is currently achieved by the GSM-R system, and is expected to be replaced by Future Railway Mobile Communication System (FRMCS). The onboard ETCS system now contains a Driver Machine Interface that communicates speed limits, MA and other signalling information to the driver. Furthermore, the train has a radio antenna to receive the MA and report its position. The MA is calculated by the Radio Block Centre, which receives input from the interlocking system. There are still balises in the tracks that the train uses to calibrate its position. Track release reporting can be performed by trackside track-clear detection (which is generally through either track circuits or axle counters), but this becomes obsolete when train integrity is supervised by the train.
Hybrid train detection indicates a system that is designed to handle situations where some trains have train integrity monitoring while others use track release reporting via trackside track-clear detection. The remaining trackside track-clear detection can also be used to start back up after a communication system failure.
Onboard ETCS systems like the onboard computer, antennas, train integrity monitoring and Driver Machine Interface are indicated with ETCS Onboard (ETCS-OB) while trackside systems like Radio Block Centres and balises fall under ETCS Trackside (ETCS-TS). The track-clear detection and interlocking are both separate systems interacting with ETCS-TS.

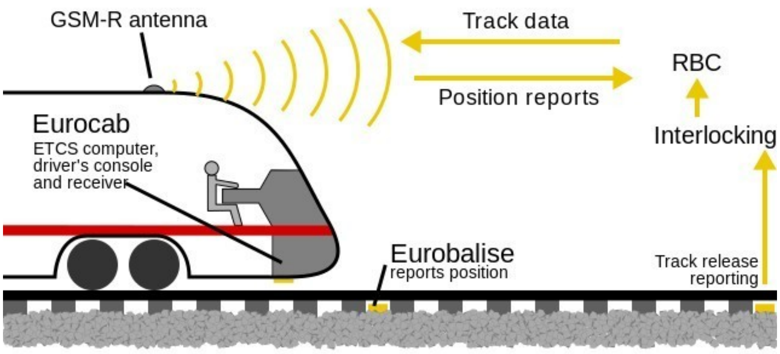


Figure 1.3: Schematic overview of ETCS Level 2 (Mobility and Transport, 2023)

In advanced systems, including ETCS, the signalling system assumes responsibility for safety during operations. This enables the TMS to focus on optimising the traffic flow within the safety boundaries defined by the signalling system.

1.3.2. Driver Advisory Systems and Automatic Train Operation

Traditionally, the movements of trains are controlled by train drivers based on trackside signals. However, as train speeds increased, it became impossible for a train driver to sight signals on time, prompting the development of cab signalling, where the instruction to brake is directly displayed to the driver through the Driver Machine Interface.

The ability of the driver to run the train effectively is limited by the information they have about the preceding trains, which can be improved by Driver Advisory Systems (DAS). DAS provides all the necessary functionality to calculate and deliver drivable advice to the train cab, on the basis of conflict-free time targets which are set outside the system's boundaries and are communicated to the DAS in a synchronous or an asynchronous fashion.
DAS can function by only taking its own operational data into account; this is called stand-alone DAS. A stand-alone DAS system is able to optimise the energy usage of the train while staying as punctual as possible based on the pre-loaded data. The real-time situation is not taken into account.
Connected DAS is a more advanced implementation that calculates the train trajectory and generates advice based on real-time monitored traffic states by the TMS. With Connected DAS, it is possible to adapt the train trajectory dynamically to avoid conflicts with other trains, for example, by slowing down early to avoid a full stop.

One step beyond DAS is ATO, where driving advice is implemented directly by automated traction and braking controls.

The automation of railway operation is classified into different **Grades of Automation (GoA)**, which range from manual operations (GoA 1) to fully autonomous systems (GoA 4):

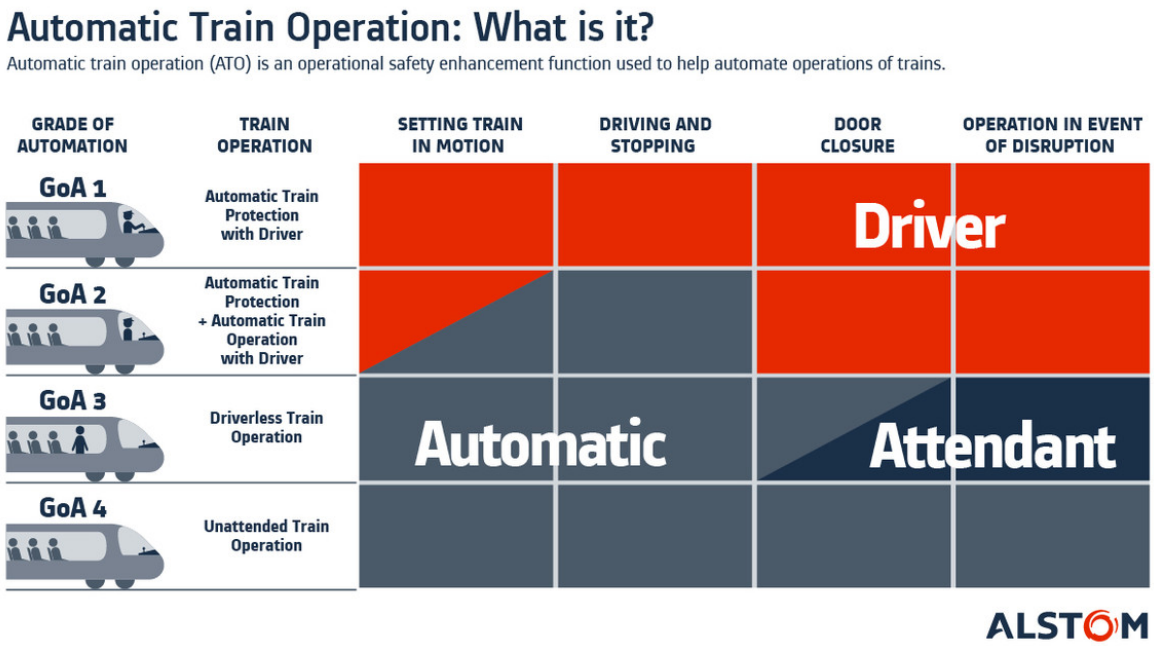


Figure 1.4: Overview of responsibilities in the four Grade of Automation levels for rail operation made by Alstom. DAS systems with Automatic Train Protection fall under ATO Grade of Automation 1.

With ATO Grade of Automation 1, the train driver is responsible for starting the departure, driving the train, closing the doors at stops and operating in the event of emergencies. DAS systems fall under ATO Grade of Automation 1.
The next automation step of ATO Grade of Automation 2 focuses on driving and stopping the train automatically while the driver stays responsible for setting the train in motion, closing doors and operating under emergencies. The train driver does not have a task in driving and stopping the train in normal circumstances.
ATO Grade of Automation 3 can be described as a driverless train operation where the train driver gets replaced by an attendant who is generally not located in the driver's seat but with the passengers on the train. The attendant is responsible for the door closure and gets called upon to drive in emergency circumstances. The ATO is responsible for the full driving procedure, including starting and stopping under normal circumstances.
ATO Grade of Automation 4 is a fully unattended train operation with no driver or attendant on the train.

ATO is already extensively used in metro systems, including fully autonomous Grade of Automation 4 implementations. Examples include lines 1 and 14 of the Paris metro, lines 9 and 10 of the Barcelona metro, the entire Doha metro system in Qatar, the North West line of the Sydney metro and lines 3 and 6 of the metro in Santiago, Chile.
ATO is easier to apply in metro systems due to the controlled, closed environment with no level crossings and the ability to fully control access to the tracks through platform barriers. Furthermore, the traffic is homogenous with standardised rolling stock travelling short distances between stops.
In contrast, on mainline railways, the development of ATO is more challenging due to the absence of this controlled, closed environment, while traffic is inhomogeneous, often with different operators present in the same area Rao et al. (2016).

1.4. Characterisation of operational processes during disturbances under ATO-over-ETCS

Disturbances are characterised by the ability of the IM to handle them without rescheduling crew and rolling stock, and thus generally without involvement of the impacted RU (s). Situations that need rescheduling of resources are called disruptions; these are out of scope in this thesis.

1.4.1. Robust timetabling for mainline railways

The highest utilisation of railway infrastructure can be achieved if all trains on a track have the same direction, stopping pattern and rolling stock characteristics (acceleration, braking, maximum speed). This is often the case for metro systems and dedicated high-speed lines.

Most mainline railway systems, on the other hand, have heterogeneous traffic, including, for example, intercity, regional and freight trains with different stopping patterns and train characteristics. On single-track sections common on regional lines, there are trains from both directions, leading to inefficient capacity utilisation. This thesis, however, focuses on heterogeneous mainline traffic operating on double or multi-track infrastructure. This environment presents operational complexity due to the variability in planned and realised driving. Conceptual and computational advancements in railway operations research allow modelling of this environment, which enables reaching deeper insights. Furthermore, the heterogeneous mainline network with double tracks is common on busy networks in the Netherlands and around the world. In the Netherlands, there are some sections with four tracks to separate fast from slow trains, but the majority of connections on the national core network consist of mixed traffic on double tracks (one track per direction).

When unidirectional trains with different stopping patterns share a single track, there is often an alternating pattern between types of services, for example, a fast intercity train followed by a slower regional service, followed by another intercity, etcetera. This results in regions where the headway between trains is reduced as the faster train “catches up” to the slower train in front of it. Initial delays are called primary delays and can propagate to the following train(s) if there is insufficient buffer planned between the trains; the delay of these following train(s) is called secondary delay. This results in the affected trains needing to brake and therefore generally results in loss of punctuality and/or increased energy use.
It is best practice to add room to the schedule to minimise delay propagation and allow for energy optimisation. This can be done through buffers between trains and running time supplements for a single train run. These running time supplements can be used during operation to reduce the maximum speed and introduce coasting in the absence of disturbances, thus reducing energy consumption. In the Netherlands, the average running time supplement used in practice is around 10% (Scheepmaker et al., 2020).

1.4.2. Real-time traffic management during disturbances

Traffic management is based on translating the timetable into commands for the infrastructure (route setting) and information for the trains with allocated train paths. In real-time traffic management, the timetable is kept up to date with the (near) real-time location of trains, resulting in the Real-Time Traffic Plan (RTTP). This allows for updating the arrival and departure times at further stops, but also connecting or interfering with the train runs of other trains. Rescheduling is based on a predetermined running time calculation between each stop.

An update in the RTTP during operations is communicated to the train through the Train Path Envelope (TPE) which contains timing points as a mechanism to influence the trajectory of the train. Timing points are predefined locations along the route which can have a time window attached in which the train can arrive, depart or pass. The TPE contains the earliest arrival time and permissible time range to each timing point, which can be changed during an update. See also Chapter 2 for more details.

1.4.3. Mainline ATO operations during disturbances

The ATO system is built around the ATO algorithm which calculates traction and braking actions based on the TPE it receives from the TMS. The priorities of a train during operation can be summarised as:

1. Safety
2. Punctuality
3. Energy use

On mainline systems with ATO-over-ETCS, safety is monitored and controlled by ETCS while punctuality and energy use belong in the realm of the ATO system (traditionally: train driver). The main priority of the ATO algorithm is punctuality, with energy consumption being a secondary priority to optimise.

The RTTP ideally eliminates scheduled conflicts, but there can still be additional unforeseen disturbances during operations, e.g., longer dwell time due to unexpectedly large crowd boarding and/or alighting. Due to the prioritisation of punctuality, there are two possible situations when a delayed train causes secondary delays:

1. If the secondary delay of the following train is small enough for it to arrive at the next stop on time => optimise total energy use as much as possible while arriving punctually.
2. If the secondary delay of the following train is big enough to result in a late arrival at the next stop => follow the minimum time trajectory to minimise arrival delay and disregard energy use.

1.4.4. Challenges with static timing points in ATO-over-ETCS with disturbances

The optimal energy efficiency of ATO trains depends on the degree to which the ATO algorithm can optimise its trajectory. Increasing running time supplements in the timetable allows for greater energy savings but may introduce conflicts between consecutive scheduled train runs, thus requiring more buffer time. Imposing additional time constraints can help mitigate such conflicts, however, it also adds complexity to achieving and resolving optimal trajectories. Therefore, the RTTP of energy-efficient ATO-over-ETCS systems will limit the number of timing points.

Scheduled stops are usually seen as necessary timing points called *timetable points* (although it can be argued that the exact timing of arrival and departure at halt stations does not need to be pinned down). Scheduled passage times at, e.g., junctions are also timetable points. These timetable points are timing points with a scheduled time window in which the train should arrive/depart/pass.
Adding timing points at other locations limits the ATO algorithm from freely choosing the most punctual and energy-efficient path and should therefore be avoided.

Operations based on timing points can, however, deteriorate during disturbances, especially if the traffic is heterogeneous.
The updating of timing points is currently envisioned as a result of an RTTP update. However, RTTP updating is done when traffic prediction and/or conflict detection predict a delay or a conflict. The updating process is resource-intensive and needs calculation time, therefore, some operational situations with disturbances would not trigger RTTP updating:

- The RTTP would not be updated for small disturbances that do not affect the RTTP, i.e., with delayed trains still able to arrive on time. This occurs when the delay can be solved en route between stops using the running time supplement and/or buffer.
- It could be acceptable to have delays at short stops on the open track as long as the arrival times at main stations are respected.

In these situations, the trajectory of the initially delayed train can result in an unresolved conflict with the following train. The signalling system will subsequently refuse to extend or revoke part of the existing MA of the following train, forcing it to brake early for this new EoA. When the track is clear again, the following train will accelerate and attempt to catch up to the schedule.
When this situation covers a fast train following a delayed slower train, it can result in multiple conflicts since the following train catches up to the slower train multiple times in a row. Without additional timing points, the faster train cannot be instructed to slow down in between stops because the stop is in itself not conflicted, and the stops are the only timing point locations.

1.5. Research objective and questions

The ATO mechanism is effective in following a conflict-free timetable, optimising its trajectory based on the restrictions of the timetable.
When disturbances occur, ATO keeps attempting to follow the timetable and will only adapt its trajectory when an updated timetable arrives. Timetable updates are computationally costly and take time to construct, leaving space for optimisation of the traffic flow for small delays.
Any timing point that is inserted during operation is indicated with *dynamic timing point* in this thesis.

The primary objective of this work is:

> *Develop and evaluate an approach for dynamic timing points to reduce the impact of disturbances on mainline railways under ATO-over-ETCS.*

The main research question is formulated as:

> What is the potential of dynamic timing points to improve ATO-over-ETCS operations during disturbances?

In order to answer the main research question, the interaction between TMS and ATO during disturbances needs to be designed.

> **Sub question 1:** What interaction mechanism between TMS and ATO enables dynamic timing point updates to improve performance during disturbances?

The next step is to gain understanding of the relevant types of interactions between trains (operational circumstances) in which the proposed dynamic timing point approach can be applied.

> **Sub question 2:** Which operational circumstances potentially benefit from dynamic timing points?

Subsequently, the proposed approach should be tested. In this initial phase, a computer model is most suitable to simulate scenarios and calculate the effect of dynamic updates to the timing points on the performance (including arrival delay and energy usage).

> **Sub question 3:** How can the proposed dynamic timing point approach be modelled and implemented computationally to evaluate its impact on train operation?

Finally, operational improvement can be quantified to identify the potential of the proposed approach.

> **Sub question 4:** To what extent can dynamic timing point updates reduce energy consumption and improve punctuality in ATO-over-ETCS operations?

1.6. Structure of this thesis

The rest of this thesis is structured as follows.

Chapter 2 contains a literature review of the latest advances in railway traffic management and railway operation. It places special emphasis on the link between traffic management and train operation under ATO-over-ETCS. This chapter includes a description of the research gap that was identified.

Chapter 3 develops the proposed approach of dynamic updating of timing points during operation with ATO-over-ETCS and details how this approach can be implemented in a computational model. It starts with describing the problem and assumptions, introduces the concept and details the strategy for calculating and handling dynamic timing points during operations. This chapter focuses on the methodology as it could potentially be applied in real life, including the implementation in the traffic management and train operation systems, describing the information flows that are needed. It furthermore describes how each part of the conceptual approach is translated into a formalised model structure and outlines the algorithms used to perform the calculations.

Chapter 4 focuses on applying the dynamic timing point approach to a case study between Utrecht and Den Bosch. It describes the situation on this corridor regarding infrastructure, train characteristics and traffic patterns. The delay scenarios to be studied are introduced, followed by a description of the modelling methodology and the assessment criteria of the output.
Furthermore, this chapter contains the results of the case study, including the effects of delays in different locations and durations. Additionally, the effect of TPE updates during delay accumulation is studied while a sensitivity analysis is conducted on communication delays by varying the processing time between the delay occurrence reported by the first train and the trajectory adaptation of the second train.

Finally, **Chapter 5** contains the conclusions with answers to the research questions. It describes recommendations for further research based on the limitations of this study.

2

Literature review

This chapter examines the available literature on the role of timing points in ATO-over-ETCS as a method to bridge traffic management and train operation during disturbances. It first describes advances in railway traffic management and train operation, followed by the existing knowledge based on linking traffic management to operation. The latter section briefly introduces the integrated approach common in metro systems and subsequently describes the timing point approach used in ATO-over-ETCS. Finally, the research gap is articulated.

2.1. Advances in railway traffic management

Railway traffic management systems are central to the day-to-day operations of the railway system, ensuring the safe and efficient movement of trains across the network. These systems handle the complexities of route setting, (re)scheduling, and resolving track occupation conflicts if there are any.

Implementing changes in such an essential system is challenging. Despite these challenges, scientific progress has focused on developing train rescheduling as a central component of real-time traffic management (Corman and Meng, 2015; Fang et al., 2015).
Such advancements aim to improve the railway system's performance during disturbances (i.e. relatively minor delays without resource modification) and disruptions (i.e. relatively major delays with resource modification), with the potential to have a significant positive impact on resilience and service reliability. Train rescheduling during disturbances generally covers the time window from real-time (or near real-time, e.g. < 2 minutes) up until 45 to 60 minutes in the future. This is referred to as real-time rescheduling as opposed to delay management and disruption management that cover later time windows and bigger impact (Fang et al., 2015).

2.1.1. Real-time traffic management

Real-time traffic management indicates systems capable of dynamically responding to changes in operational conditions, such as disturbances or disruptions, in (near) real-time. This dynamic response enables train rescheduling measures like train reordering, rerouting, or retiming, which help minimise delays and improve network performance (Fang et al., 2015).

The implementation of real-time traffic management relies on three key components (Tschirner et al., 2014; Corman and Meng, 2015; Quaglietta et al., 2016):

1. Real-time data exchange to accurately monitor train status (i.e., position, time and speed), and to support traffic prediction where applicable
2. Advanced conflict detection and resolution algorithms capable of processing large datasets and generating optimal rescheduling solutions
3. A robust framework with a standardised communication format to integrate these components into the operational production process

For optimal collaboration between IM and RU, it is essential to ensure that data is integrated and the RTTP is shared consistently. This allows the RTTP to be the single source of truth, avoiding contradictory actions due to divergence of interpretations between parties (Tschirner et al., 2014).

Real-time traffic management can be classified as static with a single optimisation based on perfect information or dynamic with repeated optimisation as updated information becomes available (Corman and Meng, 2015). Dynamic real-time traffic management can be reactive or proactive in obtaining information.

The real-time traffic management problem can be solved centrally through a TMS, more recently decentralised approaches using e.g. self-organisation are being explored (D'Amato et al., 2024). One notable effort in the central approach is the European project ON-TIME (Optimal Networks for Train Integration Management across Europe), conducted between 2011 and 2014. ON-TIME developed a framework for automatic real-time traffic management based on the computation and sharing of the RTTP (Quaglietta et al., 2016).
The framework uses a web-based Service-Oriented Architecture with a standardised RailML interface in order to achieve a flexible and scalable system, see Figure 2.1 that gives an overview of the system components.

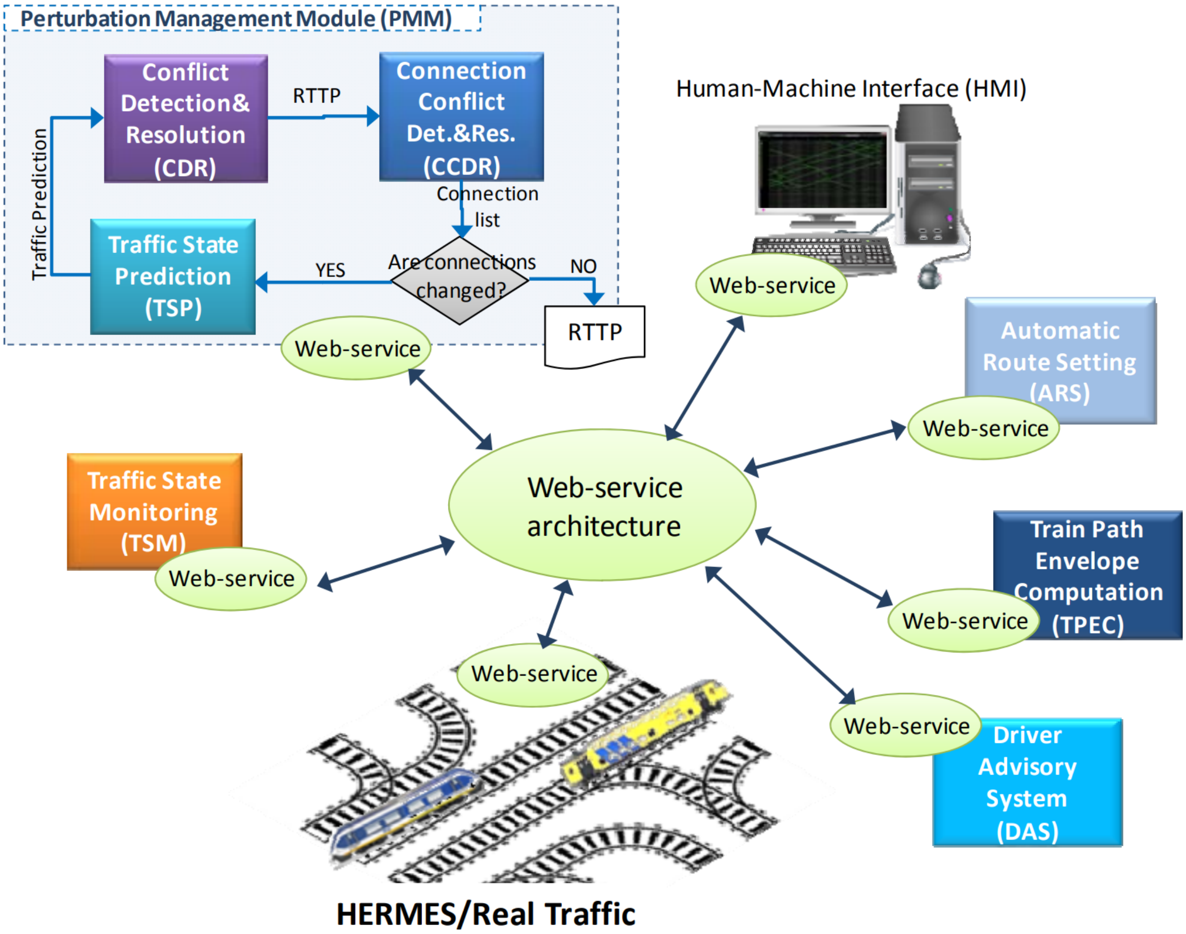


Figure 2.1: Overview of the system architecture of the ON-TIME framework for real-time management of railway traffic perturbations (Quaglietta et al., 2016). HERMES is a simulation environment used in place of real traffic for testing.

The framework was tested using simulations of real-life case studies in the UK, Netherlands and Sweden, showing that it can be applied across diverse national contexts (Quaglietta et al., 2016).

The real-time traffic management process flow can be summarised as follows (see Table 2.1):

Table 2.1: Overview of the process of real-time traffic management flow, including current situation (general situation in Europe, varies between countries). Based on (Quaglietta et al., 2016) and expert insight.

Step	Input	Output	Responsibility	Current situation
Traffic State Monitoring	Train statuses incl. location	Overview of current traffic	IM	Automatic
Traffic State Prediction	Overview of current traffic	Prediction of traffic into the future	IM	Manual or automatic
Conflict Detection	Prediction of future traffic	Prediction of future conflicts	IM	Manual
Conflict Resolution	Current and predicted conflicts	Solution to minimise conflicts	IM (& RU)	Manual
Conflict-free Real-Time Traffic Plan (RTTP)				
Route Setting	RTTP	Routes set in infrastructure	IM	Manual or automatic
Trajectory Calculation	RTTP	Train control	RU	Manual

1. The status of the trains within the sphere of control is continuously monitored by the IM, including localisation and ideally speed. The **Traffic State Monitoring** automatically combines the relevant information into an overview of current traffic. This can be limited to only the track section occupation, but in some systems with advanced signalling like ETCS Level 2, it includes more exact localisation and real-time speed.
2. The **Traffic State Prediction** takes the output from the Traffic State Monitoring and extends this into the future either automatically or manually. The prediction mechanism can be basic with fixed travel times between stops or more advanced with calculations based on actual train parameters, partial sections and even expected traffic control measures.
3. **Detecting conflicts** effectively and efficiently is essential for real-time traffic management. Most IM in Europe require manual work to detect conflicts, either based on arrival/departure time predictions based on the current delay or additionally using a visual representation of the future forecast in time-distance diagrams. This illustration is helpful for dispatchers, but manual detection is still required.
4. **Conflict resolution** is currently manual. Smaller delays (i.e. disturbances) can be handled by the dispatchers of the IM without involvement of the RU.
   Managing disruptions that affect rolling stock and crew resources requires involvement from both the IM and RU. In general, this is handled in the network control centre with significant communication between IM and RU. Disruptions are out of scope for this thesis; refer to (Ghaemi and Goverde, 2015) for more information.

There are various conflict detection and resolution algorithms in the academic literature on railway traffic management (Corman and Meng, 2015; Fang et al., 2015). Two illustrative examples are:

(a) **ROMA** (Railway traffic optimisation by Means of Alternative graphs) developed at Roma Tre University and TU Delft (D'Ariano et al., 2007)
ROMA is a conflict detection and resolution module that uses a truncated Branch and Bound algorithm to identify the optimal train passage orders and shifts in departure/arrival times that minimise the maximum consecutive delay. It then employs a tabu search to identify alternative routes that can further reduce delays.

(b) **RECIFE** (Railway traffic rescheduling integrated with fixed and forecasted delays) (Pellegrini et al., 2015)
RECIFE is a Mixed-Integer Linear Programming formulation that concurrently detects conflicts and solves them to optimality by retiming and reordering trains. It optimises the total delay of all trains at scheduled stops and at the exit of the network.

In practice, it is common to use **heuristic methods for conflict resolution**. This method enables a passenger-centric approach focused on service quality to be integrated into operations management (Corman et al., 2017).

5. Conflict resolution leads to a **conflict-free RTTP** that is shared with all trains in the form of a TPE
6. The IM uses the RTTP to **set routes** for the trains. More advanced systems can automatically set routes when there are no disruptions, for example, ARI or Automatische Rijweg Instelling at ProRail in the Netherlands. Other regions and countries are still dependent on manual route setting for all trains, e.g., much of the German network.
7. Finally, each train (driver or ATO) is individually responsible for driving according to the RTTP to the best of their ability.

Traditionally, traffic management is focused on traffic flow exclusively but recent approaches attempt to integrate the perspective of the traveller using a microscopic passenger-centric model with travel time as a central KPI (Corman et al., 2017).

2.2. Advances in train operation

Relevant innovations in train operation on mainline railway systems include energy-efficient operation based on optimisation of the trajectory and the development of DAS and ATO.

2.2.1. Energy-efficient train trajectory computation

The trajectory of a train is determined by the timetable, track characteristics (including gradients and speed limits) and vehicle-related factors. The equations of motion to calculate the trajectory of a train are generally formulated as a function of the traversed distance, focusing on spatial dynamics rather than temporal progression. This aligns better with infrastructure constraints, is able to efficiently handle resistance forces and is compatible with the available computational optimisation techniques.

$$\begin{aligned} M_{\text{eq}} * \frac{dv(s)}{ds} &= \frac{F_m - (F_r(v) + F_g(s))}{v(s)}, \\ \frac{dt(s)}{ds} &= \frac{1}{v(s)} \end{aligned} \tag{2.1}$$

where M_{eq} is the train mass including the effect of rotary inertia, $v(s)$ the velocity at traversed distance s, F_m the motor traction/braking effort, $F_r(v)$ the running resistance at speed v and $F_g(s)$ the force associated with gravity and track gradients. Track curvature forces are ignored in this thesis since they are small compared to the other resistive forces.

The running resistance is calculated using the Davis Equation:

$$F_r(v) = A + B \cdot v + C \cdot v^2 \tag{2.2}$$

The coefficients A, B and C are part of the train characteristics.

The gradient force can be modelled using:

$$F_g(s) = -mg\,\sin(\theta(s)) \tag{2.3}$$

with θ being the angle of the slope at location s. Usually, $\sin(\theta)$ is approximated to n in m/km (the slope measured as rise in m per 1000 m). This slope is given in the gradient profile.

The train dynamics equations can be used to minimise energy consumption, which is calculated using

$$E = \int_{s_i}^{s_f} F_{\text{m}}\, ds \tag{2.4}$$

with $F_{\text{m}}(s)$ being the traction force of the motor (thus zero during coasting and braking) between stops i and j (s_i and s_j respectively). The energy use of a train run with multiple stops is the sum of the energy use between each pair of consecutive stops. Regenerative braking is not taken into account in this thesis.

The following constraints are applicable:

- **Timetable**
  This is the most important requirement limiting the energy efficiency of train trajectories since customers desire fast travel times. The timetable constraints between any two stops at locations s_i and s_j can be summarised as

$$v(s_i) = 0, \quad t(s_i) = T_{d,i} \quad v(s_j) = 0, \quad t(s_j) = T_{a,j} \tag{2.5}$$

  where $T_{d,i}$ represents the departure time at stop i and $T_{a,j}$ the arrival time at stop j. The difference between the departure and arrival times at the same stop is the dwell time.
  Note that the equation of motion (Equation 2.1) is undefined when $v = 0$ since $1/0 \to \infty$. This can be solved by using a small speed $v(s_i) = \epsilon \neq 0$, for example, 1 m/s, just before and just after all stops. Acceleration from or deceleration to then needs to be correct to 1 m/s. The acceleration/deceleration from 0 to 1 m/s and vice versa can be computed using the dynamic equation with time as the independent variable because this does not need division by zero.
- **Traction and braking power**
  The traction and braking power are limited by the motor of the train as well as the adhesion between the rails and the wheels.

$$\begin{aligned} F_{\text{min}} \leq F_m(s) \leq F_{\text{max}}, \\ 0 \leq F_m(s) \cdot v(s) \leq P_{\text{max}}, \end{aligned} \tag{2.6}$$

  where $F_{\text{min}} \leq 0$ represents the maximum braking force and P_{max} is the maximum traction power. Note that the tractive effort curve of a train consists of two distinct regions. At lower speeds, the tractive effort is limited by adhesion (in the simplified form $F = F_{\text{max}}$) while the tractive effort at higher speeds is constrained by power limits ($F = P_{\text{max}}/v$).
- **Speed limits**
  The train shall not exceed the speed limit at any location along its path:

$$0 \leq v(s) \leq V_{\text{max}}(s), \tag{2.7}$$

  where $V_{\text{max}}(s)$ is the speed limit at position s.

The four typical driving regimes of a train trajectory consist of acceleration, cruising, coasting and braking, with the ideal switching points between these regimes as the prime optimisation goal. This is generally solved through Pontryagin's Principle, see chapter 3 of Su et al. (2023) for more information.

2.2.2. Modelling train-path conflicts with blocking time theory

The trajectory of an individual train is modelled using the dynamic equations of motion, infrastructure characteristics, timetable, motor characteristics and speed limits. When a second following train is added, it is essential to consider the section blocks that the route consists of. Blocking time theory is often visualised in a time-distance graph as shown in Figure 2.2.

The setup time represents the time needed to move and lock the physical infrastructure for the route. For radio-based systems like ETCS L2, the setup time also includes the generation of the MA in the ETCS-TS, the time to communicate the MA to the train, the subsequent computation of the dynamic speed profile by the ETCS-OB and finally the time to display the new MA on the Driver Machine Interface (if train is driven by a train driver). The reaction time represents the time from the appearance of the instruction to the reaction from the ATO system or train driver. The approach time corresponds to the distance needed to brake from the current braking indication point to a standstill, while the running time is the time to traverse the block. The clearing time is the time needed to get the tail of the train out of the block after the nose moves into the next section, and the release time corresponds to the time needed to release the locked infrastructure for the next path to be set.
Each block is occupied for the *occupation time* corresponding to the sum of the running time and clearing time, but blocked for a total of the *blocking time* defined as the sum of all six components.

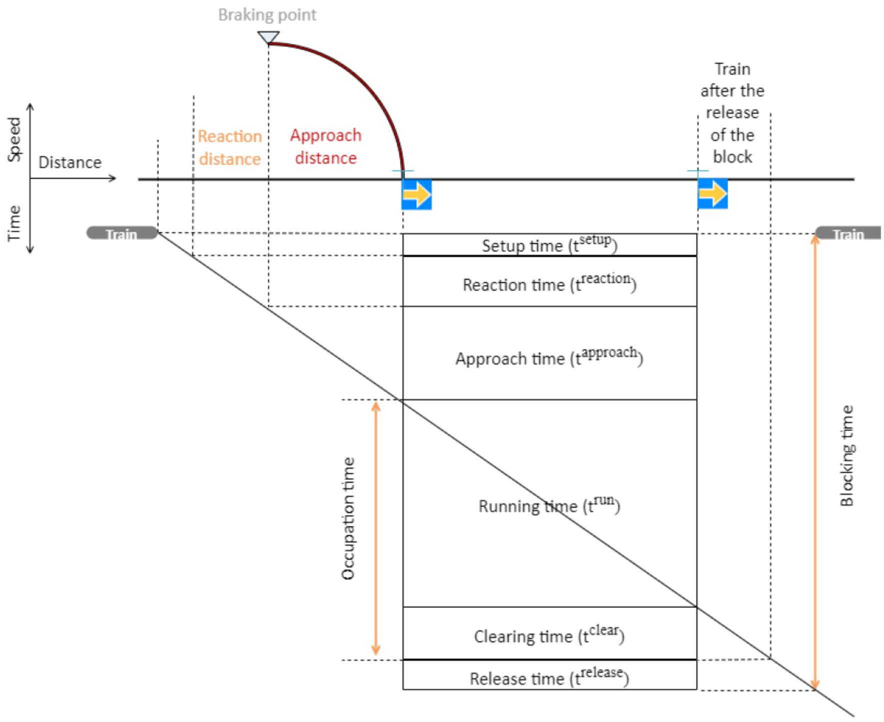


Figure 2.2: Overview of blocking time components under ETCS showing occupation time and blocking time (FP1 MOTIONAL, 2024b).

Note that some blocking time components depend on the speed of the train (approach time, time in section and clearing time) while other components are generally modelled with fixed values in the order of seconds (setup time, reaction time, release time).
The speed-dependent components are calculated by integrating the inverse of the speed over the relevant distance:

$$T = \int_{s_0}^{s_1} \frac{1}{v(s)} ds \tag{2.8}$$

with s_0 and s_1 corresponding to the appropriate start and end for the respective time component (for example, the start and end of the block for the running time in this block).

Using a blocking time model results in a blocking time stairway for each train, see Figure 2.3. The smallest buffer between the two blocking time stairways is called *critical buffer time*. The blocking time of any train cannot overlap with the blocking times of the same block created by other trains.

Blocking times are modelled in this way to ensure accurate modelling of the train trajectory under the influence of the signalling system. During real-life operations, the ETCS system will extend the MA to include a block only after the previous train has left this block. If the following train nears the block directly before the block occupied by the preceding train, the ETCS system will keep the EoA at the start of the occupied block. The ATO-OB of the following train will subsequently guide the train to a stop before the Stop Marker Board indicating the start of the occupied block. Only after the preceding train leaves the block does it become possible to extend the MA for the following train.

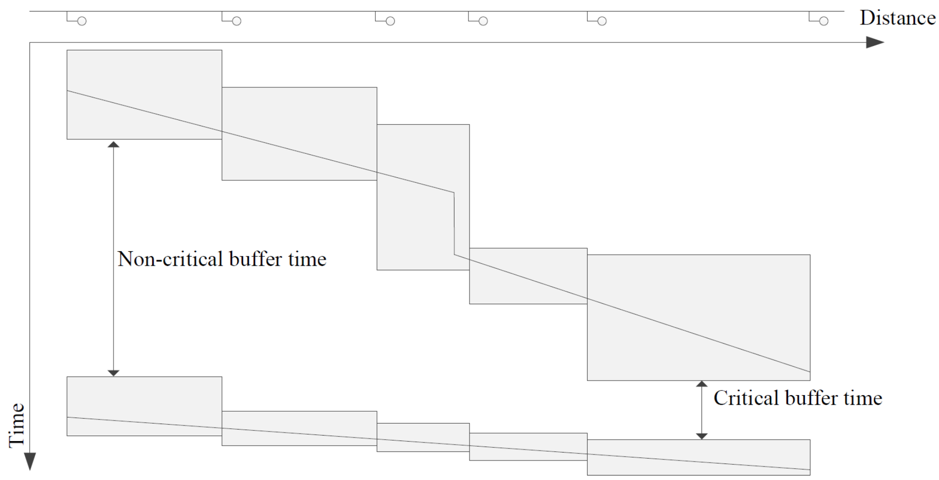


Figure 2.3: Blocking time stairways for a slower train with a stop (above) and a subsequent faster train sharing a track (Jensen, 2015).

2.2.3. Energy-efficient operation

Energy-efficient train operation research mainly focuses on traction energy consumption, which accounts for around 80% of the total energy consumption of railway systems (González-Gil et al., 2014). Relevant references include the book *Energy Efficient Train Operation* by Su et al. (2023) and the literature review by Scheepmaker et al. (2017).
The most important measures to reduce traction energy consumption include regenerative braking, traction efficiency (for example, by reducing train mass and energy losses in the catenary), and energy-efficient driving (Su et al., 2023). The latter is the focus of this study. Within energy-efficient driving, the measures are proactive traffic management, eco-driving techniques, ATO and DAS.

A train run between two stops typically consists of four different driving regimes: acceleration, cruising, coasting and braking (Howlett, 2000).
The **acceleration regime** involves applying maximum traction power to increase the speed of the train until it reaches a desired level. Once this speed is achieved, the train enters the **cruising regime**, during which the speed is maintained by balancing the applied traction effort against resistive forces, such as air resistance and track gradients, with extra traction force for uphill slopes and less traction for downhill slopes. In the **coasting regime**, both traction and braking are disengaged, allowing the train to decelerate gradually due to resistive forces alone. The effectiveness of coasting depends on the remaining time and distance before the next scheduled stop. Accurate timing of the coasting point allows coasting to be used optimally, enabling energy savings without compromising punctuality.
Finally, the **braking regime** brings the train to a controlled stop, typically using service braking to ensure passenger comfort and safety. Note that braking does not use traction power. Some railway systems use regenerative braking, which is not considered in this thesis.

Not all regimes are present in every journey. For shorter distances, such as those in urban transit systems, the cruising and coasting phases may be omitted entirely, with the train transitioning directly from acceleration to braking. The specific combination and duration of these regimes depend on operational priorities, such as maintaining tight schedules or minimising energy consumption.

The specific use of driving regimes towards a defined operational goal is called a *driving strategy*. The driving strategy, which focuses on using as little time as possible, is called Minimum Time Train Control (MTTC) and corresponds to the minimum running time on the specific segment between two stations. The MTTC does not use coasting as this slows down the train earlier than needed and thus extends the travel time.
A *running time supplement* is added on top of the minimum running time to allow for natural variance in operation time (e.g. due to dwell time extension). The running time supplement is typically 5-7% of

the minimum running time and allows for eco-driving if there are no disturbances.

Beyond MTTC there are three typical driving strategies, see Figure 2.4 and overview below (Scheepmaker et al., 2020):

1. **Maximum coasting (MC)** seeks to reduce traction energy by maximising the use of coasting. This strategy accelerates the train to its maximum speed as quickly as possible, then cuts off traction, allowing the train to coast for the remainder of the journey, using the running time supplement in the timetable to arrive as planned.
2. **Reduced maximum speed (RMS)** where the train maintains a lower maximum speed to cruise during the journey in order to arrive exactly on time at the next stop without the coasting regime.
3. **Energy-efficient train control (EETC)** combines optimised cruising speeds with controlled coasting to achieve the lowest possible energy consumption while meeting operational constraints.

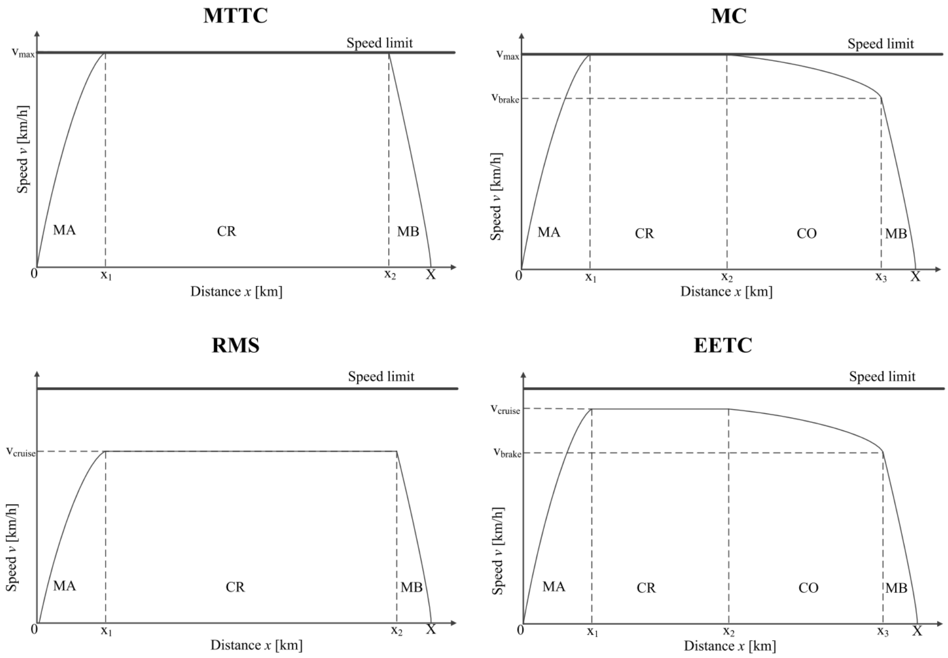


Figure 2.4: Overview of typical driving strategies (Scheepmaker et al., 2020). MA indicates maximum acceleration, CR indicates cruising, CO stands for coasting and MB for maximum braking. Note that the graphs should be seen as independent: each of these speed profiles will have different running times.

The EETC strategy can be described as the theoretical optimum of eco-driving and is the focus of much academic literature on eco-driving. It is based on maximum acceleration to the optimum speed, followed by maximising the coasting stage at the expense of the cruising stage while staying punctual. Both the optimum cruising speed and the point to switch from cruising to coasting are variables to optimise based on the objectives (Scheepmaker et al., 2020). Using a (longer) coasting phase results in a longer running time and therefore depends on using the running time supplement in the timetable. In practice, algorithms search for the optimal coasting point given a cruising speed, and then iterate to find the optimal cruising speed and corresponding coasting point(s) such that the train arrives at the scheduled arrival time.

The energy impact of maximising coasting can be significant: adding 1% time supplement to the minimum running time of a mainline train on a route of approximately 50 km between Arnhem and Nijmegen results in around 20% energy savings (Scheepmaker and Goverde, 2021). Additional running time supplements see diminishing returns in terms of energy savings. Intercity trains (only calling at large stations) benefit more from adding the same nominal running time supplement than regional

trains (stopping at all stations): a 10% running time supplement results in 50.2% decrease in energy consumption for the Arnhem-Nijmegen intercity and 48.9% for the Arnhem-Nijmegen regional train, both compared to the situation with zero running time supplement.
Maximum energy savings are lower for metro rail operation because the distance between stops is much lower, reducing the maximum speed reached and thus the coasting potential. For a stop distance of 1000 metres, the energy savings of optimal coasting are 3.78% (Sopharak et al., 2020).

2.2.4. Train trajectory optimisation methods

The trajectory of a train follows from the control inputs, i.e. throttle and braking force. Optimising the train trajectory is therefore formulated as an optimal control problem.
The Pontryagin Maximum Principle is a cornerstone of optimal control theory, providing necessary conditions that any optimal control must satisfy. Pontryagin's Maximum Principle leads to a set of differential equations for the state and additional co-state variables, with multiple solution approaches available (Yin et al., 2017):

1. **Analytical methods** are able to obtain the theoretically optimal solution but are rigid in the formulation and mainly used with simplified operational circumstances (Albrecht et al., 2016)
2. **Numerical algorithms** including Mixed-Integer Linear Programming, dynamic programming and the pseudospectral method that make a trade-off between computation time and optimisation performance. This approach transforms the optimal control problem into a non-linear programming problem solvable with standard solvers. This solution approach is common in literature, examples include work by Wang et al. (2013), Pellegrini et al. (2015) and Goverde et al. (2021)
3. **Evolutionary algorithms** like genetic algorithms, which are the most flexible, but cannot guarantee that the solution is optimal

Analytical models are quick to solve but often inadequate to capture the complexities of operations. The fast optimisation required in real-time situations is therefore best performed with dedicated numerical algorithms. A typical example of numerical algorithms is the use of the Quadratic Search Method for calculating the coasting point when modelling a single train (Sopharak et al., 2020). Another example is the PseudospectRal Optimal train control MOdel (PROMO) to compute driving strategies (Scheepmaker and Goverde, 2021; Wang et al., 2013; Wang and Goverde, 2016). More recently, the analytical solutions of the train motion model's differential equations resulted in a computationally efficient coasting advice algorithm (Cunillera et al., 2023).

2.2.5. Advances in Automatic Train Operation for mainline railway systems

ATO refers to the system that automates the driving of trains, handling functions such as acceleration, cruising, braking, and stopping with varying levels of human supervision depending on the automation grade.
Most of the development of ATO for mainline railway systems focuses on ATO-over-ETCS and is led by the European Union. The current program is Europe's Rail Joint Undertaking, a continuation of the Shift2Rail Joint Undertaking. Other initiatives include recent efforts to develop ATO for lines with Positive Train Control by the US Federal Railroad Administration (Hunter and Bryant, 2023) and Communication-Based Train Control which describes the well-developed ATO system for metro networks (out of scope for this thesis).

The performance of the ATO algorithm for speed profile computation and traction/brake control is the basis for ATO system performance as a whole. KPIs for evaluating the performance of ATO algorithms include schedule adherence (including departure, passing point and arrival locations), stopping position accuracy, energy consumption and comfort (Bochmann and Jaekel, 2022).

2.2.6. Overview of current ATO-over-ETCS standardisation

The central body in the development of technical specifications for all ERTMS systems is the European Union Agency for Railways (ERA).
Another important player is the Union of Signalling Industry (UNISIG), an industrial consortium that currently consists of nine railway manufacturers (Alstom, AZD, CAF, Hitachi, MERMEC, Progress Rail, Siemens Mobility and Thales, see (UNISIG, 2014)). They assist the Working Groups of the European Railway Agency with Mirror Groups and hold monthly meetings to discuss technical progress.

The Union of Signalling Industry is related to Union des Industries Ferroviaires Européennes (UNIFE), which represents the railway manufacturing industry.

The ERTMS specifications consist of a set of documents called subsets that are developed by UNISIG and provide detailed specifications to support the implementation of the TSI CCS (Technical Specifications for Interoperability for Control-Command and Signalling). Subset-026 is the central ETCS document and contains the System Requirements Specification with the functional and technical requirements of ETCS.
ERTMS is developed in **baselines** which represent a collection of subsets consolidated into a coherent, standardised version of ERTMS. Each baseline specifies which versions of subsets must be used. Within one baseline, there can be multiple **releases** that enhance or update the baseline with, e.g., error corrections or minor functional adjustments.

The recent history of ERTMS specifications includes:

- The European Commission adopted the specifications of **Baseline 3 Release 2**, including SRS 3.6.0, in May 2016. This version provided a mature and stable version of the system, serving as a foundation for many current ERTMS implementations.
- **Baseline 4**, including SRS 4.0.0, was adopted as part of the TSI CCS in March of 2023. This version introduced formal support for ATO, including ATO Grade of Automation 2+, which is the focus of this research.

2.2.7. ATO functional architecture

The ATO system architecture for railway systems consists of the ATO-trackside (ATO-TS) and ATO-onboard (ATO-OB) subsystems. In ATO-over-ETCS, the safety-critical functions are located in the ETCS system: the ATO-OB relies on ETCS to ensure the safety. See Figure 2.5 for an overview of the ATO functions.

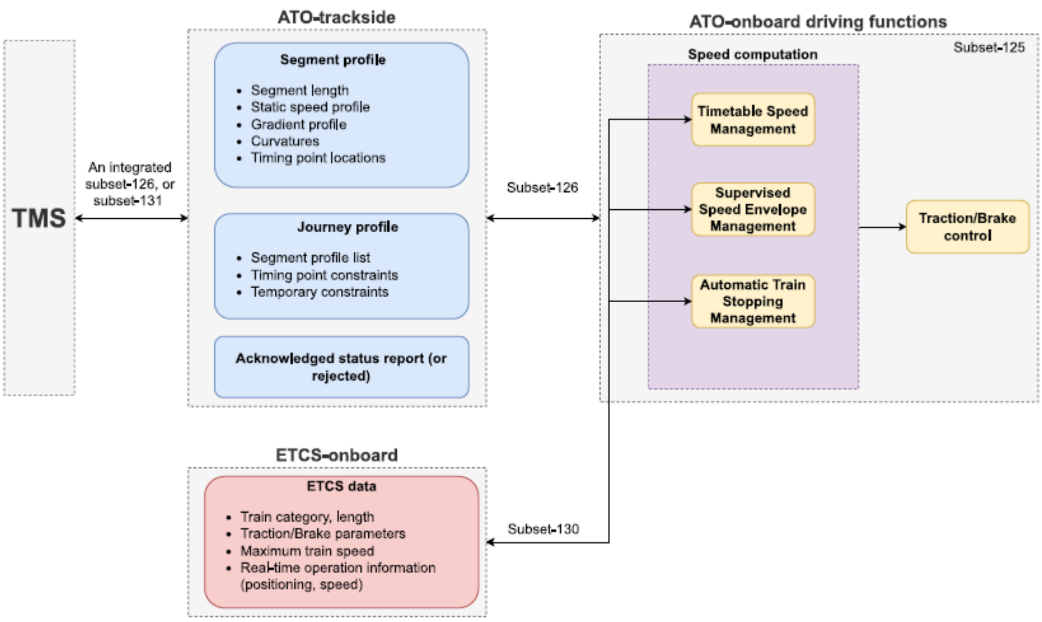


Figure 2.5: Driving functions of ATO-over-ETCS system architecture (Wang et al., 2022).

Both ATO-TS and ATO-OB can be active and passive regarding their computational functions, leading to four different potential architectures (FP1 MOTIONAL, 2024b). The ERTMS/ATO Subsets currently assume an active ATO-OB that computes its train trajectory, leading to a connection between the ATO driving functions and operational process that can be described as:

1. TMS sends the timetable and infrastructure information to the ATO-TS subsystem, most commonly when there is a change in the existing timetable.
2. ATO-TS uses this information to calculate an (updated) Journey Profile (JP) for each train, which includes all information that a train needs to know to calculate its trajectory: an ordered list of the infrastructure segments that will be covered, the timing point constrains and any temporary constraints that are applicable. The JP is subsequently sent to the ATO-OB along with the relevant segment profiles (SP) for the infrastructure information. The SP is a detailed description of a track segment, including length, speed limits, gradients and curvature, to enable accurate trajectory calculations
3. The (active) ATO-OB subsequently manages the train trajectory computation and controls the traction and brakes to follow the calculated train trajectory.

.

The JP that will be exchanged between ATO-TS and ATO-OB according to the specifications contains the Train Path Envelope (TPE): a discrete set of timing points and their associated time windows that serve as constraints for generating each train's conflict-free trajectory. This concept is used extensively in this thesis.
The TPE can either be calculated by the ATO-TS, which is then classified as active, or already in the TMS that sends it over to the passive ATO-TS (FP1 MOTIONAL, 2024a).

The interactions between ATO-OB, ATO-TS, TMS and train are visually represented in Figure 2.6. The interaction with ETCS-TS and ETCS-OB is also included.

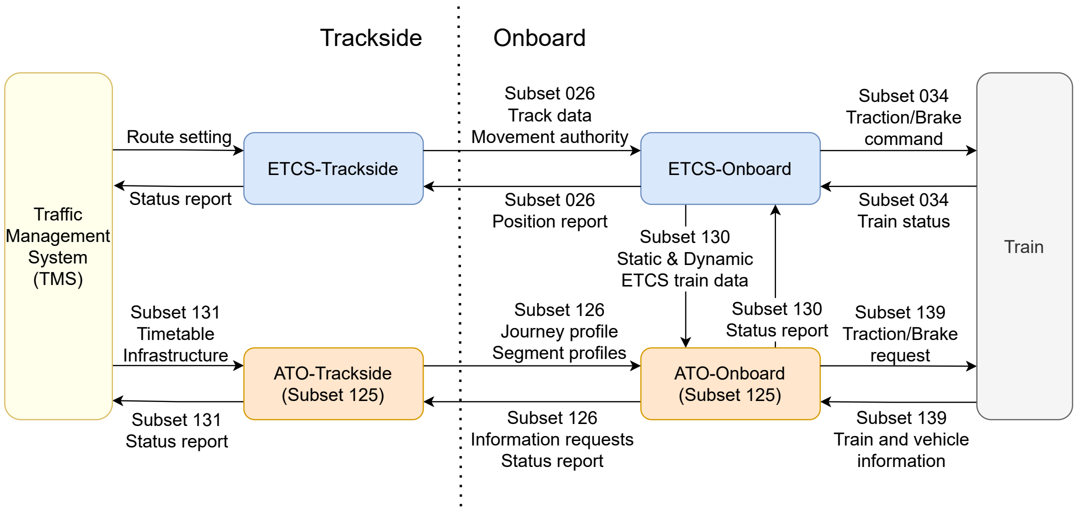


Figure 2.6: ATO-over-ETCS reference architecture (Wang et al., 2022)

The RTTP is calculated by the TMS and used to trigger route setting via ETCS-TS. Under ETCS L2 the ETCS-TS represents the Radio Block Center, which subsequently gives the ETCS-OB movement authority. The timetable and infrastructure information from the TMS are used by the ATO-TS to calculate JP and SP for the trains on the network, providing the context for ATO-OB to calculate and follow its trajectory via traction and brake commands. Position and status reports are used to check the state of the system and trigger a recalculation of the timetable if required.

Each component and interaction between components has an ETCS subset describing its function. The main specifications for ATO are in subset-125 and 126. Subset-131 is replaced by the TMS-CCS (Command Control and Signalling, referring to the systems that supervise, control and authorise train movements) interface developed by the System Pillar in Europe's Rail based on the Reference CCS Architecture. The ERTMS Users Group (EUG) and EULYNX consortium have developed the Reference CCS Architecture to make the Control Command and Signalling systems modular (ERTMS User Group, EULYNX consortium, 2022). A modular CCS system reduces the tendency towards vendor

lock-in and therefore reduces lifecycle costs while facilitating migration from legacy to advanced architectures.
This thesis aims to provide methodologies that can be used for developing the TMS-CCS interface specifications (bottom left of Figure 2.6).

The International Union of Railways (UIC) independently published a specification for the interaction between TMS and a Connected DAS system. This specification is indicated with IRS 90940, with the latest version published as edition 2 in August of 2022.
This specification aligns with Reference CCS Architecture and defines a standard for data exchange named SFERA (Smart communications For Efficient Rail Activities) that includes the content and format of data exchanged between TMS and Connected DAS, but not the implementation and calculation method for driving advice (International Union of Railways (UIC), 2022). IRS 90940 is modelled to be close to the ATO-over-ETCS specifications, providing a superset of Subset-126 that also supports legacy train protection systems. This requires the IRS 90940 standard to be enriched by additional data about the infrastructure, signalling system and train characteristics that would be covered by ETCS in the ATO-over-ETCS specifications.

Despite the lack of specification dedicated to the relation between TMS and ATO-TS, it is clear that timetable-related data needs to be exchanged between these two systems. This could be the RTTP as discussed in subsection 2.1.1, but also the TPE.
The ON-TIME framework developed a standard representation of the RTTP, which includes an infrastructure view and a train view, both indicating which section is occupied by which train starting from which moment in time (Quaglietta et al., 2016). Note that this is potentially restrictive because it leaves little room for trajectory optimisation.
The Reference CCS Architecture uses the term *Operational Plan* to signify the complete set of preplanned and dynamically updated actions to manage railway traffic, including real-time adjustments. It describes operational movements, restrictions, and warning measures for operational events. Operational movement events can be of type *"passage"*, *"stop"* or *"other"* and include both precise reference to position and the timing. The latter consists of at least the earliest arrival and latest arrival values (see section 4.1.1.3.2 of RCA.Doc.31 v1.5). The operational movement event thus corresponds to one or more timing points with *passage* event mapping to timing points as studied in this thesis.

2.3. Linking traffic management to train operation

There are two main approaches for the link between traffic management and train operation: an integrated approach where timetable and driving strategy are jointly optimised, and an approach using timing points to cover the link. This section will briefly introduce the integrated approach for completeness, but does not give a full review because this is outside the scope.

2.3.1. Integrated approach

Integrating traffic management and train operations is possible for vertically integrated systems where joint optimisation of timetable and driving strategy is an option. This is commonly the case in metro systems where the infrastructure owner and traffic manager are the same as the operator.
The power of integrating traffic management and train operations in metro systems was shown by Su et al. (2019) in their study on energy optimisation in multiple metro trains. This study contains a detailed train operation model including both energy optimisation of the train driving strategy as well as a substation-based energy consumption model taking into account regenerative energy. Using Dynamic Programming and Simulated Annealing to optimise energy consumption and Approximate Dynamic Programming to solve timetable optimisation resulted in a net energy reduction of up to 2.8% on the Beijing Yizhuang metro line.

A pivotal work for joint optimisation of timetable and driving strategy in mainline railways is the work by Rao et al. (2016). This paper focuses on feedback loops between train and trackside and discusses how an integrated system can improve train service quality with real-time automatic solutions for solving conflicts through reordering, rerouting or retiming as well as avoiding unplanned stops through Connected DAS and ATO. Unplanned stops are avoided using the concept of a *main-target point* with a target position, speed and time that is calculated according to the conflict type behind the unplanned stop (either a crossing conflict or a follow-up conflict). Sub-target points are introduced when necessary to achieve the main-target point and evaluated on the performance of the resulting trajectories in terms of capacity and energy saving.

2.3.2. Timing point approach

In vertically separated railway systems like European mainline rail systems, it is difficult to integrate real-time timetabling with train operation because these functions are organisationally separated. This is often complicated further by the existence of multiple RU as opposed to only one in a vertically integrated railway system. The decisions about rescheduling, therefore, necessitate communication between organisations.
The RU are responsible for the direct control of the train. The IM retains a form of control over the traffic flow in vertically separated systems through the issuing of timing points. A timing point serves as a control location in the network where the train is planned to arrive, depart or pass at a particular time. In a traditional timetable, this would be the stop location at station tracks, but also locations like stop marker boards before switches can be timing points in order to manage network bottlenecks. In addition to timing points at predefined locations, there is no functional limit to inserting timing points at any location along the track. However, according to the ERTMS specifications, the timing point has to be defined in the SP (section 6.4.3.3 of Subset 125 by ERA and UNISIG and EEIG ERTMS Users Group (2023)).

In ATO-over-ETCS, the timetable for a train is communicated from the wayside to the train using the JP. The JP contains the train route data (ordered list referring to the SP of segments that the train will drive on), temporary constraints (e.g. speed restrictions and low adhesion zones) and operational data in the form of a list of timing points (section 6.4.2.1 of Subset 125 by ERA and UNISIG and EEIG ERTMS Users Group (2023)). These timing points include the following information:

- Type of timing point: departure, passing or arrival
- Position of the timing point
- Departure time or Arrival time and tolerance

The TPE contains the set of timing points received by the train which is subsequently used by the ATO-OB to calculate the speed profile.

2.3.3. Trajectory optimisation with intermediate timing point(s)

Intermediate timing points are timing points that are not located at stops. Literature on energy-efficient driving using intermediate timing points in mainline railway systems is limited.

Early work on energy optimisation with intermediate timing points was done for driver advisory systems for freight rail in Australia (Pudney et al., 2011). This study splits the trajectory in two at the timing point location and calculates feasible hold speeds on both sides with constraints set by the arrival time and timing point. The distance for which the cruising speed was kept before the timing point was varied in an attempt to find the trajectory with minimal energy use. This study uses a search algorithm as an optimisation approach.
Recent continuation of this work uses an analytical approach to optimise the energy consumption of a fleet of trains travelling in the same direction on the same track (Howlett et al., 2023). This approach uses a clearance-time constraint formulation as opposed to blocking times in order to simplify the optimisation while ensuring safe separation.

Actual optimisation with a form of intermediate timing points is found in research by Wang and Goverde (2016). This study splits the train trajectory into segments (called ”phases”) delimited by signals in order to study the effect of a green wave policy on small disturbances. Each signal, therefore, becomes equivalent to a timing point with a target time window. The resulting multi-phase optimal control model is based on minimising energy use and is completed with linkage conditions. It uses the pseudospectral method to solve for minimising the sum of the cost functions within each phase, subject to the constraints within each phase and the linkage constraints. This research uses the same case study as this thesis with a similar timetable. The green wave trajectory is calculated by predicting the signal

aspects based on full knowledge of the prior delayed train, which requires further development of TMS. The following research of this team includes single-train and multi-train trajectory optimisation formulated through the multiple-phase optimal control problem and solved by the pseudospectral method Wang and Goverde (2019). A case study on the Dutch rail network showed that energy savings of 7-24% could be obtained.
Further work on the numerical approach by Haahr et al. (2017) used event-based decomposition of the space-speed graph to drastically reduce the search space while incorporating speed limits, gradients and timing points. This reduces the optimal control problem to the shortest path problem with time windows. The study uses a Label Setting Algorithm to find the optimal solution for the shortest path problem, which is accurate but not fast.

Fernandez-Rodriguez et al. (2020) used a Differential Evolution algorithm to generate possible solutions, which are subsequently evaluated using a detailed train simulation model to find the feasible speed profile with minimum energy consumption. This was studied in situations with one and two intermediate target times. The Differential Evolution algorithm is able to optimise the trajectory as a whole as opposed to optimising between timing points, followed by optimising the connections. Applying the Differential Evolution algorithm to a case study on the Spanish high-speed rail line between Calatayud to Zaragoza showed that there is a 7.7% difference between the most and least energy-efficient trajectories for one intermediate timing point and a 3.1% difference for two intermediate timing points. Increasing the number of timing points decreases the range of feasible trajectories and thus lowers the difference between the most and least energy-efficient. The results show that optimising the energy usage of the full trajectory is worthwhile, but the paper mentions that the execution time of the algorithm is around 6 minutes. It is therefore not suitable for real-time applications.

Another example of the evolutionary algorithm approach applied a particle swarm searching algorithm to trajectory optimisation with one fixed timing point in the middle of two stations (Fernandez-Rodriguez et al., 2021). This research focuses on the impact of varying the timing window at the intermediate location. Both the minimum and maximum feasible passing time result in increased energy use (respectively 5.8% and 25.3% with respect to the trajectory without timing point constraint).

Zhou et al. (2023) developed an approach for real-time optimisation of the trajectory with intermediate timing points based on decomposition to a train-based variant of the shortest path problem with time windows. This method uses a tour-adaptive partial-bounding pulse algorithm to reduce the computational time to below one second.

The ability of TPE to tolerate schedule deviations due to different driving styles was studied by Wang et al. (2023). A recent study of the same group at TU Delft shows that adding a timing point can be an effective strategy to remove conflicts due to different driving strategies (Wang et al., 2025a). In this study, a critical-block strategy was used to determine the location and time window of the timing point needed to resolve the conflict.

2.4. Research gap

While significant progress has been made in real-time traffic management, energy-efficient operations, and ATO development, the dynamic determination of timing points in ATO-over-ETCS remains underexplored. Existing literature focuses on static or pre-defined timing points, leaving a gap in methods for real-time, dynamically adaptive approaches.

This thesis aims to address this gap by developing and testing concepts for dynamic timing points, building on the reviewed existing studies.

Paper	Context	Timing point characteristics	Model approach	Optimisation algorithm	Real time
Pudney et al. (2011)	Freight	1 intermediate time target, fixed location	Pontryagin Maximum Principle	Search algorithm	No
Wang and Goverde (2016, 2019)	Conventional	All signals as timing point	Multi-phase	Pseudospectral	No
Haahr et al. (2017)	Conventional	Up to 10 timing points	Shortest path problem with time windows	Label Setting Algorithm	No
Fernandez-Rodriguez et al. (2020)	High speed	1 and 2 intermediate time target, fixed location	Integrated optimization	Differential Evolution	No
Fernandez-Rodriguez et al. (2021)	High speed	1 intermediate time target, fixed location	Particle swarm optimization	Eco-driving particle swarm searching algorithm	No
Howlett et al. (2023)	Conventional metropolitan	At selected signals	Clearance time constraints	Unconstrained convex optimisation	No
Zhou et al. (2023)	High speed and conventional	Up to three time windows, fixed location	Train-based shortest path problem with time windows	Tour-adaptive partial-bounding pulse algorithm	Yes
Wang et al. (2023), Wang et al. (2025a)	Conventional	Intermediate TPs at critical signals	Train Path Slot optimisation	Numerical based on Karush-Kuhn-Tucker conditions	Yes

Table 2.2: Overview of literature on intermediate timing points in mainline railway systems

3

Approach for traffic harmonisation with dynamic timing points

This chapter presents the proposed methodology for dynamically updating timing points. It begins by describing the scope and underlying assumptions of the approach, followed by a step-by-step explanation of dynamic timing points, demonstrated based on an illustrative use case. It furthermore details the operational procedures needed for the suggested approach. The chapter includes a translation of this approach into a computational model. The application of this model to a case study in the Netherlands is detailed in the next chapter.

3.1. Problem description and assumptions

This section describes the problem and highlights the scope that is studied by explicitly listing the operational assumptions. It furthermore highlights how the problem has been translated into a computational model.

3.1.1. Problem characterisation and illustrative use case

As described in subsection 1.4.4, the functioning of ATO-over-ETCS is suboptimal for certain situations with disturbances. When disturbances are small enough to be absorbed by the running time supplement (either until the next stop or until the next larger station in the case of halt stations), there is no recalculation of the RTTP needed. The disturbance can, however, trigger blocking time conflicts and thus force braking for one or more EoA. The resulting train trajectory is less efficient in terms of punctuality and energy consumption. The approach suggested here aims to minimise the inefficiency by calculating and inserting in real-time one or more extra timing point(s) for the train affected by the small disturbance in order to let the ATO-OB optimise the speed profile and avoid excessive braking.

The rest of this thesis describes the proposed methodology as applied to an illustrative use case that occurs regularly during operations. Note that the proposed dynamic timing point approach can be applied in any situation where trains follow each other. It is potentially also beneficial for bidirectional conflicts; additional research is needed to explore this further.
The illustrative use case of this thesis (see Figure 3.1) is characterised by two trains following each other on the same track in the same direction, with the first train being slower than the second train (either due to maximum speed, stopping pattern or both). This is visible in the time-distance diagram: the slope of Train 1's trajectory is steeper than Train 2's. Train 1 experiences a small delay induced by increased dwell time, which is sufficient to result in a path conflict with the second train. This is common during operations: small dwell time delays make up approximately 90% of delay time in commuter train systems (Palmqvist et al., 2020).

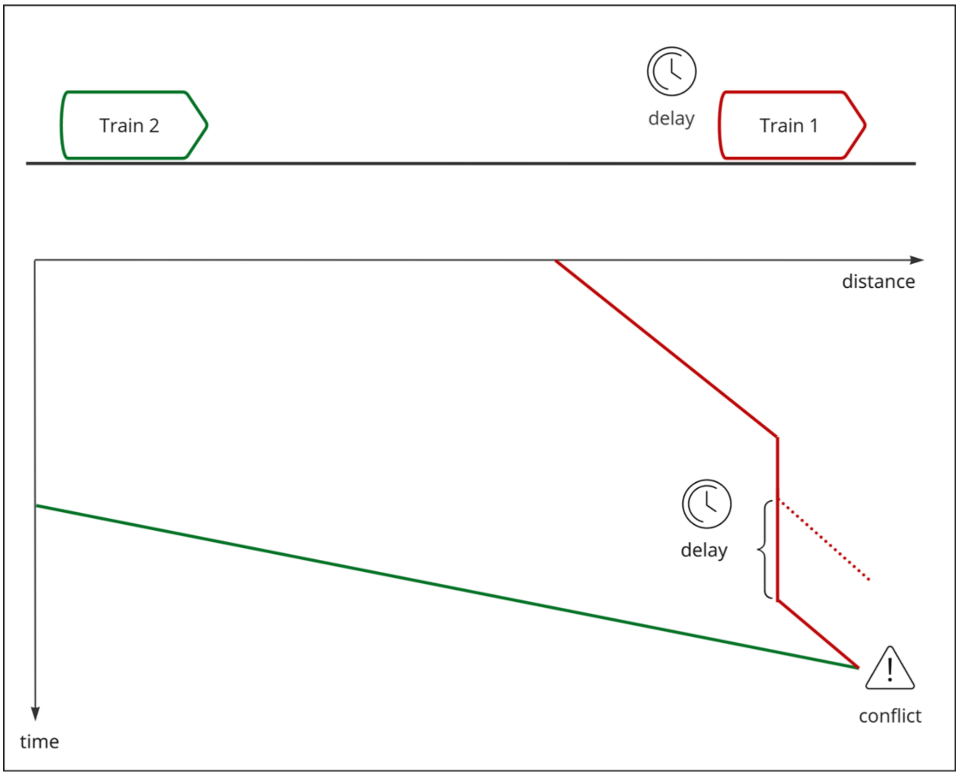


Figure 3.1: Overview of the illustrative use case.

This thesis will consider different scenarios characterised by a specific departure delay. Two variants of responding to this delay scenario will be modelled and contrasted:

1. The **uncontrolled case** where no action is taken by the trackside systems. Train 2 can not pre-emptively adapt its trajectory to the delayed preceding train and is therefore forced to brake for the EoA resulting from the conflict.
2. The **dynamic timing point case** where one or more timing points are inserted during operation and inform speed profile adaptation by the ATO-OB.

Both variants will be further explained and detailed in this chapter.

3.1.2. Scope and operational assumptions

The operational scope consists of the following assumptions:

- **Mainline railway system** equipped with **ETCS Level 2** based on **fixed blocks**
  This generally describes the future European mainline railway system. Note that advanced concepts of hybrid train detection and moving block are not explicitly considered, but the methodology can be generalised for these cases.
- **Conflict-free RTTP**
  The RTTP computed by the TMS is assumed to be conflict-free before the disturbance occurs.
- **All trains use EETC as driving strategy**
  This means that the timetable does not contain conflicts when all trains are using the EETC driving strategy. It also means that runtime calculation and traffic state prediction can be done accurately by assuming EETC.
- **Disturbances are modelled in real-time** as they occur
  Disturbances are modelled as extended dwell times at a stop of the first train. They are known only when they occur and cannot be anticipated. The following train reacts after the delay is known and communicated through the interaction mechanism.
- **Dynamic updating of trajectory** is possible
  The train is able to re-optimise the trajectory while driving based on an updated TPE. The train

control is subsequently adapted to follow the newly optimised speed profile. Note that the train trajectory encompasses position, speed and acceleration over time, while the corresponding speed profile focuses on speed variation over distance which is used to study and optimise energy consumption.

- Trains are equipped with **ATO Grade of Automation 2+**
  This thesis assumes that the EETC trajectory based on the latest TPE is followed precisely, implying the use of ATO Grade of Automation 2+. Note that this is not strictly necessary, as the concept of dynamic timing points can also be applied with Connected DAS. However, the impact of the resulting human factors in following the optimal trajectory will decrease the benefits. Therefore, this research assumes ATO Grade of Automation 2+.
  Moreover, the train trajectory modelling is specific to each train, accounting for speed limits, gradients, and train-dependent parameters such as traction characteristics, braking rates, and mass.

3.1.3. Introduction to computerised model

The dynamic timing point approach is implemented in a computerised model to study the effects of the approach on a detailed case study. The code base (implementation in computer programming code Python) used for this research is the same as in (Wang et al., 2023, 2025a,b). This code base is built to optimise the train trajectory on a chosen line, generally between two larger stations with one or more stops in between. There is no simulation involved, resulting in the need to carefully consider how real-time updates can be modelled accurately.
The code base focuses on calculating and comparing possible train trajectories for one or multiple trains. Blocking times can be calculated, but are not automatically enforced, so this needed to be built into the algorithms to handle the scenarios and control variants, see the following computational implementation notes. The setup time was defined as 1 second plus 5 seconds for every switch traversed (zero on open track). The reaction time was set to 1 second, and the release time to 2 seconds.

The speed profile optimisation algorithm discretises the continuous train control problem by discretising the distance and using this as a basis to calculate the trajectory according to the EETC model. The optimal control structure of the EETC problem is obtained by using Pontryagin's Maximum Principle and the Karush-Kuhn-Tucker conditions (see Goverde et al. (2021) for more details). The speed profile is subsequently generated from the optimal control structure.
The procedure is able to deal with discontinuities in gradients and speed limits along the tracks and relies on the sequence of switching regimes derived from the optimal control structure.

The research design of this thesis requires a conceptual translation from real-time insertion of dynamic timing points to a static model. The real-time insertion of a dynamic timing point needs to be modelled accurately while recalculating the full train trajectory.
The setup of the algorithm is based on minimising arrival delay first and subsequently optimising for energy use within the punctuality constraint for all trains. The focus on punctuality can lead to situations where the additional constraints of the disturbed scenarios result in increased cruise speeds with respect to the undisturbed scenario. This is inconsistent with a real-time scenario where the train has already travelled along the route with the cruising speed adapted to the undisturbed scenario. The model is therefore adapted to avoid this behaviour.

To ensure consistency in the driving behaviour of all trains, the speed limit of the line is reduced to the cruising speed of the undisturbed scenario for the relevant part of the line. The relevant part of the line can be described as:

- The speed limit in the original situation was higher than the cruising speed of the undisturbed situation (thus leaving room for inconsistency of higher initial cruising speed after recalculation).
  AND
- The first braking action due to the impact of the disturbance has not occurred yet.

This process forces both the first and the following train to follow the original undisturbed trajectory until the train is delayed or is able to respond to the delayed preceding train respectively. Note that this is relevant for all scenarios and variants.

The central function of the model is to recalculate train trajectories. The algorithm for this is schematically summarised in Figure 3.2. The first train uses only the top row and is not iterative. The results of this train trajectory are used for trajectory recalculation of the following train.

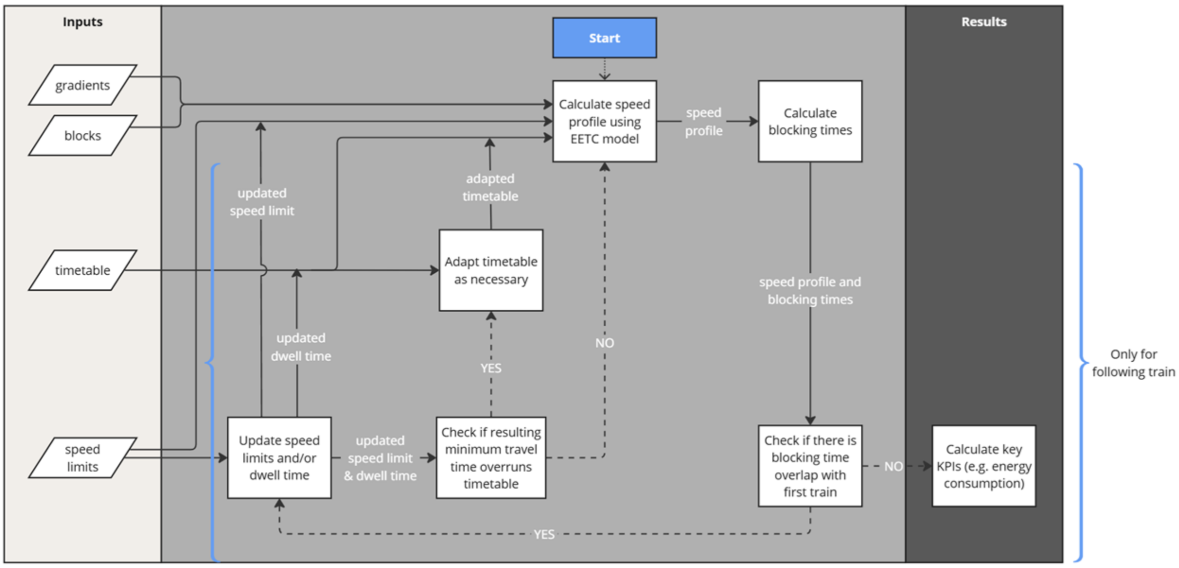


Figure 3.2: Schematic representation of the iterative train trajectory recalculation process.

The basic process of **train trajectory recalculation** is the same for all trains, including the delayed first train and the following train.

1. Load the infrastructure information (routes, gradients, blocks and stops) and train characteristics of the train in question.
2. Specify speed limits and timetable (including dwell times) in accordance with the situation.
3. Calculate train trajectory:
   (a) Calculate the speed and time trajectory as a function of distance using MTTC.
   (b) Check if this MTTC speed profile exceeds the scheduled arrival time and if so: adapt the scheduled time to guarantee a feasible train trajectory.
   (c) Use the EETC model to calculate the optimal cruising speed and coasting points, recalculate train trajectory including coasting (if available, otherwise: keep MTTC trajectory). The block in which the train stops at the station is always used to brake.
4. Calculate blocking times corresponding to the computed train trajectory.
5. For following train only: check for overlaps in blocking times of the delayed and following train. If there are one or more overlaps remaining, then the speed limit and/or dwell time during the forced stop needs to be adapted to force a later arrival at the conflicted block. The process of adapting the speed limit and/or dwell time depends on the control variant to be modelled and will be explained in the rest of this chapter.
   Repeat steps 2 to 5 iteratively for the following train until the overlaps are solved.
6. Calculate and add power consumption based on the control regimes of the final re-optimised train trajectory.

In the rest of this chapter, the reader finds **computational implementation notes** with explanations of how the train trajectory recalculation mechanism was used to quantify the impact of the proposed timing point approach.

Computational implementation note: Modelling uncontrolled case

The delay is modelled as an increase in the departure time of the preceding train of a train pair at a specific station. This simulates the effect of a longer dwell time, for example, due to an unexpectedly large number of passengers, extra time needed to assist a passenger with reduced mobility, or a malfunction in the train itself.
The trajectory of the delayed preceding train is recalculated using steps 1 to 4 as described above.

The resulting shift in the blocking time pattern of the preceding train can subsequently cause a blocking time conflict with the following train. In real life, this blocking time overlap would be a conflict which leads to not extending the MA when approaching the occupied section.

The algorithm models the forced braking of the following train by adapting the speed limit at this location and re-optimising the speed profile. In this calculation, the speed is limited to the cruising speed in the original trajectory up until the conflict to keep the driving behaviour consistent with the past trajectory.

The **speed limit/dwell time adaptation** (step 2 of train trajectory recalculation) to model the uncontrolled variant of the scenario can be described as:

i. Find the first block (in the direction of travel) in a shared section that has an overlap in blocking time between the two trains.
ii. Add a virtual speed limit for the first 10 metres of this conflicted block, reset to the original line speed limit directly after this 10 metres.
iii. The virtual speed limit is decreased in steps of 1 or 0.1 m/s (3.6 and 0.36 km/h), and the speed profile is recalculated and checked according to steps 1 to 5 of the train trajectory recalculation procedure described above. The reduction of the virtual speed limit is repeated iteratively as necessary until it eliminates the blocking time overlap between the delayed preceding train and the re-optimised trajectory of the following train at this block. If the virtual speed limit converges to zero, a stop of 1 second is inserted at the start of the conflicted block. In the following iterations, the dwell time at this forced stop is increased in steps of 1 second until the blocking time overlap is resolved.

The re-optimised trajectory is then checked for any subsequent blocking times that may overlap with those of the delayed preceding train. If blocking time overlaps are found, the procedure with virtual speed limit for the following train is repeated for these locations until no overlaps are left on the tracks shared by the two trains.

3.2. Dynamic timing point approach

This section introduces the concept of dynamic timing points as an approach to address the challenges faced when two successive ATO-equipped trains operate on the same corridor, particularly under disturbances caused by the preceding train.

3.2.1. Traffic harmonisation through dynamic computation of timing point

The concept of harmonising traffic through dynamic timing point insertion during operation is based on informing the following train of a potential future conflict as soon as possible. The dynamic timing point serves to communicate and resolve the anticipated conflict by leveraging accurate and rapid predictions, derived from (near) real-time localisation and speed data provided through train status reports fed back from the ATO-OB system to the ATO-TS system. The dynamic timing point is added while the train is running, aiming to solve the conflict by restricting the TPE and therefore the train trajectory of the following train.
The goal of this dynamic timing point is to increase punctuality and decrease energy use. This is possible because early warning allows the following train to reduce its speed earlier and pass the timing

point location at a higher speed compared to scenarios without such a dynamic timing point. The higher speed at the location where the conflict is resolved results in less time lost for acceleration, thus reducing the running time from the conflict to the next stop (i.e. increasing punctuality). Additionally, the reduction in acceleration also reduces energy consumption.

The proposed process of harmonising operations through inserting a dynamic timing point can be described as follows (see Figure 3.3):

1. Delay occurs: The preceding train (Train 1) experiences a departure delay, potentially causing a conflict with the following train (Train 2).
2. After its delayed departure, Train 1 sends an ATO Status Report from ATO-OB to the ATO-TS, including location, speed and delay. Note that it could be beneficial for Train 1 to send intermediate Status Reports while the departure delay is accumulating to inform Train 2 earlier. This will be explored in the case study.
3. The trackside intelligence uses the Status Report to detect potential conflicts based on real-time localisation, speed data, and trajectory predictions of both trains. If a conflict is identified, it calculates the optimal location and time window for a dynamic timing point (in the TPE of Train 2) to minimise the conflict.
4. The trackside intelligence updates the TPE of Train 2 to include the dynamically inserted timing point and communicates this information to Train 2.
5. Train 2 recalculates its trajectory using the updated TPE, adjusting its speed profile to harmonise operations, meet the (re)scheduled next target arrival time, and minimise energy consumption.

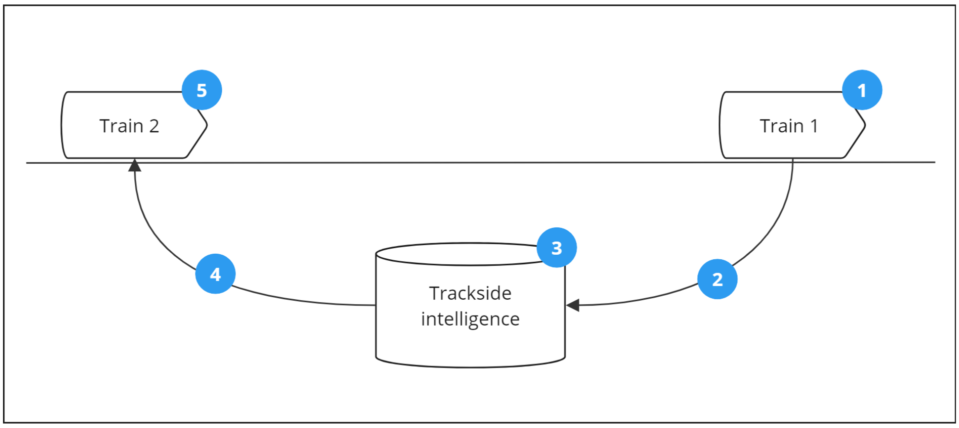


Figure 3.3: Overview of the operational process for dynamic timing points.

Note that the task of the trackside intelligence is to generate the updated TPE. Options for potential system architectures will be discussed in section 3.4.

Communication between Train 1 to the trackside intelligence (step 2) and from the trackside intelligence to Train 2 (stop 4) is essential and should happen as quickly and reliably as possible. Variations in the TPE processing time between the delayed departure of Train 1 and the trajectory adaptation of Train 2 will be explored in the case study.

3.2.2. Dynamic timing point calculation

The process described in subsection 3.2.1 requires the calculation of the dynamic timing point by trackside intelligence. The process is triggered when a potential blocking time conflict is detected between trains.

The dynamic timing point(s) is calculated by the trackside intelligence using the overlap in blocking time of the same block by subsequent trains:

1. **Blocking time stairways**
   The trackside intelligence calculates the independent blocking time stairways of the future paths

of the two trains based on the RTTP, current location and current speed of the train as contained in the Status Report. The blocking time stairway of Train 1 includes the delay, but conflicts are not taken into account for the following train: the calculation assumes that no other trains are present.

2. **Blocking time overlap per block section**
The trackside intelligence compares the independent blocking time stairways of the two trains and calculates the overlap per block section on any paths shared by both trains. An overlap in these blocking time stairs indicates a conflict that would be dealt with by the signalling system, and thus forced braking if no action is taken.
3. **Critical block**
The trackside intelligence identifies the critical block as the block on the shared path with the largest blocking time overlap. Let $p \in P$ denote a train, and $(p_i, p_j) \in P_2$ represent a successive pair of trains, where p_i precedes p_j while $b \in B_{p_i}$ represents a block section on the route of train p_i. The critical block is then

$$b = \arg\max_{b \in B_{p_j}} \left(c_{p_i,p_j,b} \;\middle|\; (p_i, p_j) \in P_2 \right) \tag{3.1}$$

where $c_{p_i,p_j,b}$ represents the blocking time overlap in a shared block section $b \in B_{p_i} \cap B_{p_j}$ (see Wang et al. (2025b), currently in final review before publishing).
4. **Timing point location and time window**
The location and time window of the dynamic timing point are chosen to correspond to solving the blocking time overlap at the critical block:
   - The timing point is placed at the start of the critical block
   - The time at which the first train clears this block is used as the earliest time of the timing point window
5. **Check for additional timing point needs**
The trackside intelligence recalculates the EETC trajectory resulting from the TPE with the dynamic timing point to determine whether this updated TPE solves all conflicts. Sometimes one dynamic timing point is not sufficient, for example if Train 2 speeds up faster than Train 1, resulting in additional run-ins. If overlapping blocking times remain, the process of adding an extra timing point is repeated until the TPE results in a conflict-free EETC trajectory. Additional timing points can be necessary both before and after the timing point to solve the initial critical block. Once all conflicts are solved the TPE, including dynamic timing point(s) can be sent to Train 2.

This critical-block strategy has been shown to be effective in resolving the full conflict between various driving strategies (Wang et al., 2023).

Computational implementation note: Modelling dynamic timing point calculation and response

The trackside intelligence uses the traffic state information (location and speed) of the two trains to calculate the dynamic timing point(s) needed. The ability to (re)calculate train trajectories is an essential function: this is used to check the need for additional dynamic timing points on top of the dynamic timing point at the critical block. The algorithmic implementation of the timing point can be summarised as follows:

Input: disturbed trajectories and blocking times of delayed and following train

1. Record the departure time of the delayed train and add the processing time of 30 seconds to the departure of the delayed train to calculate the time at which the following train receives the updated TPE. The size of the processing time would be standardised in the future, while this is not the case as of the writing of this thesis. 30 seconds was chosen as a generous limit in order to avoid overestimating the benefits of the dynamic timing point approach.
2. Find the location and speed of the following train at this receiving time using the disturbed

speed profile.

3. Ensure consistent driving behaviour prior to receiving the updated TPE by reducing the speed limit to the original cruising speed if necessary.
4. Calculate the dynamic timing points needed for the TPE update by optimising the train trajectory with one or more timing points as necessary to remove conflicts:
   (a) Find the critical block with the biggest blocking time overlap: its start is the **location of the timing point**.
   (b) Find the end of the delayed train's blocking time at the timing point location: this is the **start of the timing window** for the following train.
   (c) Find the required speed reduction of the following train to enable passing the timing point within the timing window: this is needed to check whether more timing points are required:
      i. Start with a trial speed just below the recorded speed at the moment that the TPE update arrives (see step 2).
      ii. Calculate the braking curve from the current location and speed, and record the location where the braking curve reaches the trial speed.
      iii. Insert a speed limit from this location until the location of the timing point.
      iv. Calculate the train trajectory and blocking times resulting from this sequence of speed limits (see explanation of train trajectory recalculation on page 28).
      v. Check for overlap between these blocking times and the blocking times of the delayed train at the timing point location.
         - If the conflict remains: reduce the trial speed and repeat. If the trial speed converges to zero: add a forced stop and increase dwell time by 1 second per iteration.
         - If there is no remaining conflict: the correct speed reduction or dwell time is found.
   (d) Check for remaining blocking time overlaps in other locations:
      - If there are further conflicts: repeat step 4 (a) to (c) with the speed profile including the response to the previously treated timing points.
      - If there are no remaining conflicts: the update to the TPE consists of the timing points that were calculated. The final conflict-free train trajectory is assumed to correspond directly to the speed profile that the ATO-OB calculates in response to this updated TPE.
5. Calculate the energy use and arrival delay of the train trajectory in response to the updated TPE.

3.3. Regime choices in train trajectory recalculation

This research focuses on the use case of two trains following each other on a track, with the former train experiencing a small delay. Determining the best strategy for re-optimisation of the trajectory of the following train after receiving the updated TPE belongs to the ATO-OB system and is out of scope for this thesis. It is considered briefly in this section for completeness, as it impacts the resulting benefits of the dynamic timing point methodology. An assumption for the onboard trajectory recalculation is made and motivated at the end of this section in order to be used in the dynamic timing point calculation.

The following train is located at X_{disturb} when the delayed train departs at time T_{disturb} (see Figure 3.4 for a graphical overview). The following train continues on its original path during the processing time $T_{\text{processing}}$ until it receives the updated TPE at location X_{update} at time $T_{\text{update}} = T_{\text{disturb}} + T_{\text{processing}}$. At that point in time and space, the following train needs to account for the warning time T_{warning} and the warning distance X_{warning} when adapting its trajectory. This is the time respectively distance that the following train has from T_{update} and X_{update} until the (first) dynamic timing point located at X_{tp} with clearing time (i.e. start of timing point window) T_{tp}.

The warning time can thus be described as

$$T_{\text{warning}} = T_{\text{tp}} - T_{\text{update}} \tag{3.2}$$

while the warning distance is correspondingly described as

$$X_{\text{warning}} = X_{\text{tp}} - X_{\text{update}} \tag{3.3}$$

There is a range of possible options for the trajectory of the following train from the point it receives the updated TPE with the dynamic timing point(s) included, see the option space in Figure 3.4. The option space is bounded by two extreme strategies:

1. **Latest possible trajectory adaptation**
   Let the following train follow the existing trajectory and brake just before the end of the MA. This strategy is akin to not implementing changes in train trajectory due to the updated TPE, indicated with *uncontrolled*.
2. **Direct full stop to maximum speed**
   Let the following train brake to a standstill directly and wait until it can speed up and pass the timing point at the maximum line speed at the earliest passing time. This allows for quicker covering of the distance between the timing point and the next stop. It also requires energy to accelerate from standstill and cruise at maximum line speed.

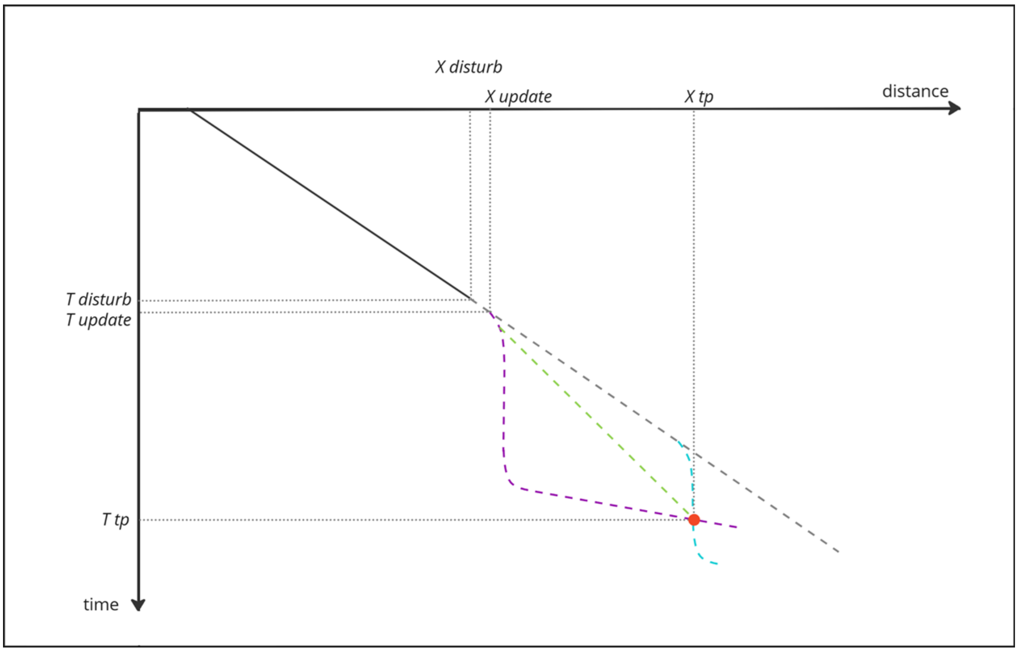


Figure 3.4: Distance-time diagram of the option space for the trajectory of the following train with one dynamic timing point, indicated in red. The "Latest possible trajectory adaptation" is drawn as the grey and blue trajectory, while the purple trajectory corresponds to "direct full stop to maximum speed". The green trajectory is described below.

Many other trajectories are possible, depending on the timing and strategy of speed reduction and re-acceleration. The options for speed reduction and optimisation of the speed while passing the timing point are discussed in the rest of this section.

3.3.1. Speed reduction strategy

Moving from the original trajectory (grey in Figure 3.4) results in increasing energy inefficiency due to the need to re-accelerate. The two main strategies for speed reduction are braking and coasting, although more complex combinations of braking, cruising and coasting are also possible. This section discusses the two most straightforward options:

1. **Braking**
   Full braking to the highest reduced cruising speed that will allow passing the timing point directly and driving with this speed up until the timing point. This minimises the braking and re-accelerating required for the situation and therefore minimises the energy usage associated with re-acceleration.
2. **Coasting**
   Reduce speed through coasting to the highest speed, which will allow the timing point to be passed directly. Coasting takes longer than braking, therefore, a coasting train will approach the timing point faster than a braking train. This strategy can thus only be used when the warning time and distance are sufficient.

The difference between these two potential strategies is illustrated in the speed-time sketch of Figure 3.5 with the curve of the minimum acceleration trajectory in green and the curve of the maximum coasting trajectory in yellow. This sketch assumes linear deceleration for both scenarios. This is appropriate for the braking scenario, which usually has an assumed constant deceleration equal to the service braking deceleration, e.g., 0.66 m/s^2 for intercity trains in the Netherlands. The coasting deceleration corresponds to the resistance experienced by the train and depends on both the train and track characteristics (curves and gradients). The train resistance varies with speed and is often modelled with the Davis equation as described in Equation 2.2. Typical values for speeds around 30 m/s (i.e., 108 km/h) are 0.05 m/s^2 and decreasing with slower speeds (Kim et al., 2006). Comparing this with the service braking deceleration shows that coasting will take over ten times more time to coast down than to brake to a given speed ($\frac{0.66}{0.05} = 13.2$). This value will increase at lower speeds because of the decreasing resistance.

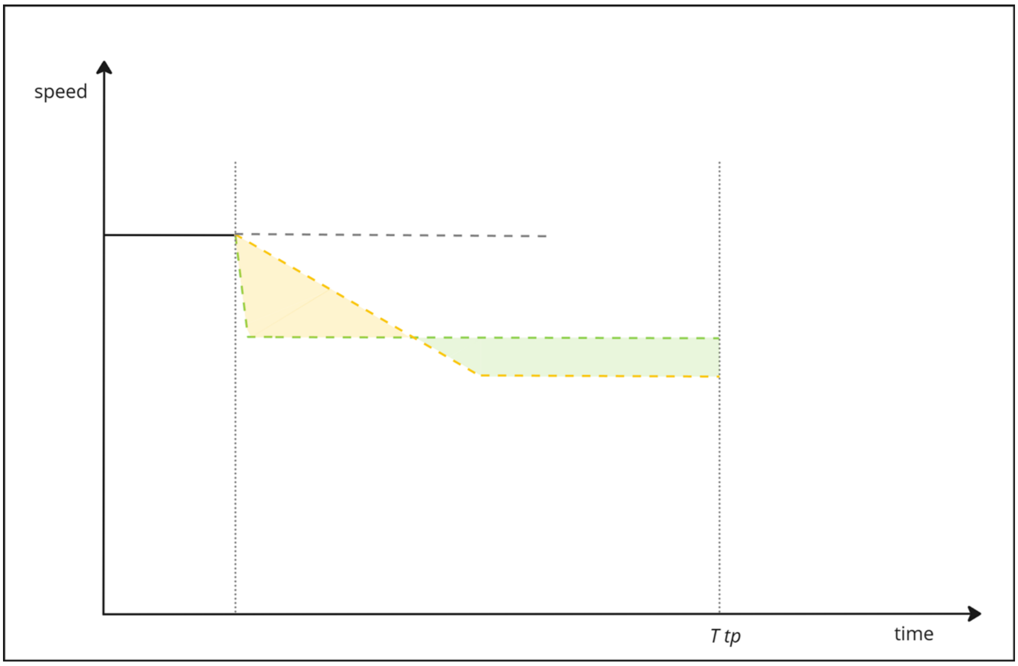


Figure 3.5: Sketch of the speed-time diagram on a flat track corresponding to braking (green dashed curve) and coasting (orange dashed curve) strategy after being notified of the timing point T_{tp}. The orange and green area is the extra distance travelled by the coasting strategy and braking strategy, respectively.

Both trajectories cover the same distance and can only arrive at the timing point location at the start of the timing point window T_{tp} (i.e., earliest possible time). The area under a speed-time curve represents the distance covered. Therefore, the areas shaded in yellow and green must be equal. As a result, the coasting strategy leads to a lower reduced cruising speed when compared to braking. Furthermore, it does reduce the duration of the cruising phase, thus saving cruising energy. However, it could require additional re-acceleration to strive for punctuality. It is not clear whether coasting will result in overall optimal energy consumption and if so, in which circumstances.

Note that the braking strategy increases the flexibility to react to further delays because the train reduces speed quicker and therefore, moves less quickly towards the area with conflicts. It could therefore be preferred over the coasting strategy even when coasting would save energy in the initial TPE update to increase flexibility to handle further TPE updates.

3.3.2. Passing speed optimisation

In addition to the strategic choice of initial braking versus coasting, the ATO-OB will consider the passing speed at the timing point location as a degree of freedom. There are two regimes that influence the need for optimisation of the passing speed, depending on the timetable, but also the distance and time to be covered after the timing point to the next stop:

1. **No punctuality challenges**
   The train has enough running time supplement after the timing point to arrive at the next planned stop on time. The passing speed does not need to be optimised to be punctual, but optimisation could be beneficial in order to increase coasting after the timing point.
2. **Punctuality in danger**
   The running time supplement after the timing point is not enough to arrive on time at the next stop. In order to minimise the delay, the passing speed needs to be increased. To achieve this, the train needs to brake deeper before the timing point and start re-accelerating before the timing point, passing it at a higher speed. This results in additional energy use for re-acceleration. There are two situations:
   (a) **Punctuality can be achieved**
   There is a passing speed below or at the speed limit that allows for punctual arrival at the next stop. In this case, this passing speed needs to be optimised to just allow for punctuality but minimise the additional energy use. The EETC trajectory can be applied after the timing point to optimise coasting if there is running time supplement left.
   (b) **Punctuality cannot be achieved**
   Passing the timing point at the speed limit is not enough to result in a punctual arrival at the next stop. In order to minimise the delay, the timing point will be passed at the speed limit, and the MTTC trajectory will be followed.

Computational implementation note: Modelling train trajectory re-optimisation

This thesis assumes that the trajectory re-optimisation is based on braking to a lower cruising speed directly after the updated TPE is received and does not use optimisation of the passing speed, see Table 3.1.

		Speed reduction strategy	
		Braking	Coasting
Passing speed optimisation	Yes		
	No	X	

Table 3.1: Overview of the possible strategic choices for trajectory optimisation after receiving a TPE update. The approach of this thesis indicated with X.

The choice for braking as an initial strategy is made to enable consistent use of a strategy across multiple scenarios, including scenarios where coasting is not possible because the warning time and distance are too small. Passing speed optimisation was not used because the methodology is intended for small disturbances, which would not result in delays at stops that may require the involvement of TMS to ensure a conflict-free RTTP.

3.4. Allocation of TPE generation with dynamic timing points

The generation of a TPE with dynamic timing points as proposed in this thesis needs to be allocated to a suitable location within the system architecture. The best location will have access to the information needed and be able to send the new TPE to the following trains.

3.4.1. Information requirements

The information needed to be able to dynamically update timing points includes:

- **Real-time traffic state**
  - **Description:** The location and speed of all relevant trains at this moment based on ETCS and ATO Status Reports provide the foundational input for dynamic adjustments.
  - **Key aspects:** In ETCS L2, positioning data can originate from radio communication as opposed to trackside detection systems like axle counters or track circuits. Onboard sensors provide speed and acceleration data, which combine with positioning information from balise passages. The resulting localisation and speed are sent via standardised ETCS and ATO Status Reports
- **Current Real-Time Traffic Plan (RTTP)**
  - **Description:** Current timetable that is conflict-free as far as possible
  - **Key aspects:** Containing information on timetable, route and possible rescheduling decisions on a network level
- **Current Journey Profile (JP)**
  - **Description:** Corresponding to the timetable that is currently followed, including the TPE with stopping locations and planned arrival times in the future, it serves as the baseline reference for operations.
  - **Key aspects:** The JP contains the timetable information at timing points detailing scheduled arrival, departure and dwell times at stations, as well as the allocations of trains to routes and current infrastructure conditions like temporary speed restrictions
- **Segment Profiles (SP)**
  - **Description:** Characteristics of the relevant infrastructure defining the physical and operational constraints of the route, affecting train paths and timing recalculations
  - **Key aspects:** The Segment Profiles are static and contain track geometry and signalling system information, including version and layout of blocks and location of timing points
- **Train characteristics**
  - **Description:** Knowledge of train-specific attributes ensures accurate recalculations for braking distances, acceleration, and maximum speeds. These characteristics significantly influence operational decisions
  - **Key aspects:** Key attributes include train length and weight, traction type and power but also braking systems and deceleration rates
- **Environmental variables**
  - **Description:** Environmental conditions that can directly or indirectly affect train operations, for example, through lower adhesion or a temporary speed restriction
  - **Key aspects:** Examples include weather conditions (e.g., rain, snow, fog, wind), seasonal influences (e.g., leaf fall reducing adhesion in autumn), temperature (e.g. track expansion in heat)

3.4.2. Options for TPE generation allocation

The generation of updated TPE, including the dynamic timing point(s), is preferably implemented in a system that is defined in the ETCS and ATO specifications. Adding a new system to the specification would be difficult due to the long development process, and therefore, only advisable when no other option exists.

Figure 3.6 gives an overview of the operational chain including systems, information concepts and actors as given by the European project MOTIONAL (FP1 MOTIONAL, 2024a).

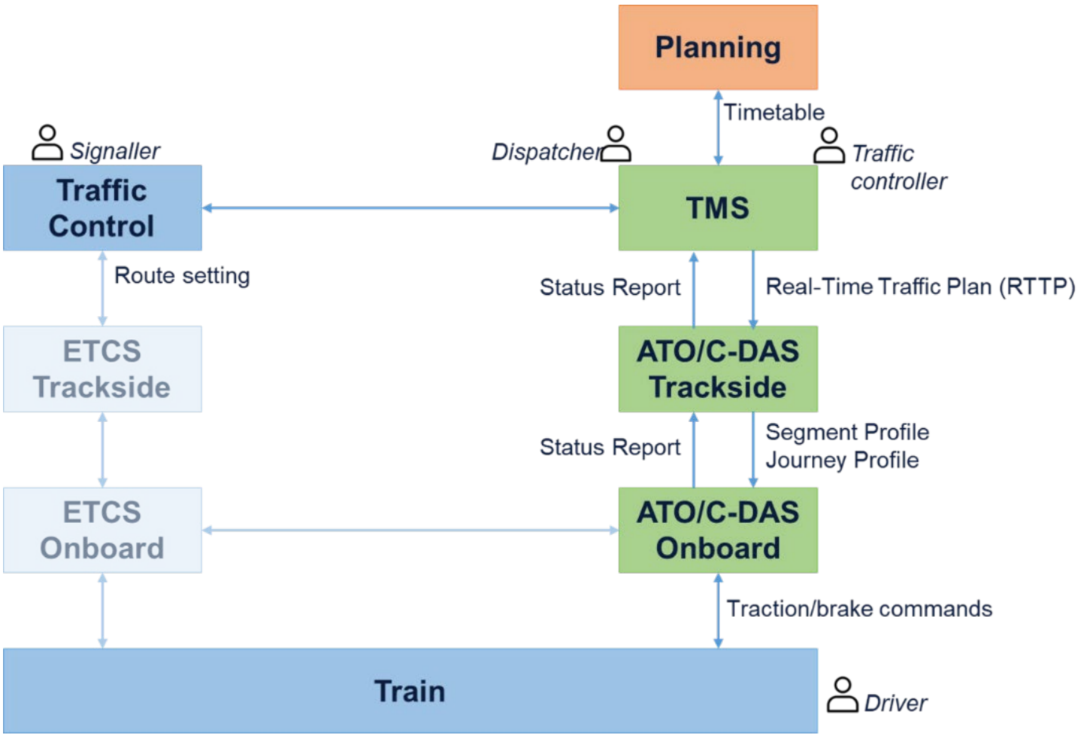


Figure 3.6: Overview of the operational chain from planning to train operation (FP1 MOTIONAL, 2024a).

The two potential locations for TPE generation described in this thesis are the TMS and the ATO-TS, since each of these can be developed to have the information required, and a timing point can be added both in the RTTP and TPE computation phases. Note that the ATO-TS still needs to be implemented in reality.

There are three essential concepts for information exchange in this framework: the RTTP that the TMS sends to the ATO-TS to communicate a conflict-free schedule with set route information for the network; the TPE that the ATO-TS sends to the ATO-OB to communicate the conflict-free train path as part of the JP based on the RTTP; and the train trajectory that details the speed profile that the train will (attempt to) follow. Table 3.2 gives an overview of the characteristics of each of these three concepts.

	RTTP	TPE	Train trajectory
Level of optimisation	Network	Corridor	Train
Goal	Conflict-free schedule	Conflict-free train path	Punctuality and energy optimisation
Input	Original schedule, Status Reports from ATO-TS and ETCS-TS	RTTP, ATO and ETCS Status Reports from train	TPE, train-based sensors
Scope	Full timetable control with rescheduling strategies	Train trajectory generation constraints	Reference train trajectory
Runtime calculation	Pre-calculated fixed values or simple calculation	To be determined	Detailed modelling with accurate parameters
Subsystem allocation	TMS	To be determined: TMS or ATO-TS	To be determined: ATO-TS or ATO-OB

Table 3.2: Summary of characteristics for RTTP, TPE and train trajectory in ATO-over-ETCS.

The full specification of the TPE and especially the system that calculates it is to be determined because the current system specifications do not define it.

The two architectural options for TPE generation are at the TMS level and at the ATO-TS level, as visualised in Figure 3.7.

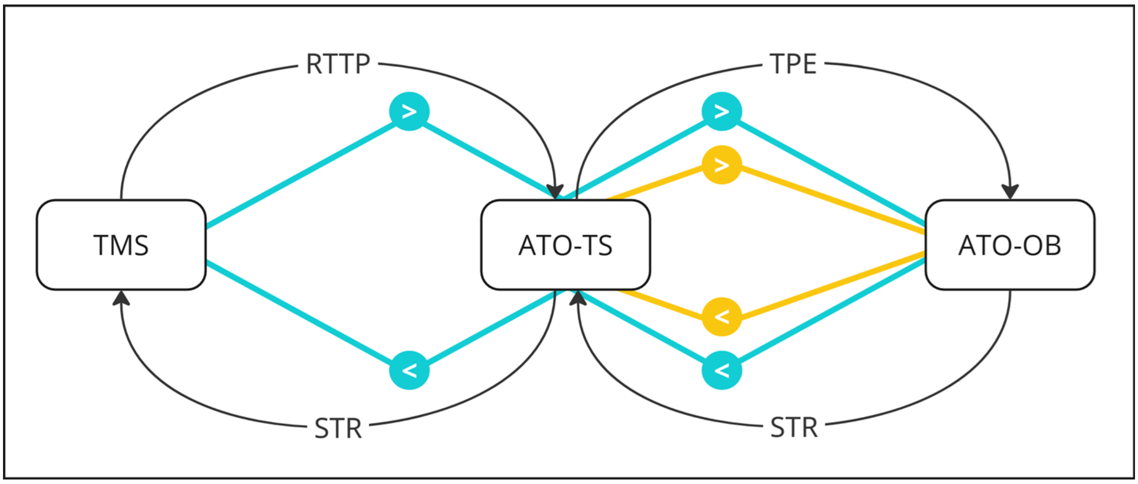


Figure 3.7: Overview of the systems, communication lines and feedback loops for trackside intelligence allocation at the TMS (blue feedback loop) and at ATO-TS (yellow feedback loop).

Option 1: TPE generation at TMS level

The TMS functions at a network level, taking into account all corridors and trains in its sphere of influence. The goal of the future TMS is to develop, update, and compute conflict-free RTTP that is constantly available. This can involve retiming train paths, local and global re-routing, cancellation, stop-skipping, short-turning, as well as reordering.
In this scenario, the TMS would generate and update the TPE while the ATO-TS is reduced to a server that receives the TPE and sends it onwards to the correct train.

The advantage of trackside intelligence for dynamic timing points at the TMS level is that it enables the use of the full array of traffic management measures, including reordering. It is possible to jointly optimise the delayed arrival time and trajectory since the TMS is able to adapt the scheduled arrival times at stops. This gives the IM a stronger degree of control.

Furthermore, the implementation of ATO-TS will be less complicated and therefore less costly.

However, there are some challenges with this approach, including the difference in level of detail in optimisation and runtime calculation, and the length of the feedback loops:

- The TMS is currently conceptualised as optimising the traffic on (a larger part of) the network. Locating TPE generation in this framework would result in computationally expensive situations where (a part of) the network governed by the TMS is re-optimised when one train reaches the pre-defined rescheduling threshold. This would be unworkable. Re-optimisation needs to be performed at the corridor level first and only expanded to the network level if the impact spreads beyond the corridor.
  Therefore, the TMS would need to be adapted to translate between the network level and the corridor level efficiently, optimising at both levels of detail and with a clear internal handover between the two. This distributed optimisation with network coordination would be similar to network-wide conflict detection and resolution, which has not been implemented yet.
- The level of detail in the runtime calculation is limited in current TMS systems: either a fixed value or a simple calculation when in between stops. This significantly speeds up the performance, which is essential for real-time operations with tens to hundreds of trains active at any point in time. However, this limited level of detail in runtime calculations would not be enough for the generation of dynamic timing points. Therefore, the level of detail in runtime calculations would need to be increased in the TMS to enable TPE generation and updating.
- The process of adding dynamic timing points would be (at least slightly) slower since the feedback loop goes from the ATO-OB via a Status Report to the ATO-TS and then with a Status Report to the TMS, see Figure 3.7. This feedback loop requires communication between ATO-OB, ATO-TS, and TMS to be active and with limited latency. Any additional time spent in communicating cannot be used to slow down early, limiting the effectiveness of the dynamic timing point methodology. The extent of the delay depends on the implementation and could range from negligible to impactful, especially when there are communication problems between TMS and ATO-TS while communication between ATO-TS and ATO-OB is functioning well.
  This challenge is limited when ATO-TS is implemented as a server with a fixed wire link to TMS, in which case the communication delay is minimal.

Option 2: TPE generation at the ATO-TS level
A second option for the allocation of TPE generation is the ATO-TS system.

The exact role of ATO-TS in the operational chain is not determined yet in the ERTMS/ATO specifications. Subset-125 defines the functional requirements for ATO-OB and ATO-TS with a general process for interaction, including:
"[An] "Update" JP specifies that the JP has been updated by the Control Centre, modifying or adding information to previously transmitted ones" (ERA and UNISIG and EEIG ERTMS Users Group, 2023).
This wording implies that the TMS ("Control Centre") is responsible for any updates in the information contained in the JP, but no explicit information is included.

The potential of ATO-TS, on the other hand, is to optimise traffic at the corridor level, aiming for a conflict-free schedule on this smaller scale. It could be set up to deliver an up-to-date TPE for all trains active on the corridor, but it cannot influence the route-setting in the infrastructure, which always requires TMS intervention. The enablement of runtime calculations with real-time positioning of the trains would be an essential function to build into ATO-TS in order to deliver the potential of the ATO-TS.

Allocating the TPE generation at the ATO-TS results in a shortened feedback loop, see Figure 3.7. The ATO-TS is well-placed in the system architecture to work at the corridor level and translate the RTTP to a JP for individual trains, which makes the computation simpler and quicker without needing multiple optimisation layers.
The downside of this allocation is the lack of infrastructure control, which makes holistic traffic management impossible by only using the ATO-TS control and limits the traffic management measures to restricting the train path between scheduled stops. Note that it is not possible to update the schedule

of arrival and departure at stops in the ATO-TS. When either of these is needed, it is necessary to involve the TMS to update the RTTP.

Conclusion TPE generation

Based on the characteristics of the two potential locations for TPE generation, the ATO-TS seems preferable: working at the appropriate level of optimisation with a quicker result due to a shorter feedback loop.
The limited scope of retiming only is generally not a problem for small disturbances, as considered in the scope of dynamic timing points. When the disturbance grows into a situation that results in delayed arrivals, the TMS would need to be involved to adapt the RTTP. This is outside the scope of the proposed methodology.
More research is needed to fully determine the best placement of TPE generation.

4

Case study of dynamic timing points between Utrecht - 's Hertogenbosch

This chapter quantifies the benefits of dynamic timing points using the computational model described in the previous chapter. The first section explains the case study, including the infrastructure, trains and corresponding traffic patterns and scenarios. The chapter then discusses the results of the case study in terms of energy savings and punctuality.

4.1. Case study description

The case study focuses on one of the busiest railway corridors in the Netherlands between Utrecht and 's-Hertogenbosch.

4.1.1. Infrastructure

The corridor from Utrecht to 's-Hertogenbosch is about 48 kilometres long and includes seven smaller stations in between the two large stations: Utrecht (Ut), Utrecht Vaartsche Rijn (Utvr), Utrecht Lunetten (Utln), Houten (Htn), Houten Castellum (Htnc), Culemborg (Cl), Geldermalsen (Gdm), Zaltbommel (Zbm) and 's-Hertogenbosch (Ht). All intermediate stations are served only by regional trains. See Figure 4.1 for an overview of the layout.

The corridor is four-track between Utrecht and Houten Castellum with regional trains separated from intercity trains. The corridor converges to double-track just after Houten Castellum and continues with two tracks until Geldermalsen, where the station contains multiple tracks per direction, which can be used for overtaking. Just after Geldermalsen, the corridor converges to double-track again until 's-Hertogenbosch.
The travel directions are separated from each other with traffic driving on the right, as is standard in the Netherlands. The case study focuses only on the traffic from Utrecht towards 's-Hertogenbosch because the directions are completely independent from each other. The methodology can be generalised to the opposite direction without adaptation.

The infrastructure model contains the location of all current signals, which are assumed to be stop marker boards under ETCS L2 operation. Furthermore, the changes in speed limits and track gradient are included, with the speed limit and gradient assumed piece-wise constant between these change points.

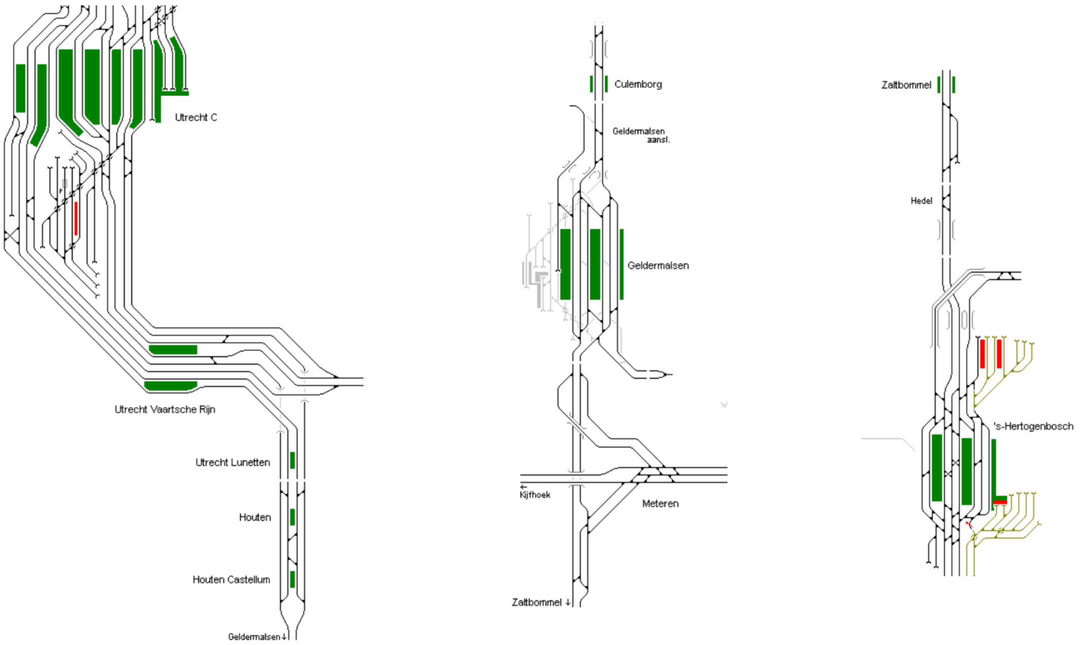


Figure 4.1: Overview of the infrastructure between Utrecht and 's-Hertogenbosch in three parts: Utrecht Centraal to Houten Castellum (left), Culemborg to just before Zaltbommel (middle) and Zaltbommel to 's-Hertogenbosch (Zeegers, 2024).

Figure 4.2 shows the layout and signal locations of the two important parts of the infrastructure. This shows that the shared blocks between Houten Castellum and Geldermalsen run from 10.434 up to and including 24.128 km.

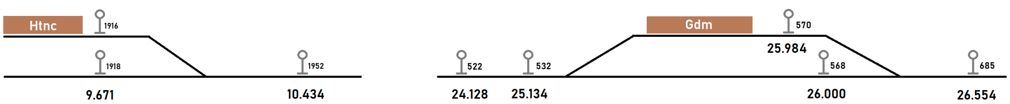


Figure 4.2: Schematic overview of the tracks after Houten Castellum and around Geldermalsen.

4.1.2. Trains and traffic patterns

The case study includes regional and intercity trains, which are called "Sprinter" and "Intercity" in the Netherlands. The characteristics of both train types are listed in Table 4.1.

Table 4.1: Train parameters used in the case study.

Parameter	Sprinter	Intercity
Train length (m)	101	162
Train mass (tonnes)	198	391
Rotating mass factor (-)	1.06	1.06
Max. traction power (kW)	1918	2157
Max. traction force (kN)	179	214
Max. speed (km/h)	140	140
Service braking deceleration (m/s^2)	0.8	0.66

The resistance of both trains is modelled using the Davis Equation with parameters A for rolling resistance, B for speed-proportional resistance (including bearing friction) and C for quadratic resistance including aerodynamic drag.

Table 4.2: Davis Equation parameters.

Parameter	Sprinter	Intercity
A (N)	1.3961	5.8584
B (N/(m/s))	0.0145	0.0206
C (N/(m/s)^2)	0.0007	0.001

The timetable used for this case study corresponds with the timetable for 2022. The basic hourly pattern consists of alternating Intercity and Sprinter trains departing between 3 and 8 minutes from each other for a total of 6 Intercity and 6 Sprinter trains per hour. See Table 4.3 for the timetable in minutes for an overview. Note that trains are planned in units of 6 seconds in the Netherlands, which allows for more detail. The exact timetable in seconds is included in Appendix A.

Table 4.3: Case study timetable in minutes with relevant trains highlighted, see scenario description.

		Spr 1	IC 1	Spr 2	IC 2	Spr 3	IC 3	Spr 4	IC 4
Ut	**d**	00	03	11	14	22	24	30	33
Utvr	**a**	02	|	13	|	24	|	32	|
	d	03	|	14	|	25	|	33	|
Utln	**a**	05	|	16	|	27	|	35	|
	d	06	|	17	|	28	|	36	|
Htn	**a**	09	|	20	|	31	|	39	|
	d	09	|	20	|	31	|	39	|
Htnc	**a**	12	|	23	|	34	|	42	|
	d		|	24	|	35	|		|
Cl	**a**		|	30	|	41	|		|
	d		|	31	|	42	|		|
Gdm	**a**		|	36	|	47	|		|
	d		|	42	|		|		|
Zbm	**a**		|	48	|		|		|
	d		|	49	|		|		|
Ht	**a**		30	57	41		51		60

Note that the pattern of trains repeats after 30 minutes while the intervals within the half hour are not regular. The intervals between subsequent Sprinter departures at Utrecht are 11 minutes, 11 minutes and 8 minutes, while Intercity travellers have subsequent departure intervals of 11, 10 and 9 minutes. Sprinter 1 stops at Houten Castellum to return to Utrecht, while Sprinter 3 leaves the track after Geldermalsen to continue on a branch line to Tiel. Sprinter 2 and all Intercity trains continue to 's-Hertogenbosch.

4.1.3. Description of scenarios

The case study focuses primarily on delay scenarios where Sprinter 2 departs with a delay of 1, 2, 3, 4 or 5 minutes from station Houten Castellum, Culemborg or Zaltbommel. These delays are common in practice and could, depending on circumstances, potentially be solved without needing traffic control measures via the TMS. For all delay scenarios at Houten Castellum and Culemborg, the directly following train (Intercity 2) passes Sprinter 2 in the four-track section before Houten Castellum, but Intercity 3 subsequently runs into the delayed Sprinter at Culemborg and needs to brake. Intercity 3 takes over Sprinter 2 at Geldermalsen. The interaction between Sprinter 2 and Intercity 3 is chosen because the respective train paths are close to each other at Culemborg. Therefore a small delay of the Sprinter already impacts the Intercity, highlighting the potential of the dynamic timing point approach. Houten Castellum and Culemborg are located in the first half of the section from Utrecht to 's-Hertogenbosch and therefore the Intercity will have running time supplement left to use.
In the case of a departure delay at Zaltbommel, it is Intercity 4 that runs into the delayed Sprinter 2.

Figure 4.3 shows the time-distance diagram of the Sprinter and the two Intercity trains along the corridor Utrecht to 's-Hertogenbosch. The IC 3 overtaking the Sprinter at Geldermalsen is visible in the centre. The grey areas indicate regions where the Sprinter follows a different track and thus does not interact with the Intercity trains.

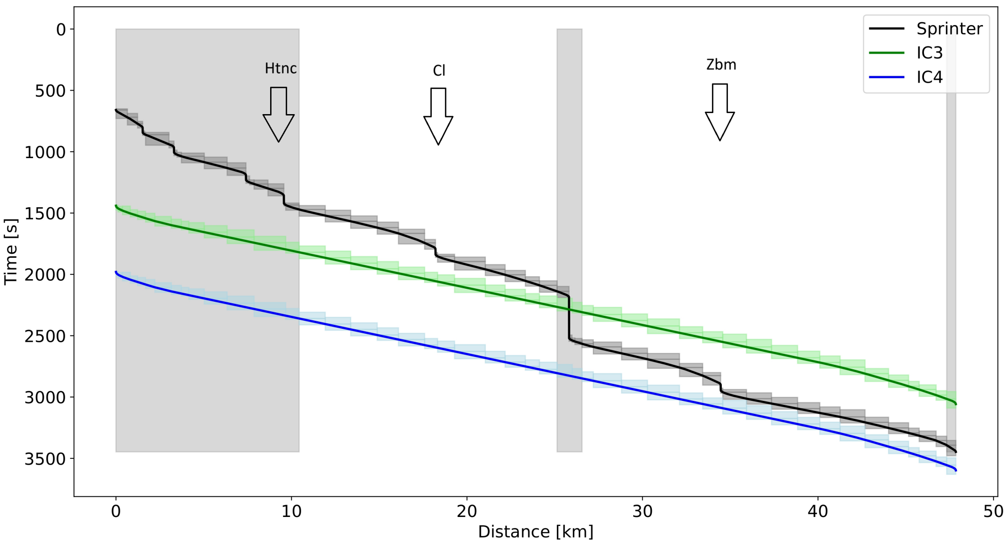


Figure 4.3: Overview of the time-distance diagrams including blocking times of three trains involved in the experiments.

The overview of all scenarios is given in Table 4.4. Note that the on-time scenario at Htnc and Cl is the same: the undisturbed running of IC 3. This is therefore only listed once.

Table 4.4: Overview of scenarios to be studied with the Intercity trains that experience a direct conflict with the delayed Sprinter 2. The scenario without delay is included as well.

Delay of Spr	**Impacted IC when delay at station**			**Variant(s) per scenario**
	Htnc	**Cl**	**Zbm**	
on time	IC 3		IC 4	undisturbed
1 min	IC 3	IC 3	IC 4	uncontrolled & timing point
2 min	IC 3	IC 3	IC 4	uncontrolled & timing point
3 min	IC 3	IC 3	IC 4	uncontrolled & timing point
4 min	n/a	IC 3	IC 4	uncontrolled & timing point
5 min	n/a	IC 3	IC 4	uncontrolled & timing point

The 4- and 5-minute delay scenarios at Houten Castellum are not investigated because the delays at Houten Castellum occur at the four-track section with the Sprinter occupying a separate track. As the delay increases, the Intercity arrives at the merging point to the double-track section closer to the Sprinter. At a 4- and 5-minute delay, the Intercity arrives at the merging point first. In practice, the trains will be reordered, and no conflict will occur between the delayed Sprinter and the Intercity at all. However, reordering is not possible in the dynamic timing point methodology proposed here because it involves a change in route setting with respect to the RTTP, thus requiring involvement of the TMS. Therefore, these delay scenarios at Houten Castellum are disregarded.

All delay scenarios are studied in the uncontrolled variant and the dynamic timing point variant:

- In the **uncontrolled case**, the Intercity does not get any information from the delayed Sprinter. It therefore follows its original EETC trajectory until the EoA forces it to brake.

- In the **dynamic timing point case**, the Status Report of the delayed departure of the Sprinter arrives at the system responsible for TPE generation, which then calculates the required timing point(s) for the Intercity, updates the TPE and sends it to the Intercity. The Intercity subsequently recalculates its trajectory to incorporate the dynamic timing point(s) and continues with this adapted trajectory.

All scenarios assume the TPE processing time from departure of the delayed train until speed adaptation by the following train to be 30 seconds.

Next to the analysis of these scenarios, two more sets of experiments are carried out:

- The scenario with a 3-minute delay at Culemborg is used to study the benefits of **intermediate updates** of the delay. This represents the situation where the delayed Sprinter sends an update that it has not departed for every minute of delay accumulation.
- The scenario with a 2-minute delay at Culemborg is used to study the **sensitivity to a change in TPE processing time**. This represents a situation where the processing of the TPE update is quicker or slower than assumed.

The resulting set of experiments therefore includes 2 ”on time” scenarios, 13 ”uncontrolled” and 13 ”dynamic timing point” variants of the delay scenarios and additionally 1 variant with intermediate updates and 1 variant with variation in the TPE processing time. This thesis thus includes 30 experiments.

4.1.4. Modelling methodology

The modelling of the scenarios follows the approach described in the previous chapter. The methodology can be summarised as follows:

1. Calculate the characteristics of the **undisturbed trajectory** of IC3 and IC4 in order to compare with the performance of the other trajectories to be calculated.
2. Choose a delay scenario
   (a) Calculate the trajectory and blocking times of the delayed Sprinter in that specific delay scenario.
   (b) Calculate the impact of the blocking times of the Sprinter on the trajectory of the Intercity as described in the previous chapter, and calculate the performance of this **uncontrolled trajectory** in terms of arrival delay, energy consumption and number of forced braking regimes.
   (c) Calculate the updated TPE for the Intercity and the resulting impact on the performance of the **dynamic timing point trajectory**.
      i. Calculate the location of the dynamic timing point(s) and the time window at this location(s).
      ii. Adapt the trajectory of the Intercity to the dynamic timing point iteratively as described in the previous chapter.
   (d) Calculate the performance of the Intercity when following the dynamic timing point trajectory in terms of arrival delay, energy consumption and forced braking regimes.
3. Repeat step 2 for all delay scenarios.

Subsequently, the performance of the two possible control variants per delay scenario (without and with dynamic timing point, outcomes of steps 3 and 4, respectively) can be compared with each other and the undisturbed scenario.

4.1.5. Performance and assessment criteria

The key KPIs to assess the performance of the dynamic timing point approach are:

1. the arrival delay (in seconds) of the Intercity in ’s-Hertogenbosch
2. the energy consumption (in kWh) of the Intercity trajectory
3. the number of forced braking regimes that are avoided for the Intercity (compared to the uncontrolled variant of the scenario)

The arrival delay and energy consumption after adaptation to the dynamic timing point(s) are compared to the arrival delay and energy consumption of the uncontrolled variant and the undisturbed scenario. The quality of the dynamic timing point(s) in the TPE update is assessed by how many forced braking regimes it can make the Intercity avoid.

The warning time and warning distance are recorded as explanatory variables and used to explain differences in energy consumption. The warning time is the difference between the time when the intercity adapts its trajectory after it received the updated TPE and the start of the timing point window. The warning distance is the distance between the location where the intercity was when it adapted its trajectory to the updated TPE and the timing point location.

4.2. Timing point benefits for 2-minute delay at Culemborg

This section describes the train trajectory results for the Intercity train in the undisturbed scenario and in the scenario where the Sprinter experiences a 2-minute delay at Culemborg, detailing both the uncontrolled and dynamic timing point variant of the latter scenario. This detailed section shows the application of the approach and computer model for this delay scenario. This scenario is explained in depth because the impact of the delay is significant without control measures but can be managed well with the timing point approach.

The results of the other scenarios (including intermediate delay updates and sensitivity analysis) and detailed comparisons between scenarios will follow in further sections. The speed profiles of these scenarios can be found in Appendix B.

4.2.1. Undisturbed scenario: on-time running

When there is no delay, the Intercity train will be able to run its energy-optimised trajectory without any interference because the timetable is conflict-free. The EETC algorithm results in an optimal cruising speed of 118.9 km/h and incorporates coasting for approximately the last 10 kilometres, see Figure 4.4. This is accurate for both IC 3 and IC 4, as their scheduled running times are identical, just shifted in time.

The resulting energy use for this run is 313 kWh, arriving punctually.

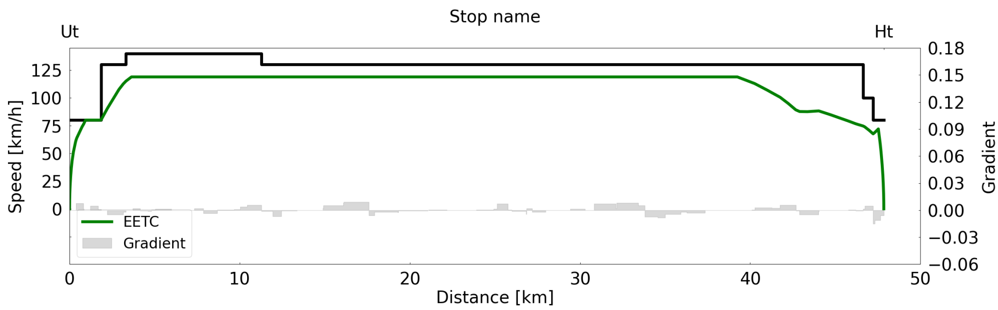


Figure 4.4: Speed profile of Intercity 3 in the base scenario without delays in green with the speed limit over the route in black. The gradients are expressed as grey boxes.

4.2.2. Uncontrolled response to 2-minute delay at Culemborg

In the uncontrolled and timing point variant of this scenario, the Sprinter train has a 2-minute delay at the departure from Culemborg (Cl). As a result of the departure delay of the Sprinter in the uncontrolled variant, the blocking times for the Sprinter prevent the Intercity train from following its trajectory, forcing it to brake.

The blocking times are calculated from both trains to identify the first block with overlap, see Figure 4.5. The section where both trains follow the same track is shown in white: from the 4-track to double-track switch at 10.434 km to the entry signal for Geldermalsen at 25.195 km, where overtaking is possible again. The grey parts of the graph indicate sections where the Sprinter and Intercity do not follow the same tracks.

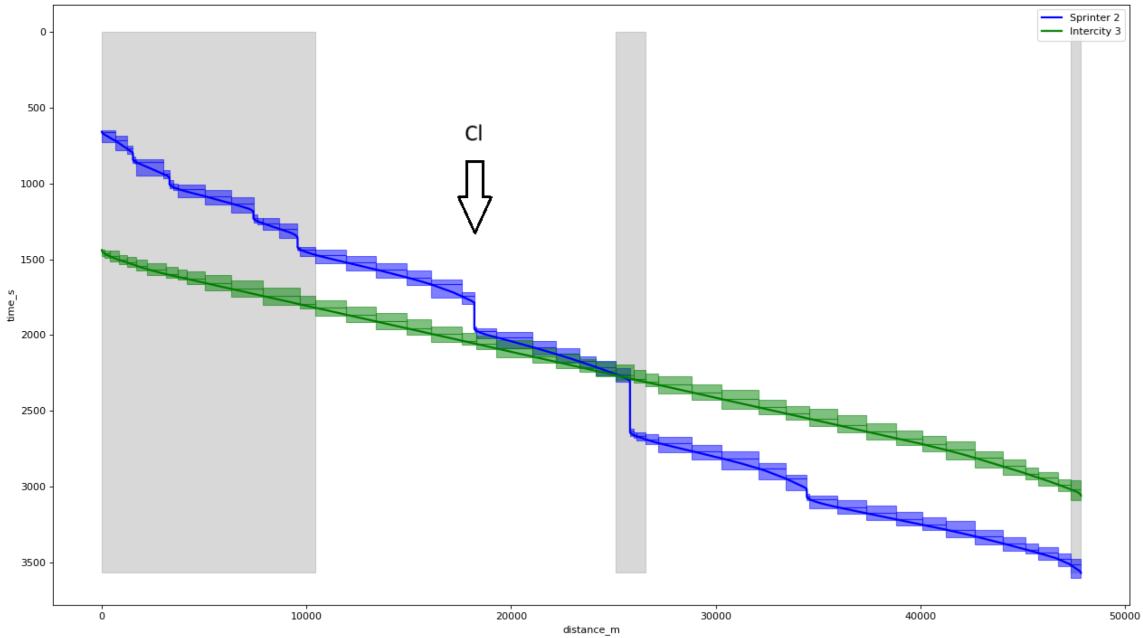


Figure 4.5: Blocking time diagram for both trains when the Sprinter departs with a 2-minute delay from Culemborg

The initial blocking time overlap starts at the block of Culemborg station at 18.324 km, but is small enough not to force the Intercity to a complete stop. After braking to 21.6 km/h, the Intercity's MA is extended and the train can continue. The Intercity does not get additional information in this variant, therefore it accelerates at full speed to the new EETC cruising speed. Due to the Intercity driving closely behind the Sprinter, there are two additional forced braking regimes: at 22.172 km, the Intercity needs to brake to 20.3 m/s, while the Intercity is forced to brake to 32.4 km/h at 24.128 km. The resulting speed profile for the Intercity then becomes Figure 4.6.

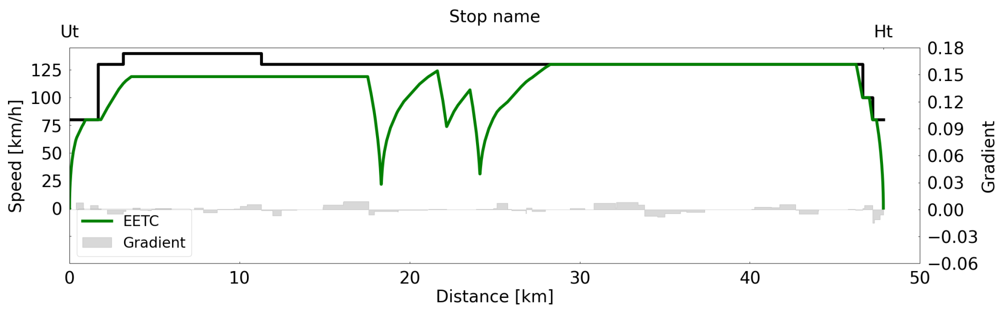


Figure 4.6: Speed profile for Intercity following a Sprinter with a 2-minute delay from Culemborg.

The energy usage of this trajectory is 535 kWh, which is a 70.9% increase with respect to the undisturbed scenario. The arrival is, however, just in time with 7 seconds to spare due to maximum usage of the remaining running time supplement. The train, therefore, follows the minimum time profile without any coasting after the last braking action.

4.2.3. Dynamic timing point response to 2-minute delay at Culemborg

Subsequently, the dynamic timing point approach is used to generate a new TPE that removes the blocking time overlap. This timing point is located at 24.128 km (just before Geldermalsen) and can be passed starting from 2125 seconds (35 minutes and 25 seconds). The Intercity receives this timing point and can start adapting its trajectory at 15.929 km and 1986 seconds (33 minutes and 6 seconds). The warning distance is thus 8.199 km while the warning time is 2 minutes and 19 seconds.
The Intercity brakes to 85.3 km/h and keeps this reduced cruising speed until the dynamic timing point, resulting in the speed profile of Figure 4.7.

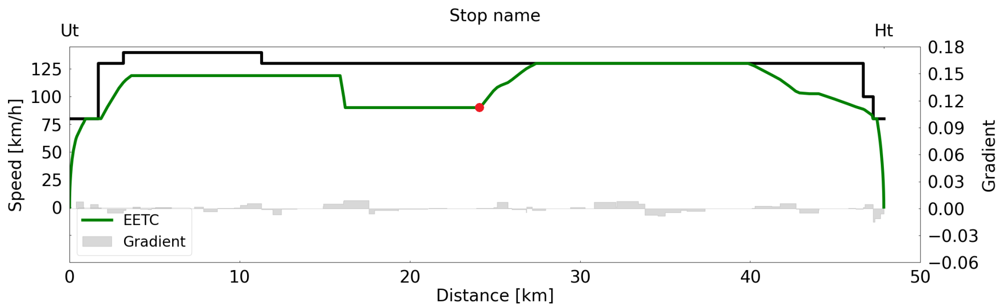


Figure 4.7: Dynamic timing point speed profile for Intercity following a Sprinter with a 2-minute delay from Culemborg, with timing point indicated in red.

With this dynamic timing point, the energy usage is 343 kWh (6.6% increase with respect to the undisturbed situation) while the arrival is punctual. Energy consumption shows a 64.3 percentage point decrease with respect to the uncontrolled variant without a timing point. This is possible due to the reduction of three forced braking phases to zero. The Intercity is only braking to reduce its speed for the dynamic timing point, while there is running time supplement left to coast towards the end of the trajectory. The shape of the speed profile indicates that the strategy of directly braking to the reduced cruising speed after receiving the updated TPE could be suboptimal. Coasting to reduce speed after receiving the dynamic timing point could potentially deliver additional energy savings while being punctual. This is not further explored to keep the results comparable to other scenarios where coasting is not possible.

4.3. Overview of timing point benefits

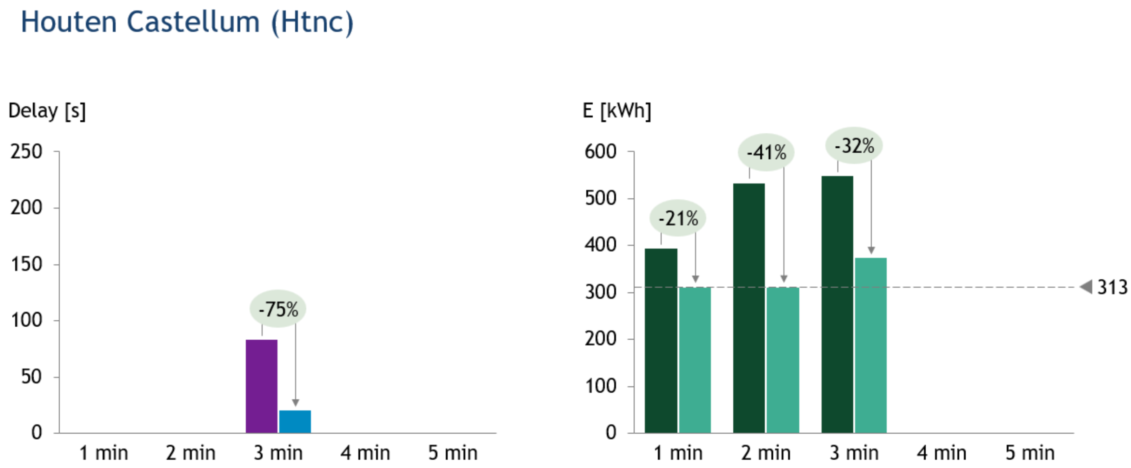
The arrival delay and energy consumption for the Intercity in all scenarios are summarised in Table 4.5. The first two columns describe the delay of the Sprinter, the rest of the table describes the Intercity situation and outcomes. The increase in arrival time and energy consumption with respect to the on-time scenario is indicated in brackets. The scheduled journey time is 1625 seconds, split in 1425 seconds minimum running time and 200 seconds (14%) of running time supplement.

All scenarios with a 1-minute delay and the 2-minute delay at Houten Castellum and Culemborg do not experience any arrival delays at 's-Hertogenbosch and are thus suitable for treatment without updating the RTTP. The scenarios with a 3-minute delay at Houten Castellum and Culemborg and the 2-minute delay at Zaltbommel have an arrival delay of less than 1 minute with the dynamic timing point approach.

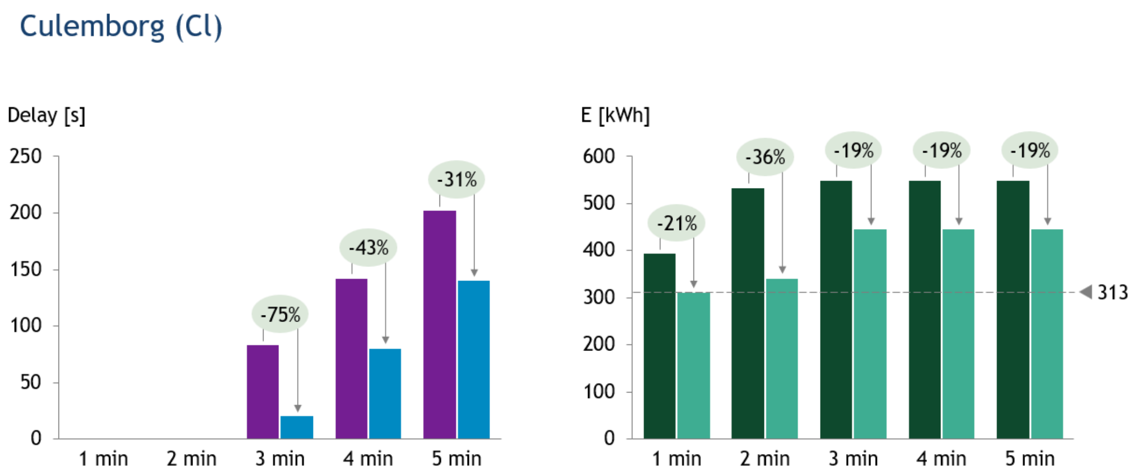
Figure 4.8 visually represents the key KPIs of arrival delay and energy consumption to compare the overall reduction that the dynamic timing point approach brings.

Table 4.5: Overview of arrival delay and energy consumption including comparison to on-time running of the Intercity corresponding to the Sprinter delay scenarios in the first two columns.

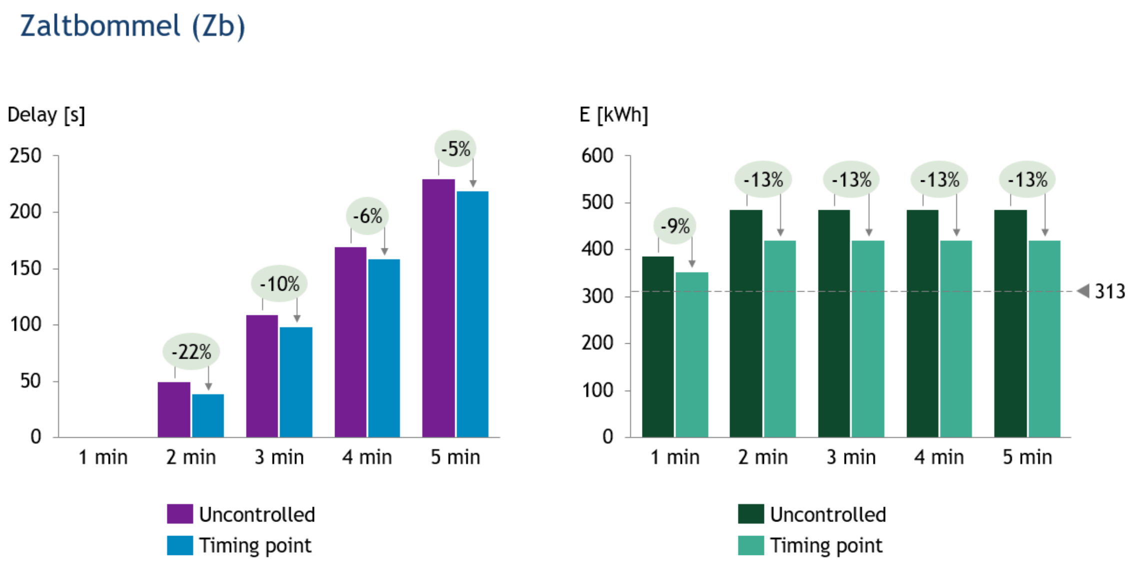
Scenario		Variant	Forced braking	Timing point	Arrival delay [s]	Energy [kWh]	
on time		undisturbed	0x	N/A	+0	313	
Htnc	1 min	uncontrolled	1x: 24.128 km	N/A	+0	395	+82 (+26.2%)
		timing point	0x	1x: 24.128 km	+0	313	+0
	2 min	uncontrolled	3x: 18.324, 22.172 & 24.128 km	N/A	+0	535	+222 (+70.9%)
		timing point	0x	1x: 24.128 km	+0	313	+0
	3 min	uncontrolled	3x: 17.585, 19.292 & 24.128 km	N/A	+84 (+5.2%)	550	+237 (+75.7%)
		timing point	0x	1x: 24.128 km	+22 (+1.4%)	375	+62 (+19.8%)
Cl	1 min	uncontrolled	1x: 24.128 km	N/A	+0	395	+82 (+26.2%)
		timing point	0x	1x: 24.128 km	+0	313	+0
	2 min	uncontrolled	3x: 18.324, 22.172 & 24.128 km	N/A	+0	535	+222 (+70.9%)
		timing point	0x	2x: 19.292 km & 24.128 km	+0	343	+20 (+6.6%)
	3 min	uncontrolled	3x: 17.585, 19.292 & 24.128 km	N/A	+84 (+5.2%)	550	+237 (+75.7%)
		timing point	1x: 17.585 km	2x: 19.292 km & 24.128 km	+22 (+1.4%)	447	+134 (+42.8%)
	4 min	uncontrolled	3x: 17.585, 19.292 & 24.128 km	N/A	+143 (+8.8%)	550	+237 (+75.7%)
		timing point	1x: 17.585 km	2x: 19.292 km & 24.128 km	+82 (+5.0%)	447	+134 (+42.8%)
	5 min	uncontrolled	3x: 17.585, 19.292 & 24.128 km	N/A	+203 (+12.5%)	550	+237 (+75.7%)
		timing point	1x: 17.585 km	2x: 19.292 km & 24.128 km	+142 (+8.7%)	447	+134 (+42.8%)
Zbm	1 min	uncontrolled	1x: 33.429 km	N/A	+0	387	+74 (+23.6%)
		timing point	0x	2x: 34.574 & 40.262 km	+0	354	+41 (+13.1%)
	2 min	uncontrolled	1x: 33.429 km	N/A	+50 (+3.1%)	486	+173 (+55.3%)
		timing point	1x: 33.429 km	2x: 41.259 & 46.315 km	+39 (+2.4%)	421	+108 (+34.5%)
	3 min	uncontrolled	2x: 33.429 & 41.259 km	N/A	+110 (+6.8%)	486	+173 (+55.3%)
		timing point	1x: 33.429 km	2x: 41.259 & 46.315 km	+99 (+6.1%)	421	+108 (+34.5%)
	4 min	uncontrolled	2x: 33.429 & 41.259 km	N/A	+170 (+10.5%)	486	+173 (+55.3%)
		timing point	1x: 33.429 km	2x: 41.259 & 46.315 km	+159 (+9.8%)	421	+108 (+34.5%)
	5 min	uncontrolled	2x: 33.429 & 41.259 km	N/A	+230 (+14.2%)	486	+173 (+55.3%)
		timing point	1x: 33.429 km	2x: 41.259 & 46.315 km	+219 (+13.5%)	421	+108 (+34.5%)



(a) Intercity arrival delay and energy consumption in a scenario with Sprinter delay originating in Houten Castellum.



(b) Intercity arrival delay and energy consumption in a scenario with Sprinter delay originating in Culemborg.



(c) Intercity arrival delay and energy consumption in a scenario with Sprinter delay originating in Zaltbommel.

Figure 4.8: Resulting arrival delays and energy consumption for all delay scenarios, the dotted line represents the energy consumption for the undisturbed scenario.

The dynamic timing point approach added 1 or 2 timing points in each of the 13 delay scenarios. It was able to give performance benefits for all scenarios in terms of energy consumption. All scenarios where the delayed Sprinter induced arrival delay for the Intercity experience benefit in decreasing this delay. There are clear patterns in the data based on location and increasing delay. The benefits of timing points are most significant for the scenarios with small delays of the preceding Sprinter.

4.3.1. Effect of warning time and warning distance

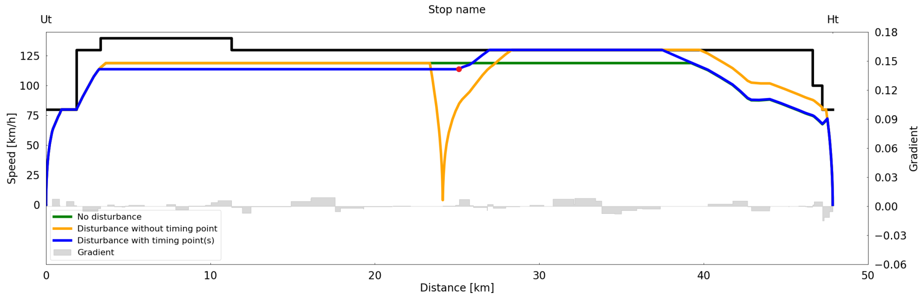
The performance of the dynamic timing point approach benefits from warning time where the headway between the Intercity and Sprinter allows the Intercity sufficient time to brake after it receives the updated TPE with dynamic timing point.
Table 4.6 shows the warning time and distance of the dynamic timing point variants of the delay scenarios. The warning distance and time are zero if the Intercity was forced to brake before the updated TPE arrived (see next section).

Table 4.6: Overview of warning time and warning distance to the first timing point or forced braking for the Intercity under dynamic timing point variant including performance data.

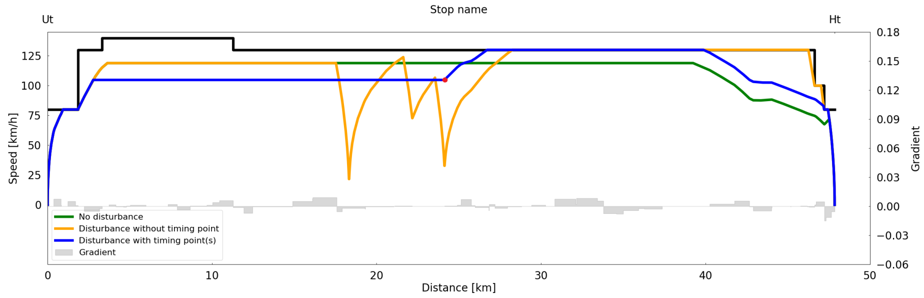
Spr scenario		**Warning time [s]**	**Warning distance [km]**	**Arrival delay [s]**	**Energy [kWh]**	
Htnc	1 min	754	23.252	+0	313	+0 (+0%)
	2 min	754	21.897	+0	313	+0 (+0%)
	3 min	754	20.089	+22 (+1.4%)	375	+62 (+19.8%)
Cl	1 min	405	10.180	+0	313	+0 (+0%)
	2 min	139	8.199	+0	343	+20 (+6.6%)
	3 min	0	0	+22 (+1.4%)	447	+134 (+42.8%)
	4 min	0	0	+82 (+5.0%)	447	+134 (+42.8%)
	5 min	0	0	+142 (+8.7%)	447	+134 (+42.8%)
Zbm	1 min	92	1.998	+0	354	+41 (+13.1%)
	2 min	0	0	+39 (+2.4%)	421	+108 (+34.5%)
	3 min	0	0	+99 (+6.1%)	421	+108 (+34.5%)
	4 min	0	0	+159 (+9.8%)	421	+108 (+34.5%)
	5 min	0	0	+219 (+13.5%)	421	+108 (+34.5%)

Increase in arrival delay and energy consumption of the Intercity is smallest when there is sufficient warning time and distance. Note that warning time and distance alone do not guarantee performance. This is visible when comparing the scenario with a 3-minute delay in Houten Castellum with a 1-minute delay in Culemborg. The former scenario experiences arrival delay and increased energy consumption despite a warning time of 12 minutes and 34 seconds and a warning distance of over 20 km, while the latter has optimal performance with 6 minutes and 45 seconds and just over 10 km before the first timing point. This is because the larger delay of the Sprinter forces the Intercity to pass the timing point later, consuming more than the complete running time supplement before the timing point. The Intercity is forced to take the more energy-consuming MTTC trajectory after the timing point to minimise the arrival delay.

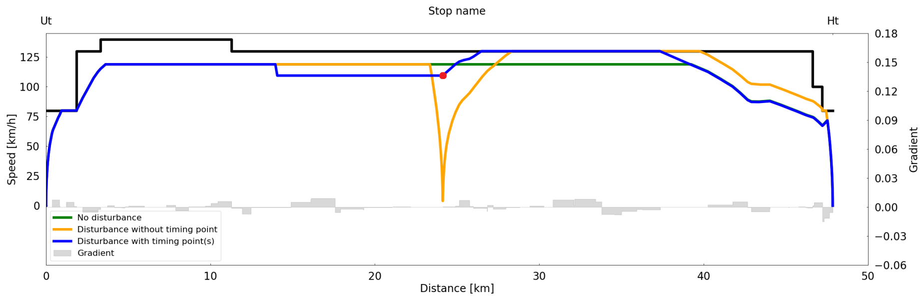
The scenarios for 1 and 2 minute delay at Culemborg and Houten Castellum result in the same conflict around Culemborg, but different solutions for the timing point approach, see the speed profiles of the uncontrolled scenarios in orange in Figure 4.9. Comparing the results of these scenarios shows the benefits of early warning.



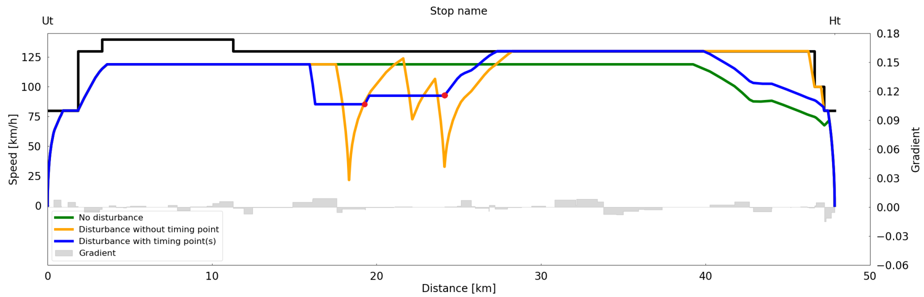
(a) 1 minute delay at Houten Castellum



(b) 2 minute delay at Houten Castellum



(c) 1 minute delay at Culemborg



(d) 2 minute delay at Culemborg

Figure 4.9: Intercity speed profiles for 1 and 2 minutes at Houten Castellum and Culemborg

Comparing the 1-minute scenarios with dynamic timing point (blue curves on the left) shows that the same timing point (red dot) can be respected with higher speed when the delay occurred at Houten Castellum. The delayed Sprinter departs earlier from this station as it lies more towards the start of the section. The departure headway at Ut between the Sprinter and the Intercity is 13 minutes and the departure of the Intercity at Ut is planned in the same minute as the departure of the Sprinter at Htnc. In the scenario where the Sprinter has a 1-minute delayed departure at Htnc, the Intercity gets a TPE update 1.5 minutes after departure, before it reaches its planned cruising speed. This results in an adapted lower cruising speed and no forced braking regime.
In contrast, when the delay of the Sprinter occurs in Culemborg, the following Intercity departs from Ut 8 minutes before the delayed departure of the Sprinter from Htnc. The Intercity receives the TPE update 8.5 minutes after its departure, at which time it is already at cruising speed and needs to brake. The speed reduction is necessarily bigger in order to clear the conflict. This results in additional acceleration and thus energy consumption.

The comparison of the 2-minute scenarios with dynamic point shows a more pronounced effect. A 2-minutes delay at Houten Castellum results in an updated TPE being communicated to the Intercity 2 minutes and 6 seconds after departure, while it is still accelerating. The accelerating can therefore be stopped at the reduced cruising speed, avoiding unnecessary initial acceleration.
When the 2-minute delay occurs at Culemborg the Intercity is already at the optimal cruising speed under the original EETC speed profile. As a result, the Intercity needs to brake to a lower cruising speed. Furthermore, an additional timing point is needed to avoid running into the Sprinter before the timing point of the largest blocking time overlap. This block with the largest overlap is the same for both delay locations (24.128 km), but the relatively late communication caused by the delay being located at Culemborg necessitates an additional timing point at 19.292 km.

4.3.2. Effect of a forced stop

The most significant increase in energy consumption is found for the scenarios with 0 warning time and distance. In these situations, the Intercity is forced to stop, and therefore a full re-acceleration is required. At the same time, the timing point benefits are limited by a forced stop of the following Intercity. The delay locations show a similar pattern with increasing delay: during small delays, the following train needs to brake but can continue without stopping, and at a certain delay, the following train is forced to brake to a complete stop.

A full stop can be recognised from the speed profile by braking to zero speed. The speed profile is limited to speed versus distance and thus does not show the dwell time. The train trajectory of this speed profile does include this information.
The scenarios with a full stop can furthermore be identified from Table 4.5: when the arrival delay for scenarios with more delay increases with the increase of the Sprinter departure delay, while there is no increase in energy consumption. See also the rows with zero warning time and distance in Table 4.6 and the speed profiles in Appendix B for confirmation.
In this situation, the Intercity stopped just behind the (still stationary) delayed Sprinter before being informed of the delay because the Sprinter would communicate its delay when departing. As a result, the speed profile does not change with additional delay. All additional delay leads to an equivalently longer unplanned stop.

The absolute benefits (in terms of energy consumption) of the timing point approach are constant in situations with a full stop. The benefits of the dynamic timing point approach compared to the uncontrolled variant are related to regulating the re-acceleration after the forced stop and avoiding multiple braking regimes. The trajectories of the Sprinter and subsequently of the Intercity shift back in time with increasing Sprinter delay but do not change shape, resulting in identical energy consumption.

4.3.3. Effect of avoiding forced braking regimes

The train trajectories for scenarios where one or more forced braking regimes were avoided show the largest percentage point difference between uncontrolled and timing point variant, see Table 4.7. The corresponding speed profiles can be found in the appendix (see 1 and 2 minutes in Culemborg in Figure B.4, Figure B.5 and 1, 2 and 3 minutes in Houten Castellum in Figure B.1, Figure B.2 and

Figure B.3).

Table 4.7: Overview of the number of avoided braking regimes for Intercity due to dynamic timing point(s), including resulting energy consumption and energy reduction in percentage points compared to the uncontrolled variant of the same scenario.

Spr scenario		# avoided braking regimes	Arrival delay [s]	Energy [kWh]		Energy red. [percentage point]
Htnc	1 min	1	+0	313	+0 (+0%)	26.2
	2 min	3	+0	313	+0 (+0%)	70.9
	3 min	3	+22 (+1.4%)	375	+62 (+19.8%)	55.9
Cl	1 min	1	+0	313	+0 (+0%)	26.2
	2 min	3	+0	343	+20 (+6.6%)	64.3
	3 min	2	+22 (+1.4%)	447	+134 (+42.8%)	32.9
	4 min	2	+82 (+5.0%)	447	+134 (+42.8%)	32.9
	5 min	2	+142 (+8.7%)	447	+134 (+42.8%)	32.9
Zbm	1 min	1	+0	354	+41 (+13.1%)	10.5
	2 min	0	+39 (+2.4%)	421	+108 (+34.5%)	20.8
	3 min	1	+99 (+6.1%)	421	+108 (+34.5%)	20.8
	4 min	1	+159 (+9.8%)	421	+108 (+34.5%)	20.8
	5 min	1	+219 (+13.5%)	421	+108 (+34.5%)	20.8

The decrease in energy consumption compared to the uncontrolled scenario is highly dependent on the extra energy needed in the uncontrolled scenario. The largest differences occur when the uncontrolled scenario runs into multiple forced braking regimes and therefore accelerates multiple times. For example, when three minutes delay occurs in Houten Castellum (see Figure 4.10 and the others in the appendix) where the uncontrolled scenario results in three forced braking actions (orange speed profile) while the speed profile shows one partial braking action under the dynamic timing point strategy (blue speed profile). Note that the Intercity receives the TPE update at 3 minutes and 6 seconds after departure, approximately 13 seconds after the optimal cruising speed of the original EETC profile. Therefore there is an initial braking phase to reach the reduced cruising speed in order to pass the timing point. The green speed profile shows the original undisturbed EETC speed profile.

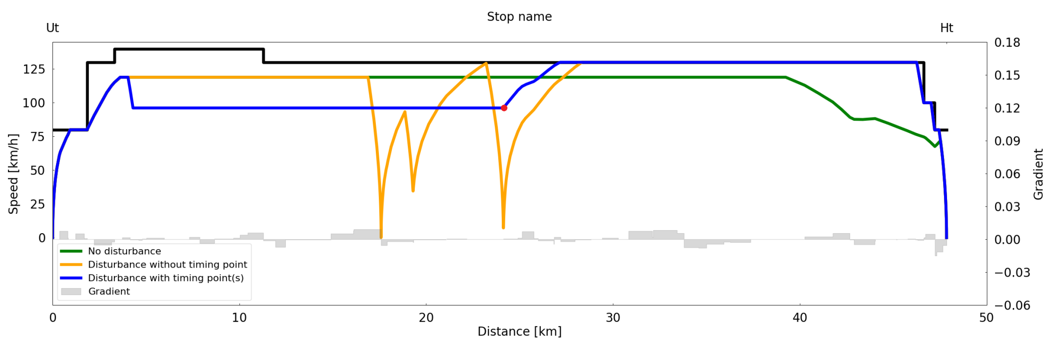


Figure 4.10: Intercity speed profiles of variants of the 3-minute delay of the Sprinter at Houten Castellum.

4.3.4. Effect of progress along route

The ability to save both delay and energy consumption is limited by the degree of remaining running time supplement that is left when the following train receives the updated TPE. The remaining running time supplement is generally reduced over the train run, and therefore also the difference between the time-distance paths of different strategies targeting the scheduled arrival time. As a result, the room for a dynamic timing point shrinks, decreasing the ability to guide the ATO-OB.

This is relevant in the case of Zaltbommel, which is located towards the end of the section between Utrecht and 's-Hertogenbosch. The following Intercity train has driven the EETC speed profile along the route, using a portion of the running time supplement for a decreased cruising speed. The ability for recovery in terms of punctuality and energy consumption is therefore limited. The challenge to handle disturbances at Zaltbommel is compounded by the small planned buffer at this location, resulting in a lack of warning time.

4.4. Intermediate TPE updates through communication during delay accumulation

When the delay of the Sprinter increases, there will be an increase in blocking time overlap, forcing the Intercity into additional braking to avoid the conflict. Additionally, the communication of this delay moves forward in time with the delay itself, resulting in decreasing warning time and distance for the Intercity to adapt its trajectory.
This section, therefore, reports on the situation where the Sprinter sends a status report every minute with an update on the lower bound of its departure delay. This methodology is applied to the scenarios of 2- and 3-minute departure delay at Culemborg, because these scenarios show a significant improvement of energy consumption with respect to the uncontrolled variant (64.3 and 32.9 percentage points respectively). The 2-minute departure delay at Culemborg has a warning time of 139 seconds and avoids a full stop using the dynamic timing point approach, while the 3-minute departure delay at Culemborg forces the Intercity to come to a full stop. The 2- and 3-minute delay scenarios at Culemborg thus form a suitable selection for this additional experiment.

In this experiment, the TPE is adapted every minute. Subsequently, the following Intercity changes its trajectory at every update.
The updating frequency is set to 1 minute to reflect the balance between riding comfort and benefits in terms of energy saving. A higher updating frequency could result in bigger energy savings, but frequent changes in speed could be uncomfortable for passengers, especially when braking is used as opposed to coasting to reduce speed.

Figure 4.11 shows the resulting Intercity speed profile for a 2-minute delay of the Sprinter at Culemborg. The Intercity receives two TPE updates during the journey. The first TPE update (after 1 minute of accumulated delay) results in a reduced cruising speed, see the first step down in the graph. The second TPE update (after the delayed departure of the Sprinter) includes two timing points. Additional reduction of the cruising speed is required, resulting in a second step down in the graph. After the first timing point, a third section of reduced cruising speed is needed up until the second timing point.

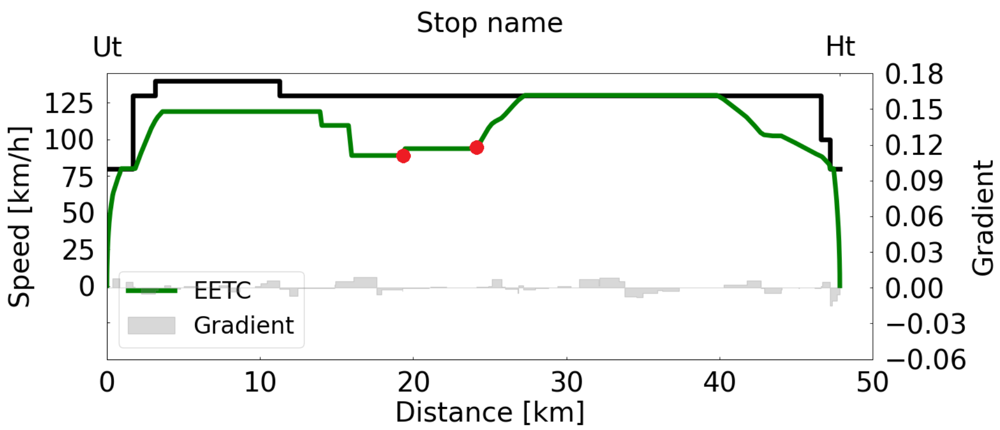


Figure 4.11: Final speed profile for Intercity following a Sprinter with a 2-minute delay from Culemborg with an intermediate warning at 1 minute of accumulated delay. Timing points are indicated with red dots.

The energy and time savings are summarised in Table 4.8.

Table 4.8: Overview of the arrival delay and energy consumption comparing dynamic timing points without and with intermediate updates, compared to the on-time scenario.

	Without updates		**With updates**	
	Arrival delay [s]	**E [kWh)]**	**Arrival delay [s]**	**E [kWh]**
on time	0	313	0	313
1 min	+0 (+0%)	313 (+0, +0%)	+0 (+0%)	313 (+0, +0%)
2 min	+0 (+0%)	343 (+10, +3.2%)	+0 (+0%)	341 (+8, +2.6%)
3 min	+21 (+1.3%)	447 (+134, +42.8%)	+21 (+1.3%)	427 (+114, +36.4%)

The results show that there is an increase in benefits in terms of energy consumption for these scenarios. For a 2-minute delay with an intermediate update at 1 minute, the energy consumption decreases with 0.6 percentage points while energy consumption during the 3-minute delay decreases with 6.4 percentage points due to the updates per minute.

4.5. Sensitivity analysis: effect of TPE processing time

The processing time from the delayed departure of the Sprinter to the first adapted control in the Intercity could have significant effects on the outcome. The experiments so far have assumed a processing time of 30 seconds. This section focuses on the effects of varying the TPE processing time between 15 and 60 seconds. This range is a reasonable sample of possible processing times without failures and enable the comparison of different potential standards.

Table 4.9 presents the arrival delay and energy consumption of the Intercity under different TPE processing times using the 2-minute delay scenario for the Sprinter at Culemborg. This location is chosen for the significant interaction between the Intercity and the Sprinter.

Table 4.9: Sensitivity analysis of processing times for a 2-minute delay at Culemborg with energy savings compared to the assumption of 30 seconds.

Processing time [s]	Timing Point	Braking Location	Braking to [km/h]	Arrival	Energy [kWh]	Energy Change [%]
	No	18.324 km 22.172 km 24.128 km	21.6 73.1 32.4	on time	535	
15	Yes	from 15.398 km	91.1	on time	340	-0.9
30	Yes	from 15.896 km	90.0	on time	343	0
45	Yes	from 16.389 km	88.6	on time	351	+2.3
60	Yes	from 16.884 km	87.5	on time	359	+4.7

Decreasing the TPE processing time results in earlier braking and therefore less speed reduction. With an increase in processing time, the opposite happens: later braking forces more speed reduction, and therefore more re-acceleration is needed.
As a result, energy usage goes up with increasing processing time. In this example, this is especially clear when the processing time increases compared with the 30-second standard.

Based on this experiment, the additional energy savings due to improving the processing time from 15 to 30 seconds are limited at 0.9% while the additional energy consumption due to increasing processing times to 45 and 60 is more significant with 2.3% and 4.7% respectively.

More research is needed to gain insight into the impact of even longer processing times on punctuality and energy consumption, as it could vary with the situation, especially when the increased processing time forces the following train to brake due to MA not being extended.

5 Conclusions

This thesis formulated a dynamic timing point methodology to manage minor delays on a corridor and quantified its benefits. Finally, this chapter answered the research questions as formulated in the introduction, followed by suggestions for further research.

5.1. The potential of dynamic timing points during small disturbances under ATO-over-ETCS

The research in this thesis focused on developing and testing a dynamic timing points approach to manage disturbances under ATO-over-ETCS.

> **Sub question 1:** What interaction mechanism between TMS and ATO enables dynamic timing point updates to improve performance during disturbances?

The existing literature focusing on the link between TMS and ATO during disturbances is limited. Real-time traffic management and energy-efficient train operations are studied in depth but the dynamic determination of timing points remains under-explored.

Dynamic timing points are locations on the route where a time window of passage is inserted during operations. This restricts the train trajectory to a narrower range of options which can improve the traffic flow by avoiding conflicts. This thesis considers situations where a delayed train sends one or more delay update(s) to trackside intelligence which translates this into an update to the Train Path Envelope (TPE) for the following train containing timing points with corresponding timing point windows. The Real-Time Traffic Plan (RTTP) with arrival, departure and passing times at timetable points such as stations remains unchanged because the TMS does not get involved.

The proposed approach for dynamic timing points is described in Chapter 3. The process is built to fit in the context of ATO-over-ETCS and relies on proactive communication of delay, in the example of this thesis, a departure delay. The delay information is used by the trackside intelligence to predict conflicts between the delayed train and the train directly following it, and calculate one or more dynamic timing points using the concept of blocking time overlap. The timing point locations are positioned at where the blocking time overlap is maximal, while the timing point windows are computed based on the (expected) end of the block occupation by the delayed train. The dynamic timing points are added to the TPE to form an updated TPE that is sent to the following train which subsequently adapts its trajectory and avoids unnecessary braking regimes. The updated TPE includes the dynamic timing point(s) that limit the train to pass the location in a restricted time window. This additional intelligence provides value in situations with small disturbances where there is no need to update the RTTP.

The trackside intelligence in charge of generating the updated TPE is an essential part of the process and is currently not specified. This system needs to have information about the traffic state (train positions and speeds) and RTTP and be able to calculate trajectories on a corridor level for an accurate

calculation. Ideally, the train characteristics and environmental variables should also be known. It is possible to calculate dynamic timing points with estimated parameters but the result might not solve the conflict(s) if the assumed train behaviour is too different from reality. Uncertainties and communication delays should be minimised as much as possible in order to maximise the punctuality and energy consumption benefits.
The ATO-TS subsystem could be developed into a suitable location for the dynamic timing point intelligence that responds quickly due to short feedback loops. It is already conceptualised as a subsystem to translate between the network-level operational planning (RTTP) and train operation, which makes it a suitable approach to optimise traffic at the corridor level. Another possibility is to implement dynamic timing point intelligence in the TMS, with ATO-TS as a server to receive and forward the TPE to the trains. This requires extensive development of the TMS for more detailed calculation at the corridor and train level.

Sub question 2: Which operational circumstances potentially benefit from dynamic timing points?

The proposed dynamic timing point approach could be used in operational circumstances with a disturbance that does not trigger an RTTP update. In responding to disturbances, the dynamic timing point approach can improve performance (punctuality and energy consumption) by reducing forced braking regimes. The performance can be improved if the warning time and distance to the closest timing point are sufficient to reduce forced braking. Furthermore, there is potential for (further) performance improvement if the following train re-accelerates quicker than the delayed train. In this case, the re-acceleration of the following train is guided by the timing points to avoid forced braking.

The example studied in this thesis is a faster train following a slower train on a shared unidirectional track. The scenarios that were studied include departure delays of the slower train ranging from 1 to 5 minutes at intermediate stations. In all situations, the punctuality and energy consumption of the following train improved due to the dynamic timing point compared to the uncontrolled case with the same delay (see also later in this section).

The case study showed that dynamic timing points could benefit situations where a conflict occurs due to a small delay of up to approximately 3 minutes. With a less dense timetable, the allowable delay could be longer. The power of dynamic timing points lies in avoiding conflicts by restricting the given TPE as resulting from the originally conflict-free RTTP during disturbances. Based on the case study, this is likely to give results in dense timetables with mixed traffic.

The experiments with more frequent or quicker and slower communication show that reliable communication channels are required to fully use the potential of the proposed dynamic timing point approach. The performance increases with intermediate TPE updates while delay accumulates and when processing time decreases. Operational circumstances where communication is reduced will thus decrease the potential for increased performance.

Sub question 3: How can the proposed dynamic timing point approach be modelled and implemented computationally to evaluate its impact on train operation?

A computational TPE generation model was implemented with components of trajectory (re-)optimisation based on the optimal control structure as obtained by Pontryagin's Maximum Principle and the Karush-Kuhn-Tucker conditions. Blocking time theory was used to accurately model blocking time conflicts between two trains following each other.
The trajectories of the following faster train are restricted to their original form until the location where the train adapts its speed to react to the updated TPE. This ensures that the modelling of the trajectory before the update stays consistent with real-time updates during operation.

Two versions of the model are developed: one to study the uncontrolled response of a train to a delayed train in front of it, and another version to implement the proposed dynamic timing points approach. The arrival delay, energy consumption and number of forced braking regimes of the resulting train trajectories are calculated to determine the performance of the response.

The **uncontrolled response** is implemented by restricting the speed limit of the following train for a short distance at the start of the first conflicting block. The size of the speed reduction at this location is optimised to be just enough to force the trajectory of the following train out of the corresponding conflict with the delayed train. If a complete stop is required, the size of the dwell time is optimised to be just enough to resolve the conflict.

The **dynamic timing point response** is implemented by determining the critical block with the biggest blocking time overlap between the delayed and following train. The timing point location is the beginning of the critical block while the timing point window starts as the delayed train leaves this block. Subsequently, the trajectory response to this dynamic timing point is calculated to determine if additional timing points are needed. The traffic state (location and speed) of the following train when it receives the updated TPE are determined. From this point, the following train brakes (or accelerates) to a speed that is optimised to be just sufficiently low to pass the timing point within the timing point window. The blocking time stairway of the resulting trajectory is used to check for remaining conflicts which are solved with one or more additional timing points according to the same process as described above.

Sub question 4: To what extent can dynamic timing point updates reduce energy consumption and improve punctuality in ATO-over-ETCS operations?

The dynamic timing point concept was studied on the 48-kilometre-long corridor from Utrecht to s'-Hertogenbosch, where a fast intercity train is running into a delayed slower Sprinter train.
Small disturbances (1 and 2 minutes at Houten Castellum and Culemborg, 1 minute in Zaltbommel) did not result in arrival delays for the following intercity train, both with and without the dynamic timing point strategy, indicating that there is no update to the RTTP required. However, the improvement in energy consumption due to the dynamic timing point(s) in these scenarios ranged from 10.5 to 70.9 percentage points, compared to the uncontrolled variant of the scenario where the Intercity is forced to brake due to the conflict(s).
The benefits of dynamic timing points are smaller for larger delays at Culemborg and Zaltbommel because these have already forced the following Intercity to come to a full stop behind the delayed, stationary Sprinter train. The dynamic timing point benefits that remain in these situations are due to a better re-acceleration regulation, preventing the faster Intercity from running into the Sprinter train.

Intermediate updates of the delay accumulation of the Sprinter departure resulted in additional energy savings. For this experiment, the Sprinter updates its continuing accumulation of delay every minute it remains stationary at the stop. In this case study, there is additional saving in energy consumption of 0.6 and 6.4 percentage points for 2 and 3 minutes total delay, respectively, when the delay occurs in Culemborg.
A fast processing time between the departure of the Sprinter and the trajectory adaptation of the following Intercity can provide additional benefits. Decreasing the processing time from 30 to 15 seconds results in a 0.9% decrease in energy use for the 2-minute delay scenario at Culemborg. Increasing the processing time to 45 and 60 seconds increased the energy use by 2.3% and 4.7% respectively, compared to the 30-second assumption used in the rest of the experiments. It is therefore important to aim for the processing time to be as small as possible while maximising reliability in communication.

Main research question: What is the potential of dynamic timing points to improve ATO-over-ETCS operations during disturbances?

Dynamic timing points present a powerful enhancement to ATO-over-ETCS operations, addressing a pervasive yet operationally under-managed issue in mainline rail systems: small disturbances that do not warrant a full RTTP update, yet degrade performance significantly if unaddressed.
The thesis focuses on dense mixed-traffic mainline corridors (particularly those with fast trains following

slower ones) where minor delays in trains cause avoidable braking and re-acceleration for following trains. These reactions increase energy consumption and can reduce punctuality.

The proposed approach shows that dynamic timing points, inserted by a trackside system such as ATO-TS in response to real-time delay updates, can effectively adjust the TPE for the following train. This allows ATO-OB to re-optimise the speed profile, converting reactive braking into gradual speed adaptations that reduce energy consumption and improve punctuality.
The proposed approach aligns with the modular architecture of ATO-over-ETCS and can be efficiently implemented at the corridor level within the ATO-TS subsystem, avoiding the computational complexity and communication burden associated with centralised TMS recalculations.

The computer model developed and applied to the case study of the Utrecht - 's-Hertogenbosch corridor shows that the proposed dynamic timing point approach is promising to increase punctuality and decrease energy consumption during disturbances. The approach shows improvements across a range of delay scenarios on the case study corridor, pointing towards broader applicability.
The performance benefits of dynamic timing points are sensitive to communication latency, underscoring that early communication, real-time localisation, and fast processing are not only enablers but essential to unlocking the full potential of dynamic timing points.

5.2. Suggestions for further research

The dynamic timing point concept is a promising research direction for future research in general, as it has clear energy- and time-saving potential based on the case study. The case study outcomes are, nevertheless, limited to the case study situations, with no definitive representativeness claims for the wider railway system. This section, therefore, describes topics to be researched further.

5.2.1. Insight in beneficial situations

This thesis focused on calculating the benefits of the dynamic timing point methodology on one corridor. The results make it clear that this case study significantly benefits from dynamic timing points during small delays. This insight is limited to the specific infrastructure, trains and timetable used in the calculations.
The case study corridor is typical for high-density routes in the Netherlands, but the average route in Europe has a less densely packed timetable.

The code developed for this study can be adapted to handle different case studies, for example, with different timetables and infrastructures. With the use of simulation software that can model the train trajectories and automatically insert timing points, it becomes easier to study the dynamic timing point methodology in various circumstances, for instance, multiple delays following each other in the same scenario. Potential topics to clarify include the impact of timetable density, interacting delays, infrastructure complexity, and delays directly affecting opposing, diverging and/or converging train routes. The impact of train speeds is also an interesting topic of study, especially for railway networks like Germany, where high-speed trains share tracks with slower trains on certain routes.

5.2.2. Effect on other trains

This thesis focused on the time and energy savings for the train directly following a delayed train. Depending on the timetable density, the trains that traverse the track after the first following train can also be impacted.
The timing point strategy reduced the delay of the first following train in the case study of this thesis, indicating that the effect on the trains following afterwards would likely be positive. The extent of the expected positive effect is to be determined, as well as any situations where the timing point strategy could result in negative impacts. Considering additional trains could reveal a negative impact of the trajectory on the first following train to subsequent trains, potentially leading to sub-optimality of the dynamic timing point that was functional for a two-train environment.

5.2.3. Most suitable dynamic trajectory re-optimisation methods

The model used to calculate the benefits of dynamic timing points in the case study scenarios is built upon the strategy where the following train brakes directly to reach the reduced cruising speed and does not optimise the passing speed. This strategy is reasonable for this first research, but has not been proven to be the optimal re-optimisation method.

Based on this research, it can be argued that it is reasonable to use immediate braking to a reduced cruising speed that is held towards the dynamic timing point, as this keeps the options open in reacting to further delays. Adding the coasting regime would widen the set of possible trajectories, making dynamic timing point calculation more complex but potentially increasing performance. The dynamic timing point approach can thus be improved by including the option to coast instead of brake to the reduced cruising speed.

Two examples where the minimum acceleration strategy does not necessarily give optimal results include:

- When the timing point and corresponding time window are located in such a way that the train cannot be on time. To minimise the delay, the train needs to optimise the speed at which it passes the timing point: slowing down before the timing point and passing it at a higher speed than following the minimum acceleration strategy. This results in increased re-acceleration and higher energy consumption than the figures calculated in this thesis, but a smaller delay.
  The option of passing the timing point at higher speeds has not been explored in this thesis. This could also give additional benefits because more running time supplement will be available for coasting after the timing point.
- Dynamic timing points located far enough and with a small enough time shift compared to the original train-path do not necessitate braking and could instead use coasting to reduce speed. Coasting uses a significantly longer distance for the same speed reduction, thus reducing the required distance to cover at reduced cruising speed. As a result, the energy consumption decreases.

Future research could explore the best trajectory re-optimisation method for each situation. This would primarily be part of research on the optimal ATO-OB software but also influences the TPE calculation that guides the ATO-OB.

Note that the need to prioritise punctuality over energy use could be questioned in cases where the delay is not critical and does not influence following trains or connections for passengers. If this situation is known to the railway operator, they could choose to optimise on energy use only or only aim for punctuality up until the point where conflicts occur.

5.2.4. The effect of information uncertainty regarding onboard dynamic trajectory re-optimisation strategy

When calculating the timing points, the trackside intelligence knows the train location, train speed, JP, SP, train characteristics and environmental variables, and the train is known to follow the EETC trajectory. The regime choices of the trajectory re-optimisation method used by the ATO-OB are not necessarily known. However, the trajectory re-calculation is also performed by the trackside intelligence to check if additional dynamic timing points are needed. If the ATO-OB strategy for this calculation is unknown, assumptions need to be made.

The existence of multiple options for the trajectory towards the dynamic timing point results in challenges for the trackside intelligence to accurately calculate the required timing points in the TPE updates. Based on the current system architecture, the trackside intelligence needs to calculate the future traffic state without knowing which optimisation objective the ATO-OB chose. The trackside intelligence is informed of the estimated time of arrival at the upcoming timing points but not the intermediate trajectory. The IM could argue for standardisation of trajectory re-calculation in response to TPE updates as this would improve the TPE generation mechanism, but it does limit the freedom of RU in choosing driving strategies that are appropriate for their objective.

A potential topic to explore is the impact of these different potential trajectories on the effectiveness of the dynamic timing points that are inserted.

5.2.5. Effect of train-specific characteristics and variations in external factors

The TPE is ideally based on accurate train-specific characteristics and an accurate perception of the relevant external factors. The train-specific characteristics are also used in the calculation of dynamic timing points. The effect of an inaccuracy is unknown and could be important. A large enough inaccuracy could result in the conflict remaining to exist instead of being solved. The current standards also lack feedback on the parameters. More research is needed to understand the risk of inaccuracies and to what extent this would impact the computation.

One approach for future research on this topic could be the implementation of a tracking buffer that compensates for the variation in tracking accuracy of the ATO algorithm, see for example FP1 MOTIONAL (2024b).

5.2.6. Communication between trackside and trains

The model used in this thesis assumes that the full cycle from the delayed departure of the first train to the trajectory adaptation of the following train occurs in a set time of 30 seconds. The sensitivity analysis of section 4.5 furthermore shows the impact of varying the TPE processing time of the system.
The interaction between the effect of communication timelines and delay scenarios can be studied further. Furthermore, it is essential to understand the effect of failures in communication in terms of process and resulting effects on punctuality and energy consumption.

The initial analysis of using intermediate updates of the delay while it occurs shows that it is possible to achieve additional energy savings. More research is needed to find the optimal frequency of intermediate updates, for example, increasing this to every 10 seconds or similar. The benefits in terms of energy consumption and punctuality need to be weighed against the communication load and computational effort, and the comfort of the passengers, because a large number of speed changes is not optimal for the passenger experience.

These open questions on communication can be studied further in a simulation environment.

5.2.7. Connected DAS versus ATO: impact of human factors

The dynamic timing point methodology can be applied both as driver advice and in ATO. In the case of implementation through Connected DAS instead of ATO (as assumed here), there are human factors like reaction time and compliance to consider.
When implementing dynamic timing points through Connected DAS, the driving advice needs to be derived from the trajectory, which might bring in extra uncertainties and require more buffer in the timetable. The frequency of updates to driver advice should be limited in order to keep the driver workload feasible, which limits the ability to react as the situation develops.

Bibliography

Abderrahman Ait Ali and Jonas Eliasson. European railway deregulation: an overview of market organization and capacity allocation. *Transportmetrica A: Transport Science*, 18(3):594–618, 2022.

Amie Albrecht, Phil Howlett, Peter Pudney, Xuan Vu, and Peng Zhou. The key principles of optimal train control — part 2: Existence of an optimal strategy, the local energy minimization principle, uniqueness, computational techniques. *Transportation Research Part B: Methodological*, 94:509–538, 2016.

Patrick Bochmann and Birgit Jaekel. Measures and methods for the evaluation of ATO algorithms. *Applied Sciences*, 12(9):4570, 2022.

Francesco Corman and Lingyun Meng. A review of online dynamic models and algorithms for railway traffic management. *IEEE Transactions on Intelligent Transportation Systems*, 16(3):1274–1284, 2015.

Francesco Corman, Andrea D'Ariano, Alessio D Marra, Dario Pacciarelli, and Marcella Samà. Integrating train scheduling and delay management in real-time railway traffic control. *Transportation Research Part E: Logistics and Transportation Review*, 105:213–239, 2017.

Alex Cunillera, Harm H. Jonker, Gerben M. Scheepmaker, Wilbert H.T.J. Bogers, and Rob M.P Goverde. Coasting advice based on the analytical solution of the train motion model. *Journal of Rail Transport Planning & Management*, 28:100412, 2023.

Leo D'Amato, Federico Naldini, Valentina Tibaldo, Vito Trianni, and Paola Pellegrini. Towards self-organizing railway traffic management: concept and framework. *Journal of Rail Transport Planning & Management*, 29:100427, 2024.

Andrea D'Ariano, Dario Pacciarelli, and Marco Pranzo. A branch and bound algorithm for scheduling trains in a railway network. *European journal of operational research*, 183(2):643–657, 2007.

ERA and UNISIG and EEIG ERTMS Users Group. Subset 125 - ERTMS/ATO-System Requirements Specification, July 2023. version 1.0.0.

ERTMS User Group, EULYNX consortium. Reference CCS Architecture (RCA). In *Version 1.0*, 2022.

Wei Fang, Shengxiang Yang, and Xin Yao. A survey on problem models and solution approaches to rescheduling in railway networks. *IEEE Transactions on Intelligent Transportation Systems*, 16(6): 2997–3016, 2015.

Adrian Fernandez-Rodriguez, Asuncion P Cucala, and Antonio Fernandez-Cardador. An eco-driving algorithm for interoperable automatic train operation. *Applied Sciences*, 10(21):7705, 2020.

Adrian Fernandez-Rodriguez, Shuai Su, Antonio Fernandez-Cardador, Asuncion P Cucala, and Yuan Cao. A multi-objective algorithm for train driving energy reduction with multiple time targets. *Engineering Optimization*, 53(4):719–734, 2021.

FP1 MOTIONAL. Deliverable D15.1 - Requirements for the deployment of TMS linked with ATO/C-DAS, 2024a. version 2.0.

FP1 MOTIONAL. Deliverable D15.2 - TMS and ATO/C-DAS Timetable Test and Simulation Environment, 2024b. version 1.0.

Nadjla Ghaemi and Rob MP Goverde. Review of railway disruption management practice and literature. In *6th international conference on railway operations modelling and analysis - RailTokyo2015*. Citeseer, 2015.

A. González-Gil, R. Palacin, P. Batty, and J. P. Powell. A systems approach to reduce urban rail energy consumption. *Energy Conversion and Management*, 80:509–524, 2014.

Rob M.P. Goverde, Gerben M. Scheepmaker, and Pengling Wang. Pseudospectral optimal train control. *European Journal of Operational Research*, 292:353–375, 2021.

Jørgen Thorlund Haahr, David Pisinger, and Mohammad Sabbaghian. A dynamic programming approach for optimizing train speed profiles with speed restrictions and passage points. *Transportation Research Part B: Methodological*, 99:167–182, 2017.

Phil Howlett. The optimal control of a train. *Annals of Operations Research*, 98:65–87, 2000.

Phil Howlett, Peter Pudney, and Amie Albrecht. Optimal driving strategies for a fleet of trains on level track with prescribed intermediate signal times and safe separation. *Transportation science*, 57(2): 399–423, 2023.

David Hunter and Quinn Bryant. Automated train operation interface requirements specification development summary report. Technical report, U.S. Department of Transportation, Federal Railroad Administration, 2023. URL https://rosap.ntl.bts.gov/view/dot/71896. Report Number : DOT/FRA/ORD-23/31.

International Union of Railways (UIC). IRS 90940 - Ed. 2: Digitalisation, Data, Emerging Innovations-Exchange of Data-Data exchange with Driver Advisory Systems (DAS) following the SFERA protocol, 2022.

Lars Wittrup Jensen. *Robustness indicators and capacity models for railway networks*. PhD thesis, Technical University of Denmark, December 2015.

SW Kim, HB Kwon, YG Kim, and TW Park. Calculation of resistance to motion of a high-speed train using acceleration measurements in irregular coasting conditions. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 220(4):449–459, 2006.

European Commission Mobility and Transport. ETCS Levels and Modes, 2023. URL https://transport.ec.europa.eu/transport-modes/rail/ertms/what-ertms-and-how-does-it-work/etcs-levels-and-modes_en. Visited 2024-11-19.

Carl-William Palmqvist, Norio Tomii, and Yasufumi Ochiai. Explaining dwell time delays with passenger counts for some commuter trains in Stockholm and Tokyo. *Journal of Rail Transport Planning & Management*, 14:100189, 2020.

Paola Pellegrini, Grégory Marlière, Raffaele Pesenti, and Joaquin Rodriguez. 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, 2015.

P. Pudney, P. Howlett, A. Albrecht, D. Coleman, X. Vu, and J. Koelewijn. *Optimal driving strategies with intermediate timing points*. PhD thesis, International Heavy Haul Association, 2011.

Egidio Quaglietta, Paola Pellegrini, Rob M.P. Goverde, Thomas Albrecht, Birgit Jaekel, Gregory Marliere, Joaquin Rodriguez, Twan Dollevoet, Bruno Ambrogio, Daniele Carcasole, Marco Giaroli, and Gemma Nicholson. The ON-TIME real-time railway traffic management framework: a proof-of-concept using scalable standardized data communication architecture. *Transportation Research Part C*, 63:23–50, 2016.

Xiaolu Rao, Markus Montigel, and Ulrich Weidmann. A new rail optimisation model by integration of traffic management and train automation. *Transportation Research Part C: Emerging Technologies*, 71:382–405, 2016.

Gerben M. Scheepmaker and Rob M.P. Goverde. Multi-objective railway timetabling including energy-efficient train trajectory optimization. *European Journal of Transport and Infrastructure Research*, 21: 1–42, 2021.

Gerben M Scheepmaker, Rob MP Goverde, and Leo G Kroon. Review of energy-efficient train control and timetabling. *European Journal of Operational Research*, 257(2):355–376, 2017.

Gerben M. Scheepmaker, Helen Y. Willeboordse, Jan H. Hoogenraad, Ralph S. Luijt, and Rob M.P. Goverde. Comparing train driving strategies on multiple key performance indicators. *Journal of Rail Transport Planning & Management*, 13:100163, 2020.

Artiya Sopharak, Tosaphol Ratniyomchai, and Thanatchai Kulworawanichpong. Energy saving study of mass rapid transit by optimal train coasting operation. In *Proceedings of the 2020 International Conference on Power, Energy and Innovations, ICPEI 2020*, pages 25–28. Institute of Electrical and Electronics Engineers Inc., 2020.

Shuai Su, Xuekai Wang, Yuan Cao, and Jiateng Yin. An energy-efficient train operation approach by integrating the metro timetabling and eco-driving. *IEEE Transactions on Intelligent Transportation Systems*, 21(10):4252–4268, 2019.

Shuai Su, Zhongbei Tian, and Rob M. P. Goverde. *Energy-Efficient Train Operation*. Springer International Publishing, 2023. ISBN 978-3-031-34655-2. doi: 10.1007/978-3-031-34656-9. URL https://link.springer.com/10.1007/978-3-031-34656-9.

Simon Tschirner, Bengt Sandblad, and Arne W Andersson. Solutions to the problem of inconsistent plans in railway traffic operation. *Journal of rail transport planning & management*, 4(4):87–97, 2014.

UNISIG. UNISIG factsheet, 2014. URL https://www.ertms.net/wp-content/uploads/2021/07/ERTMS_Factsheet_8_UNISIG.pdf. Visited 2024-08-01.

Pengling Wang and Rob M. P. Goverde. Multiple-phase train trajectory optimization with signalling and operational constraints. *Transportation Research Part C: Emerging Technologies*, 69:255–275, 2016.

Pengling Wang and Rob M.P. Goverde. Multi-train trajectory optimization for energy-efficient timetabling. *European Journal of Operational Research*, 272(2):621–635, 2019.

Yihui Wang, Bart De Schutter, Ton J.J. van den Boom, and Bin Ning. Optimal trajectory planning for trains - a pseudospectral method and a mixed integer linear programming approach. *Transportation Research Part C: Emerging Technologies*, 29:97–114, 2013.

Ziyulong Wang, Egidio Quaglietta, Maarten G.P. Bartholomeus, and Rob M.P. Goverde. Assessment of architectures for Automatic Train Operation driving functions. *Journal of Rail Transport Planning & Management*, 24:100352, 2022.

Ziyulong Wang, Egidio Quaglietta, Maarten Bartholomeus, Alex Cunillera, and Rob Goverde. Conflict-free train path planning using ato timing points. In *RailBelgrade 2023: 10th International Conference on Railway Operations Modelling and Analysis*, 2023.

Ziyulong Wang, Egidio Quaglietta, Maarten G.P. Bartholomeus, Alex Cunillera, and Rob M.P. Goverde. Optimising timing points for effective automatic train operation. *Computers & Industrial Engineering*, page 111237, 2025a.

Ziyulong Wang, Egidio Quaglietta, Maarten G.P. Bartholomeus, and Rob M.P. Goverde. Sensitivity of automatic train operation timing points. In final review at IEEE Transactions on Intelligent Transportation Systems, 2025b.

Jiateng Yin, Tao Tang, Lixing Yang, Jing Xun, Yeran Huang, and Ziyou Gao. Research and development of automatic train operation for railway transportation systems: A survey. *Transportation Research Part C: Emerging Technologies*, 85:548–572, 2017.

Sven Zeegers. Sporenplanonline, 2024. URL https://www.sporenplan.nl/. Visited 2024-12-23.

Hao Zhou, Yiming Wan, Hao Ye, Borui Li, and Baoming Liu. A novel real-time algorithm for optimizing train speed profiles under complex constraints. *IEEE Transactions on Intelligent Transportation Systems*, 24(8):7987–8002, 2023.

A

Timetable

Table A.1: Case study timetable in seconds

		Sp 1	Ic 1	Sp 2	Ic 2	Sp 3	Ic 3	Sp 4	Ic 4
Ut	**d**	000	180	660	840	1320	1440	1800	1980
Utvr	**a**	144	|	804	|	1464	|	1944	|
	d	180	|	840	|	1500	|	1980	|
Utln	**a**	306	|	966	|	1626	|	2106	|
	d	342	|	1002	|	1662	|	2142	|
Htn	**a**	528	|	1188	|	1848	|	2328	|
	d	564	|	1224	|	1884	|	2364	|
Htnc	**a**	696	|	1356	|	2016	|	2496	|
	d		|	1416	|	2076	|		|
Cl	**a**		|	1794	|	2454	|		|
	d		|	1836	|	2496	|		|
Gdm	**a**		|	2184	|	2844	|		|
	d		|	2514	|		|		|
Zbm	**a**		|	2898	|		|		|
	d		|	2940	|		|		|
Ht	**a**		1800	2444	2460		3060		3600

B

Overview of all speed profiles

B.1. Houten Castellum

1 minute delay

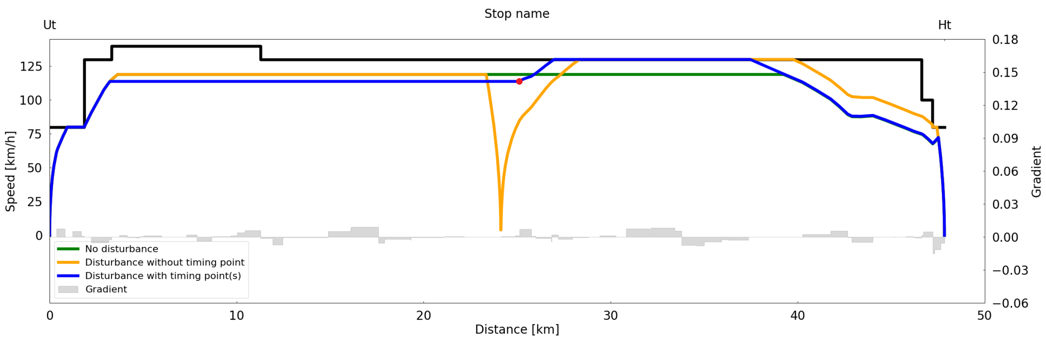


Figure B.1: Intercity speed profiles of scenarios around a 1 minute delay of the Sprinter at Houten Castellum

2 minute delay

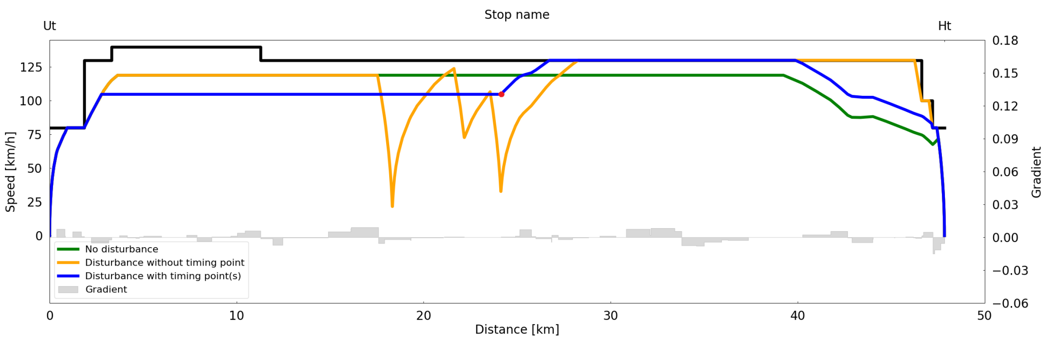


Figure B.2: Intercity speed profiles of scenarios around a 2 minute delay of the Sprinter at Houten Castellum

3 minute delay

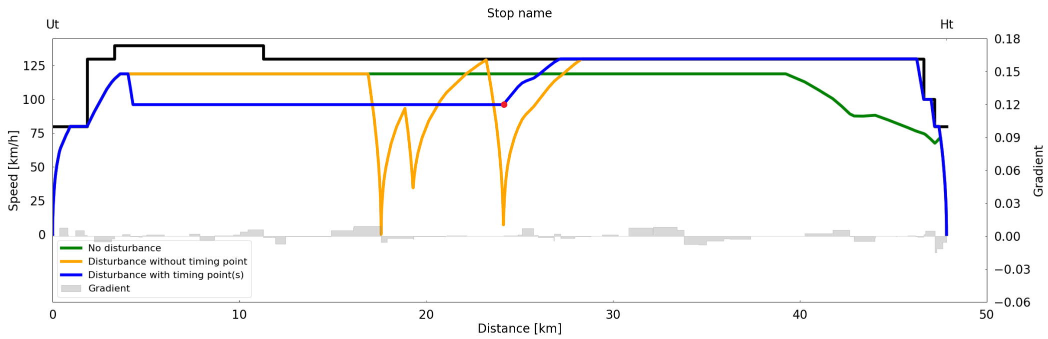


Figure B.3: Intercity speed profiles of scenarios around a 3 minute delay of the Sprinter at Houten Castellum

B.2. Culemborg

1 minute delay

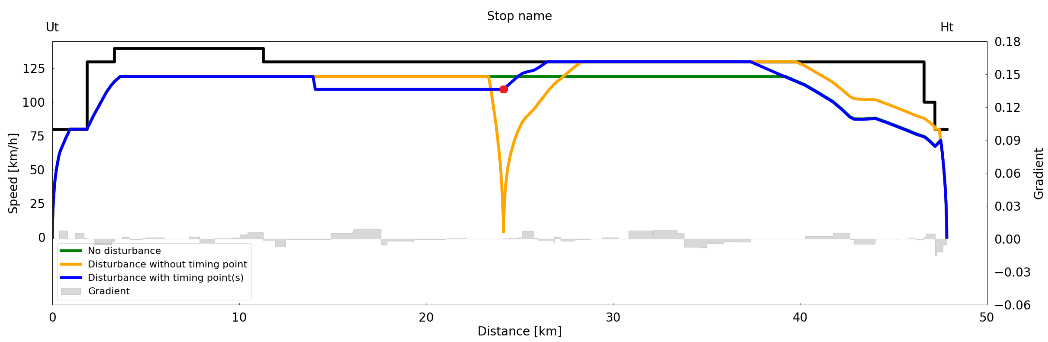


Figure B.4: Intercity speed profiles of scenarios around a 1 minute delay of the Sprinter at Culemborg

2 minute delay

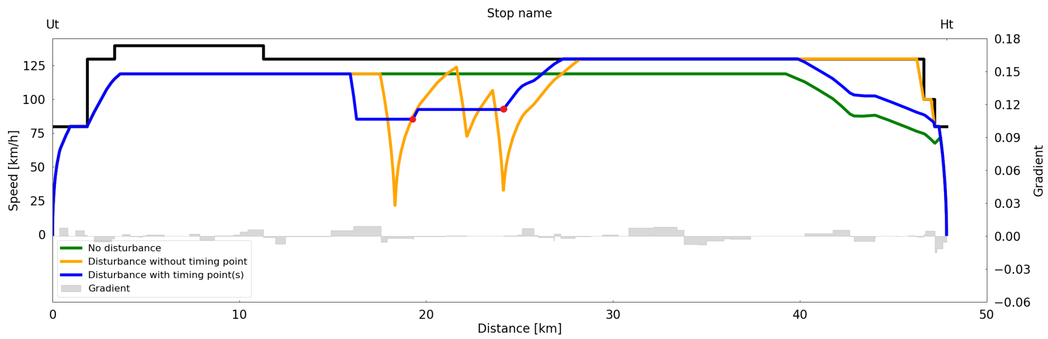


Figure B.5: Intercity speed profiles of scenarios around a 2 minute delay of the Sprinter at Culemborg

3 minute delay

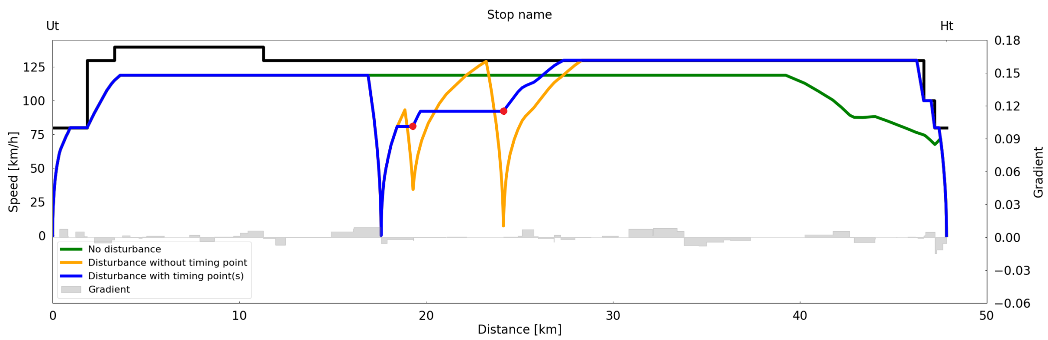


Figure B.6: Intercity speed profiles of scenarios around a 3 minute delay of the Sprinter at Culemborg

4 minute delay

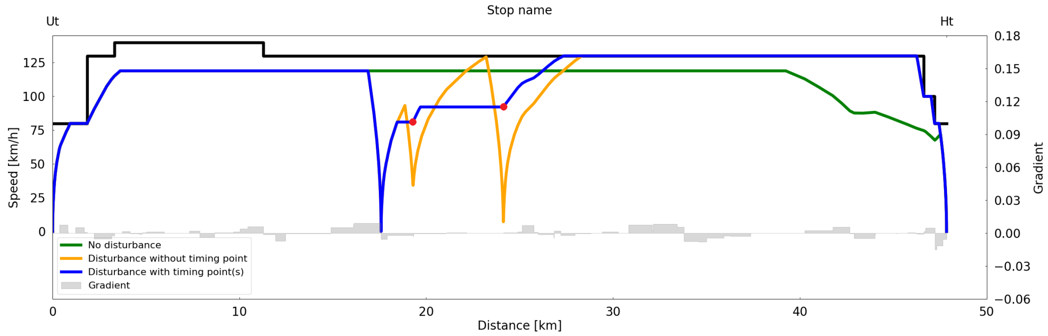


Figure B.7: Intercity speed profiles of scenarios around a 4 minute delay of the Sprinter at Culemborg

5 minute delay

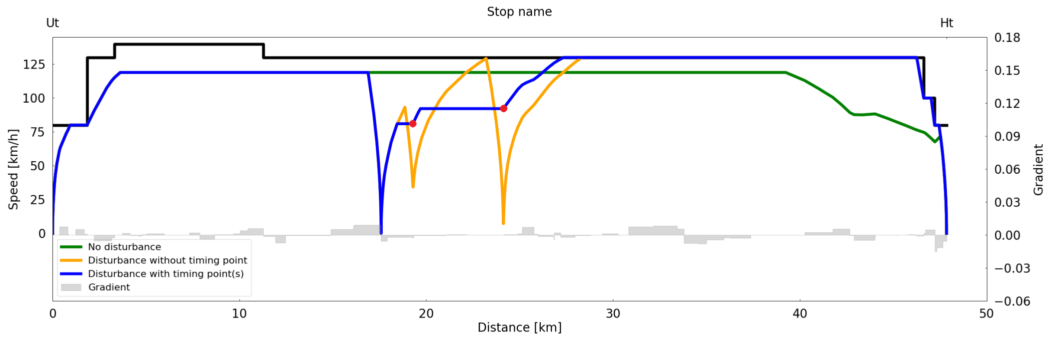


Figure B.8: Intercity speed profiles of scenarios around a 5 minute delay of the Sprinter at Culemborg

B.3. Zaltbommel

1 minute delay

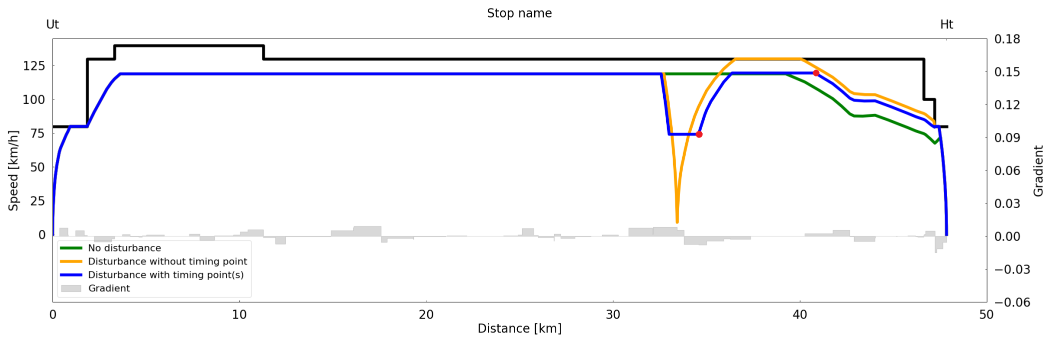


Figure B.9: Intercity speed profiles of scenarios around a 1 minute delay of the Sprinter at Zaltbommel

2 minute delay

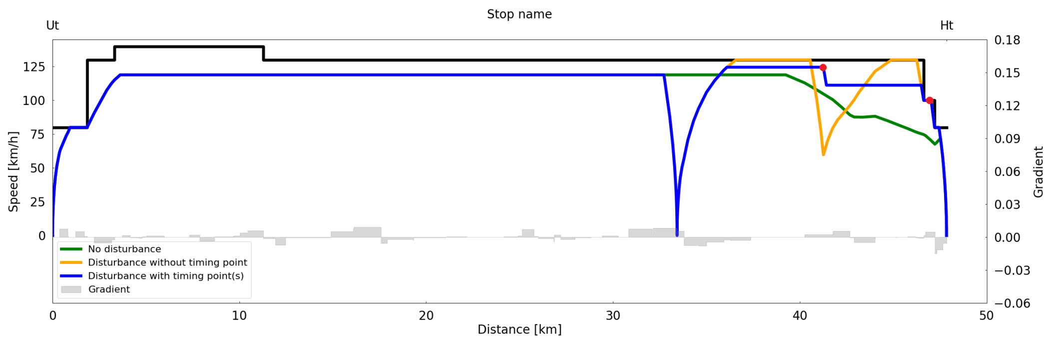


Figure B.10: Intercity speed profiles of scenarios around a 2 minute delay of the Sprinter at Zaltbommel

3 minute delay

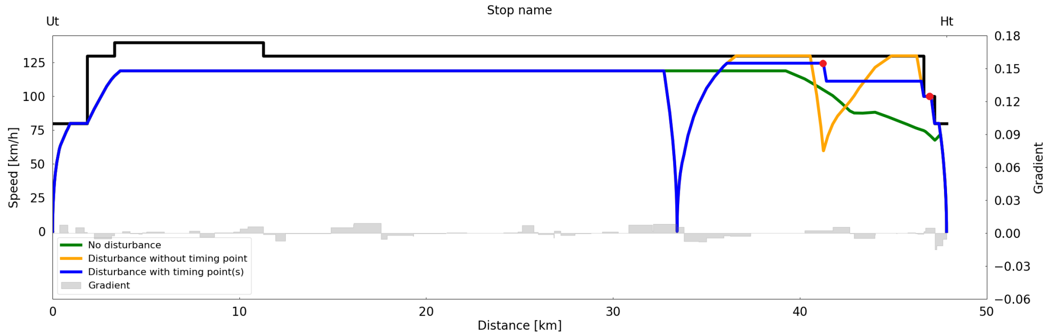


Figure B.11: Intercity speed profiles of scenarios around a 3 minute delay of the Sprinter at Zaltbommel

4 minute delay

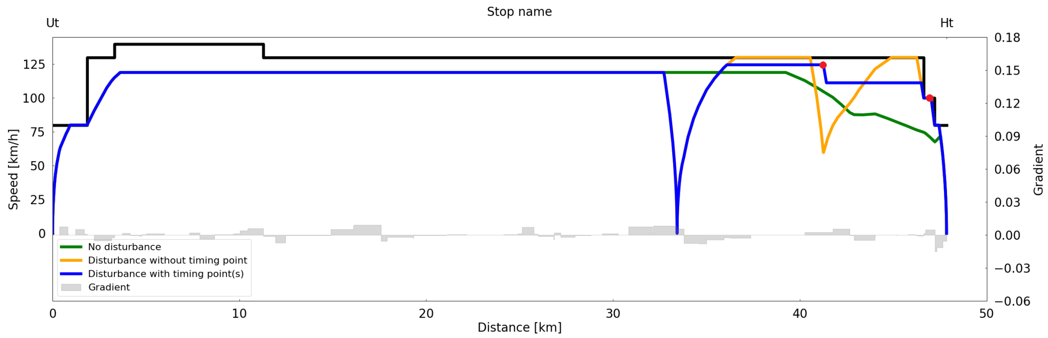


Figure B.12: Intercity speed profiles of scenarios around a 4 minute delay of the Sprinter at Zaltbommel

5 minute delay

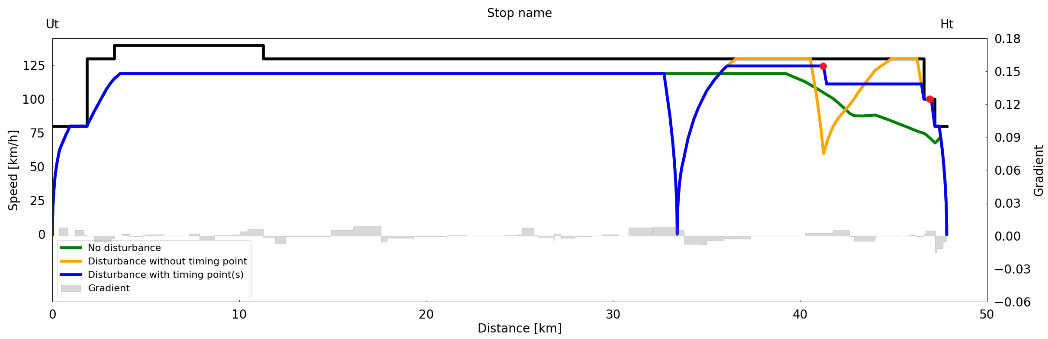


Figure B.13: Intercity speed profiles of scenarios around a 5 minute delay of the Sprinter at Zaltbommel