

Research Article

An Experimental Analysis of Hierarchical Rail Traffic and Train Control in a Stochastic Environment

Pengling Wang ^{1,2,3,4}, Annik Jeiziner ⁵, Xiaojie Luan ⁴, Valerio De Martinis ^{4,6}
and Francesco Corman ⁴

¹College of Transportation Engineering, Tongji University, Shanghai, China

²Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai, China

³Shanghai Key Laboratory of Rail Infrastructure Durability and System Safety, Tongji University, Shanghai, China

⁴IVT-Institute for Transport Planning and Systems, ETH Zurich, Zurich, Switzerland

⁵Swiss Federal Railways, Bern, Switzerland

⁶IDP-Institute for Data Analysis and Process Design, ZHAW, Zurich, Switzerland

Correspondence should be addressed to Pengling Wang; pengling_wang@tongji.edu.cn

Received 15 December 2021; Accepted 15 April 2022; Published 29 May 2022

Academic Editor: Evelien van der Hurk

Copyright © 2022 Pengling Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The hierarchical connection of Rail Traffic Management System (TMS) and Automatic Train Operation (ATO) for mainline railways has been proposed for a while; however, few have investigated this hierarchical connection with the real field. This paper studies in detail the benefits and limitations of an integrated framework of TMS and ATO in stochastic and dynamic conditions in terms of punctuality, energy efficiency, and conflict-resolving. A simulation is built by interfacing a rescheduling tool and a stand-alone ATO tool with the realistic traffic simulation environment OpenTrack. The investigation refers to different disturbed traffic scenarios obtained by sampling train entrance delays and dwell times within a typical Monte Carlo scheme. Results obtained for the Dutch railway corridor Utrecht–Den Bosch prove the value of the approach. In case of no disruptions, the implementation of ATO systems is beneficial for maintaining timetables and saving energy costs. In case of delay disruptions, the TMS rescheduling has its full effect only if trains are able to follow TMS rescheduled timetables, while the energy-saving by using ATO can only be achieved with conflict-free schedules. A bidirectional communication between ATO and TMS is therefore beneficial for conflict-resolving and energy-saving.

1. Introduction

In railway systems, intelligent rescheduling (Traffic Management System, TMS) and autonomous driving (Automated Train Operation, ATO) are not a novelty. The automation module partially or entirely takes over the tasks of dispatchers and drivers. With the support of TMS and ATO, trains have no conflicts with other trains and run with their optimal speed profiles, therefore optimizing the utilization of railway capacity and overall system efficiency. Those benefits, however, have been found so far only on metro systems: many driverless metro systems are now operating around the world. The current challenge is to apply automated systems to operations on conventional (mainline) railways. In mainline

railways, the state-of-the-art execution of both traffic control and train control happens manually. Only very few automated processes are implemented to guarantee a safe train operation. One such example is the Automatic Train Protection (ATP) supporting the train's overspeed protection and keeping a safe headway between trains. However, the increasing need for sustainable transportation of people and goods has led to a growing density of traffic on a limited railway network. Extending the network with new railway lines is rarely an option, as such projects require huge investments and long-term planning. Therefore, the capacity of the existing network needs to be increased while simultaneously keeping the energy consumption low. This can be realized by further automating processes.

The TMS is often proposed to help/replace dispatchers. TMS automatically monitors the real-time traffic state and in case of delays/conflicts computes an updated schedule that minimizes delays by changing train speeds, times, routes, and orders [1]. The TMS makes decisions at the traffic level, while the ATO takes actions at the train level, which is responsible for executing the production plan, and makes real-time driving decisions [2]. The ATO system has been typically defined as hierarchically below the TMS [3–5]. This hierarchical connection of TMS and ATO has been proposed for a while and the advantages are promising. However, few have investigated the interaction with the real field [6], like (i) stochastic external disturbances to real train operations (e.g., unforeseen dwell and/or running times extensions), (ii) missing or erroneous information about current traffic state and train parameters, etc. The main open question for both researchers and practitioners is whether automatic rescheduling and train operation tools can concretely reach the improvements found in theory, when instead they are actually interfaced with uncertainty.

Therefore, this paper is to set up a hierarchical connection between TMS and ATO for mainline railways and thus to evaluate the performances of such a system in stochastic environment in terms of punctuality, energy efficiency, and conflict-resolving with simulation methods.

In general, railway traffic is based on a timetable, defining the exact arrival and departure times for each train at each station. Meanwhile, railway traffic is subject to many different kinds of perturbations propagating quickly through the system. Therefore, real-time dispatching is needed to restore the original timetable and resolve conflicts. A conflict in that sense occurs, when two or more trains require the same piece of infrastructure at the same time. Nowadays, most of the railway traffic control is done by human dispatchers, which base their decisions on training and experience, lacking intelligent decision support. Nevertheless, the research on the topic of automated railway traffic control by the TMS is extensive. An overview can be found in [7, 8].

The basic principle is to either re-route, re-time, and re-order trains, or cancel services to restore the original timetable while fulfilling certain other objectives, such as the reduction of the total delay or the reduction of the consecutive delay. There exists a variety of approaches on how to solve this problem, differing in the following considerations:

- (i) *Consideration of future changes*: static TMS gets information on the traffic state at one point in time and computes the control measures and implements them, assuming perfect knowledge of future traffic states. Dynamic TMS on the other hand takes uncertainty into account and after a certain rescheduling interval, the control measures are computed again with updated traffic state information [6].
- (ii) *Level of detail*: macroscopic rescheduling approaches consider stations only while microscopic rescheduling approaches consider each block section and signal [9–11].
- (iii) *Dimension of disturbance*: small disturbances only request small rearrangements in the timetable, while

big disturbances might require the cancellation and/or replacement of certain trains [12, 13]. Thus, the rescheduling processes regarding big disturbances might also take into account the rearrangement of crews and rolling stock.

- (iv) *Solving algorithm*: different solving approaches have been adopted for solving the rescheduling problem, such as the alternative graph approach, mixed-integer linear program (MILP), etc. An overview can be found in [7, 8].

Despite the variety of rescheduling approaches and TMS types, many TMSs do not consider detailed train dynamics (e.g., train mass, train failure, different driving strategies, etc.). This could lead to a side effect that the real-time TMS traffic plan cannot be executed well by drivers/ATO systems. Nevertheless, being able to follow TMS timetables is essential to make automatic railway traffic control practicable [4, 14, 15]. In fact, TMS has so far been mainly a scientific topic and there are very few practical applications; showing applicability in real life conditions is very important.

ATO is the automatic real-time control of a train's acceleration, braking, cruising, and coasting commands according to the actual traffic plan [16]. The technology has emerged over the last few years with the development of control and computer technologies and is considered to be a very promising approach in order to improve energy efficiency and network capacity compared to the manual train control by a train driver [2, 17]. The latter is based on training and experience rather than exact computation, which is why it is difficult to guarantee the optimal operation, for example, in terms of energy efficiency, punctuality, and stopping accuracy.

Many metro lines are operating ATOs (e.g., Paris Metro, London Underground, Beijing Subway, and Lausanne metro). Metro railways depend strongly on an efficient operation to get the maximum capacity out of their rail network. As they have a relatively simple timetable design, homogeneous trains, and few uncertainties due to weather conditions, they are predestined for the ATO system. On the other hand, mainline railways have a much more complex timetable [18], different train types, dependent connections, and uncertainty due to weather conditions. Therefore, developing an ATO system for mainline railways is much more complex. Nevertheless, some mainline railways are equipped with driver advisory systems (DASs), which support drivers with additional advice on optimal speeds, control regimes [19, 20]. There exist practical DAS examples, such as the AF (Automatic Function) in the Lötschberg Base Tunnel in Switzerland [21], the GreenSpeed applied in Denmark [22], and the Computer-Aided Train Operation (CATO) applied in Sweden [23].

The research on integrated traffic and train control becomes more and more popular only in recent years. One research stream of integrated traffic and train control is to simultaneously solve the rescheduling problem and the train trajectory optimization problem [24, 25]. A TMS has a clear overview of the entire traffic state and possible conflicts. Therefore, the TMS can not only re-route, re-time, and re-

order trains, but also compute and transmit optimal speed trajectories directly to involved trains. Such a framework spares the train from braking or stopping at signals due to an unexpected, rescheduled timetable and it can therefore prevent both time and energy loss. In [9, 10], such a model is proposed, where the dispatching and train trajectory computation are not done sequentially, but rather both tasks are part of the entire problem-solving.

Another research stream is to specify proper system frameworks for integrated management of traffic and train operation. In other words, TMS and ATO are still independent tools. Research has been focused on optimizing the communication and interaction between the two systems and investigating efficient ways of distributing intelligence (functions) among the two systems. The European project ONTIME has developed a framework to set up standards for the inter-operable cross-border automatic real-time management [14, 15]. The ATO system has been typically defined as hierarchically below TMS [3,4,19]. With such a hierarchical framework, the TMS takes charge of conflict detection and computes conflict resolutions, while the ATO undertakes executing traffic plans. However, there is always a gap between optimized results and actual situations. The first reason is due to unexpected disruptions, such as unpredictable dwell times, inappropriate driver behaviors (in case of using DASS), communication delays, and so on. The second reason, however, results from deviations between train dynamics estimation/prediction in TMS and in the real situation. The train dynamics estimation/prediction in TMS not only relies upon the real-time information about traffic state, but also needs the real-time information about train parameters (driving strategies, train weight, train failures, etc.), as those factors lead to different train performances. If TMS has no access to real-time train parameters, these train dynamics estimation/prediction in TMS cannot represent the real case. Consequently, the differences bring inaccuracies and difficulties for ATOs/drivers to follow real-time traffic plan accurately. Only recently, [5] elaborated a more holistic theoretical framework, which uses bidirectional communication between traffic management and train automation. By doing so, the TMS knows if its traffic plan has been fulfilled and it can improve its calculation according to the real-time feedback from train automation. However, the bidirectional communication between traffic management and train automation has not been examined with detailed experiments yet.

According to the background information and literature review given above, this paper focuses on specifying a control framework for an integrated management of traffic and train operation on mainline railways and sets up a simulation environment to evaluate the integrated traffic and train control in stochastic and dynamic conditions. Compared to the other large study of interaction between traffic control and simulation, like the proof-of-concept presented in [15], the present paper uses a microsimulator tool like OpenTrack, which has been validated and used in many real life projects and is a de facto standard and connects not only TMS and simulation, but also ATO for energy efficient and smooth train control, with simulation.

Moreover, the stochastic aspects included in comparable studies in the state of the art are not going beyond a simple Monte Carlo scenario of delayed traffic, while we introduce randomness/incomplete knowledge also in train parameters, and information availability. Finally, we evaluate a range of KPIs, instead of pure delays.

This paper interfaces a rescheduling tool [26, 27] and an ATO tool [28] with the realistic traffic simulation environment OpenTrack [29]. Three different traffic scenarios are designed, including an on-time scenario, a delay scenario, and a train failure scenario. The investigation has been performed by referring to the three different traffic scenarios within a typical Monte Carlo scheme. The proposed methodology has been applied to a real case-study in the Netherlands: the railway corridor between Utrecht and Den Bosch. Results show that, in case of on-time operations, the implementation of ATO systems is beneficial for maintaining timetables and saving energy costs. Moreover, in case of delay disruptions, the TMS rescheduling has its full effect only if trains are able to follow TMS rescheduled timetables, while the energy-saving by using ATO can be only achieved with conflict-free schedules. A bidirectional communication between ATO and TMS is beneficial for conflict-resolving and energy-saving.

The rest is organized as follows. Materials and Methods presents the integrated TMS and ATO control and simulation framework. Results and Discussion presents the case study in the Dutch railway corridor Utrecht-Den Bosch, and the evaluation we perform. The last section concludes the paper and discusses the future research directions.

2. Materials and Methods

In this section, a simulation framework for integrated traffic (TMS) and train control (ATO) is presented. This involves the choice of the control scheme, the simulation platform, the rescheduling method, the speed profile optimization method, and the ATO driving command generation method.

2.1. Control Framework. Figure 1 presents a cascade control loop framework for integrated rail traffic management and train control, which is extended from the basic control framework proposed by ON-TIME project [14]. The framework consists of four basic control loops. The first loop is the TMS control loop, which is responsible for traffic state monitoring, conflict detection, and resolution. In other words, TMS computes and translates the real-time traffic plan to each train based on information from different sensors. The real-time traffic plan, here indicated as the rescheduled timetable, defines conflict-free train paths, with detailed route setting, for each train and updated arrival, departure, and passing-through times, which typically minimize delays. In the second loop (ATO/DAS control loop I), a speed profile and the associated time-distance path are computed based on outputs from the first loop. The speed profile is derivable and fulfills a set of constraints and objectives, such as minimal energy consumption, punctuality,

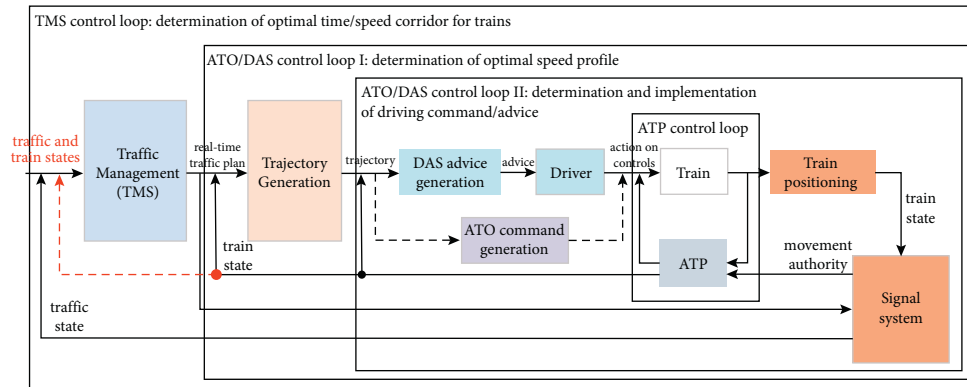


FIGURE 1: Integrated traffic and train control framework.

or riding comfort. The third control loop, ATO/DAS control loop II, is used for generating driving advice/commands to train drivers, or to directly control train movement to follow the optimal trajectory, or get back to the trajectory in case of deviations. Then the final automatic train protection (ATP) control loop supervises independently that any safety restrictions are respected by the advice implementation.

Overall, TMS is in charge of the traffic level and produces the conflict-free real-time traffic, while the ATO is in charge of optimizing train speed profiles (trajectory computation) to match the time constraints set by TMS decisions. It has to be mentioned that a TMS nowadays predominantly has to deal with non-continuous train detection, depending on certain fix-installed infrastructure elements (e.g., axle counters) (without the red dashed line in Figure 1). In other words, the TMS only knows one train's position after the train passes an axle counter. We assume that instead trains can report some status description to a control center. This requires sensor and onboard processing capabilities, which can be integrated into a supervisory control for movement, and onboard monitoring systems (e.g., for power equipment) and communication channel to a trackside control center. Concerning this latter, on top of the existing communication channels now dedicated to movement supervision (in legacy and ETCS systems) we can assume more pervasive communication capabilities, allowing the TMS to be quasi-continuously connected to every train's ATP/ATO (see, for instance, the project 5GRail: 5G for future railway mobile communication system [30]).

We therefore assume possible to enhance the future train states' predictions, conflict detection, and resolution functions with more information on current trains' performances and conditions. In addition, the ATO nowadays receives only train speed and position reports. By those channels, the information reported on train position and speed can be enriched with other information, from onboard monitoring systems, such as acceleration and deceleration, failures (e.g., doors not opening), train integrity, automatic passenger counters, remaining useful life of critical components, etc., as well as from train dynamics (estimated or calibrated train acceleration and deceleration capabilities, train weight, train failures, etc.). All this information is assumed available to the ATO, to improve control accuracy.

2.2. Simulation Framework. We construct a simulation platform to analyze the potential benefits and limitations of the integrated traffic and train control framework presented in Figure 1. The simulation framework implemented in this work is shown in Figure 2. It consists of four modules: TMS, trajectory generation, ATO command generation, and OpenTrack simulation platform. The data exchanged is detailed as follows, as described in both upper part and lower part of the figure. Infrastructure pertains length, topology, routes, switches, gradients, radii, etc. Vehicle pertains rolling stock characteristics such as maximum speed, acceleration, braking, etc. The planned timetable and the rescheduled timetable specify the operation of trains as a sequence of time and space points. A train failure can be represented by updated values for maximum speed, acceleration, and braking possibilities for the affected vehicle. The dwell time represents the time spent at a planned stop; the entrance delay represents the variation with regards to the planned departure times at any planned stop. In the paper, we also investigate different scenarios in which some information on delays and failure is not available, to provide benchmark cases. The speed profile lookup table (LUT) reports the expected consequences from a series of actions (accelerating, braking), for any considered position and speed. ATO Commands are the specific decision on changing/maintaining the speed of a train at a specific moment and space.

The most relevant input and outputs of the four modules are summarized in Table 1, divided into offline inputs (data which does not change, and is available on beforehand), online input (data which is available only at a specific moment), derived inputs (computed by another module in our proposed framework), and outputs (data made available to other modules, or implemented as control actions). We also mention when the data exchange takes place.

The TMS rescheduling function is built based on the work done in [26, 27]. We here recall briefly the main characteristics of the models in the following sections, while for formal definitions and detailed examples we refer to [26, 27]. The TMS is responsible for the rescheduling of the original timetable according to entrance delays, train failures, etc. The output of TMS is a rescheduled timetable. In case trains are controlled by ATOs, the rescheduled timetable is transmitted to the ATO command generation

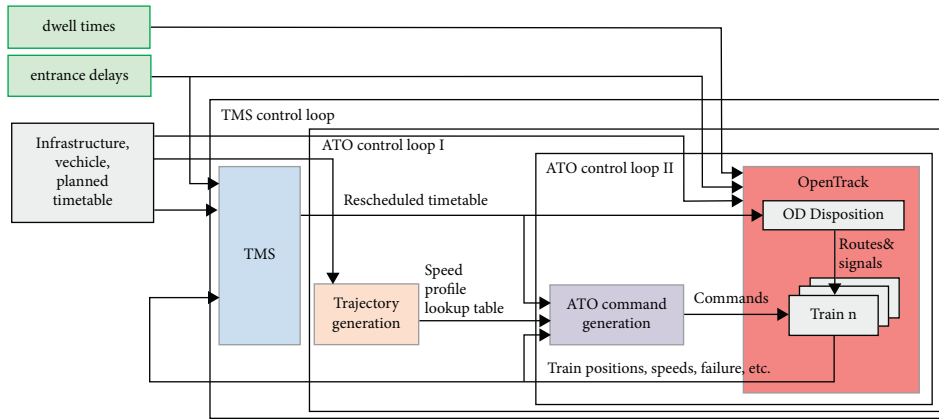


FIGURE 2: Simulation framework.

TABLE 1: Inputs and outputs.

	TMS	Trajectory generation	ATO command generation	OpenTrack
Static input	Infrastructure Vehicle Planned timetable	Infrastructure Vehicle	—	Infrastructure Vehicle Planned timetable
Dynamic input	Entrance delay Train failure	—	Train failure	Dwell time Entrance delay Train failure
Acquired	<i>Start of each run</i>	—	<i>Start of each run</i>	<i>Each simulation step</i>
Input from other modules	Train position; train speed	—	Speed profile LUT Rescheduled timetable Train position Train speed	Rescheduled timetable; ATO commands
Acquired	<i>Start of each run</i>	—	<i>Each simulation step</i>	<i>Each simulation step</i>
Outputs Generated	Rescheduled timetable <i>End of each run</i>	Speed profile LUT <i>When run</i>	ATO commands <i>Each simulation step</i>	Tran position, train speed <i>Each simulation step</i>

module. Otherwise, the rescheduled timetable is transmitted to the disposition system in OpenTrack before the run-through. On the other hand, it is assumed that the train operation can be controlled by the ATO or the OpenTrack internal speed controller.

The ATO system consists of two parts, which, respectively, correspond to ATO control loops I and II in Figures 1 and 2. The first part is an offline trajectory generation module which calculates several speed profile lookup tables with respect to the given rolling stock (different train parameters) and infrastructure. The second part is an ATO command generation module, which decides on optimal driving commands with regard to the real-time train location, speed, speed profile lookup tables, and rescheduled timetable. The ATO trajectory generation function and driving command generation function are built on the basis of [31]. More detailed descriptions regarding speed profile lookup table and ATO command generation algorithm are introduced in Sections 2.5 and 2.6.

The real-time simulation interconnects two elements: the microscopic simulation tool OpenTrack [29] simulating the train movements and railway system and the ATO command generation algorithm deciding on optimal actions for

every train at different locations. At regular time intervals, updated traffic information is gathered from OpenTrack and transferred to the ATO command generation tools to compute optimal driving commands, based on all available information. A basic introduction to OpenTrack is included in Section 2.3.

2.3. Simulation of Railway Traffic with OpenTack. The railway system is represented by the microscopic simulation tool OpenTrack [29]. OpenTrack contains detailed information about the infrastructure (signals, stations, radii, gradients, and speed limits), the rolling stock (length, mass, maximum traction power, maximum traction force, deceleration, and resistance parameters) as well as timetable entries (departure, arrival, and dwell times). As output, OpenTrack generates several data files for evaluation, such as speed, acceleration, or energy consumption per time or distance. The simulation of train movement is a time-driven model, meaning that the simulation is realized by integrating Newton’s motion equations for each time step in order to get the train’s position and speed. The trains must, however, also respect signals indicating whether the according block

section is reserved for the corresponding train and no other train is blocking this section.

The train movement is controlled by ATO driving commands or OpenTrack internal speed control algorithm. OpenTrack provides a “train performance” parameter to simulate different human driver behaviors. The train performance parameter is the scaling value for acceleration and maximum speed (initial state 100%). By default, OpenTrack simulates the train movement at maximal available acceleration, speed, and deceleration, respecting the infrastructure speed limits and signals. This results in a minimum travel time, provided that the train performance is set to 100%. OpenTrack suggests imitating human driver behavior by reducing the performance of the train; i.e., OpenTrack allows users to define their own piece-wise linear distribution functions for travel time performance of punctual (on time) and delayed individual trains. This allows OpenTrack to provide a more realistic distribution of travel times.

Besides, OpenTrack allows the connection to other programs via a programming interface (API: application programming interface) so that they can send standardized commands to OpenTrack and receive defined status messages from OpenTrack [32]. In our proposed simulation framework, OpenTrack is connected to the TMS module and the ATO command generation module (Figure 2). On the one hand, TMS rescheduled timetables can be directly handed over from the TMS to OpenTrack, so that the dispatching in OpenTrack is handled according to TMS timetables accordingly. On the other hand, OpenTrack receives driving commands from the ATO driving command generation function and sends train position reports to corresponding ATOs.

It is also possible to implement delays or other perturbations in OpenTrack. To simulate train movement in a stochastic environment, the simulation takes into account random entrance delays and random dwell times. The entrance delays are assumed to be known by the scheduler (TMS) and the simulator (OpenTrack), while the random dwell times are known by the simulator, but not to the scheduler. The scheduler uses predefined dwell times for rescheduling.

2.4. TMS Rescheduling Algorithm. The TMS algorithm used in this paper is based on the work presented in [26, 27]. Its task is to reschedule the original timetable in order to resolve conflicts and to reduce delays, which have been caused by various perturbations. As a standard assumption, the algorithm is assumed to have perfect future knowledge of all entrance (initial) delays arising during the considered time horizon [7]. The TMS maintains an estimated internal evolution of the network, therefore, by which there is no dynamic detection of the traffic state at any point in time during the run-through (train failure information is assumed to be available when relevant). In addition to the entrance delays and train dynamics, further inputs to the TMS include the infrastructure information (position of signals and stations), traffic characteristics (origin, destination, route, minimum travel time, etc.), and original timetable, to compute minimal travel times.

The TMS algorithm formulates the rescheduling process as a mixed-integer linear programming (MILP) model, based on the time-continuous formulation method. The model optimizes train orders and arrival and departure times at passing stations based on the current delay and traffic situation, minimizing the sum, over all trains, of the absolute delay times (both late arrivals and early arrivals) at all visited stations, subject to a number of constraints for ensuring the operational and safety requirements. The key constraints include the following: (1) *transition constraint* to force the spatial and temporal transition of each train as it moves on the rail network; (2) *train travel/dwell time constraint* to ensure the required minimum travel/dwell time; (3) *safety headway constraint* to define the safety time interval between trains; and (4) *capacity constraint* to guarantee that any pair of trains using the same infrastructure (track/block) are conflict-free. For the sake of space, we do not present here the detailed formulations of the TMS algorithm; interested readers may refer to the MILP model presented in [27] for those details. Note that the TMS algorithm can be solved by a standard MILP solver, e.g., CPLEX or Gurobi.

In the TMS algorithm, we consider microscopic infrastructure details (i.e., block sections), and we allow only one train to occupy a block section at any given time. Thus, the output of our TMS algorithm is a microscopic schedule, i.e., arrival and departure times for each block section, with a guarantee of the solution feasibility at the microscopic level. It is worth noting that OpenTrack timetable entries are only possible for stations (not, e.g., signals) and entries at stations are only respected by trains stopping at that respective station. Therefore, we neglect the microscopic details of the generated TMS timetable and extract only the macroscopic schedule (i.e., the train arrival and departure times at stations), used as the OpenTrack timetable entries.

2.5. ATO Trajectory Generation. The ATO trajectory generation function and driving command generation function are built on the basis of [31]. Within the ATO trajectory generation module, the train operation control problem is formulated as a Markov decision process (MDP) and solved using an offline approximate dynamic programming (ADP) method to learn the energy and time costs over iterations.

Specifically, the train driving process is formulated as an MDP by dividing the train journey into small pieces with a set of discrete locations (which correspond to the control points); see Figure 3. The discretized location is represented with the symbol d in the following description. At every discretized location d , a set of potential train speeds (S_d) are defined. Those speeds have to be non-negative and are bounded from above by the speed limit. In addition, at every speed point S_d , a set of potential driving commands are defined. The driving command at S_d is represented by $x_d(S_d)$, while the set of potential driving commands is represented by $\mathcal{X}_d(S_d)$, $x_d(S_d) \in \mathcal{X}_d(S_d)$. The driving command set is defined by following two heuristic rules: (i) the driving command is among the four optimal regimes for energy-efficient driving: maximum traction (MT), speed

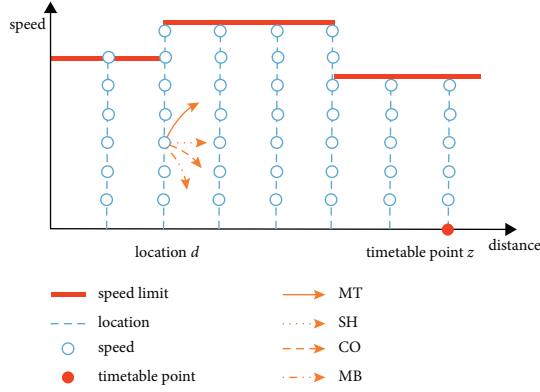


FIGURE 3: Example of MDP.

holding (SH), coasting (CO), and maximum braking (MB) [33–35]; (ii) the driving command respects the maximum accelerating/braking rate restrictions for passengers' riding comfort [36]; and (iii) performing the decision MT from the departure station, MB to stop at the arrival station, and MT, SH, CO, or MB in the middle of the journey but avoiding decisions that would violate the speed limits or result in unnecessary stops and low travel speeds.

For every combination of discretized location d , speed S_d , and driving command $x_d(S_d)$, an ADP-based method [31] is used to learn the following two values:

- (i) The single step time cost $t_d(S_d, x_d(S_d))$: the time cost incurred from state S_d to the next discretized location $d + 1$ by applying action $x_d(S_d)$
- (ii) The time estimation to reach the next timetable point z (stop stations, passing junctions, etc.) from each state S_d by using energy-efficient control laws, denoted by $T_d^z(S_d)$

The ADP-based method learns the two values through many iterations offline. The learning process also takes into account uncertainty in traction force and train resistance, and their impact on travel time and energy consumption. The learned results are stored in a so-called lookup table, as the example presented in Table 2, which is then adopted by the ATO command generator for online decision-making, as described in the next section.

2.6. ATO Command Generation. The ATO command generation module receives the up-to-date train position and speed report from OpenTrack in a regular time interval. As soon as the train reaches any predefined discretized location, the ATO needs to come up with a driving command and sends it to OpenTrack for the control of train movement. The driving command is determined according to the off-line learning results (lookup table, Section 2.3) as well as a decision-making function in (1). Assume a train is reaching a discretized location d and heads towards a timetable target point, z . Its speed is closest to the speed point S_d , then the optimal driving command at stage d is selected with the following function:

$$X_d(S_d) = \operatorname{argmin}_{x_d \in \mathcal{X}_d(S_d)} \{ |t_d(S_d, x_d) + T_{d+1}^z(S_{d+1}) - \bar{T}_d^z(S_d)| \}, \quad (1)$$

where $X_d(S_d)$ refers to the selected driving command at location d . $t_d(S_d, x_d)$ and $T_{d+1}^z(S_{d+1})$ are the learning results obtained from the offline ADP-based algorithm proposed in [31]. $t_d(S_d, x_d) + T_{d+1}^z(S_{d+1})$ is equal to the overall estimated running time from d to z . $\bar{T}_d^z(S_d)$ is the scheduled remaining time from state S_d to z . The difference between the two items refers to the ability of the train following its timetable by using energy-efficient driving commands. In case that the overall estimated arrival time is later than the scheduled time, $X_d(S_d)$ often goes for a fast driving in order to reduce delays. Otherwise, $X_d(S_d)$ goes for a cruising or a coasting, in order to arrive on time and saving energy. Function (1) ensures the selected driving command prioritizes on-time arrival over energy reduction. In addition, the ATO command generation module works online but is extremely fast as it has to take a single control decision mostly based on precomputed information.

3. Results and Discussion

3.1. Setups

3.1.1. Test Case. The corridor between Utrecht and Den Bosch ('s-Hertogenbosch) from the Dutch Railways has been used. The corridor is about 48 km long and has nine stations. The infrastructure characteristics consist of an accurate description of all track sections, points, speed signs, gradients, and signals over the entire track layout from Utrecht until Hertogenbosch. Details can be found in [28]. The infrastructure is equipped with the fixed-block signalling system NS'54 and the automatic train protection system ATB. The corridor is simulated in one direction only (Ut–Ht); however this does not affect the results as the other direction operates on different tracks. Two train types are operated on this corridor: intercity trains (IC) and sprinters (SP). The characteristics (train mass, rotating mass factor, train length, maximum traction force and power, maximum braking rate, and train resistance curves) of the sprinter and intercity can be found in [28]. Each hour, three IC services drive directly from Utrecht to 's-Hertogenbosch without any intermediate stop. Moreover, two SP services stop at each station, one terminates in Geldermalsen, the other in 's-Hertogenbosch. For the simulation a time horizon of two hours was considered with a half-hourly clocked timetable, resulting in 20 trains departing from Utrecht (12 ICs, 8 SPs).

3.1.2. Simulated Human Driver Behaviors. As we mentioned in Section, OpenTrack provides a "train performance" parameter to simulate different human driver behaviors (different levels of acceleration, deceleration, and cruising speeds). In this work, the train performances are random parameters and set based on the following rules: for trains operating punctually: 20% of trains operate with performance values between 90% and 92%; 60% of trains operate with performance values between 92% and 94%; and 20% of

TABLE 2: Speed profile lookup table example.

Symbols units description	d (m) discretized location	S_d (km/h) speed	$x_d(S_d)$ –driving command	$t_d(S_d, x_d(S_d))$ (s) single-step time cost	z (m) next timetable point	$T_d^z(S_d)$ (s) estimated time to arrive z (with eco-driving)
Examples	147	50	MT	25	5000	955
	147	50	SH	30	5000	955
					
	147	100	MT	20	5000	885
	147	100	SH	22	5000	885
	147	100	CO	35	5000	885
	147	100	MB	55	5000	885

trains operate with performance values between 94% and 96%. For delayed trains: all trains operate with performance values between 98% and 100%. The selection of train performance parameters is based on the recommendation in the OpenTrack manual.

3.1.3. Random Entrance Delays and Dwell times. We simulate the stochastic environment with random dwell times (and entrance delays). Dwell times are sampled from two different Weibull distributions depending on the size of the station considered (as in [6, 37]). For major stations like Ut, Gdm and Ht dwell times are drawn from the Weibull with scale, shape, and shift parameters equal to 252 s, 2.18 s, and 3.86 s, respectively. For minor stations (i.e., Utr, Utl, Htn, Htnc, Cl, and Zbm) instead dwell times are generated from the Weibull with 20 s, 2.56 s, and 24 s as scale, shape, and shift parameters, respectively. Entrance delays are drawn from a Weibull distribution fitted to recorded data (as in [6, 37]). The scale, shape, and shift parameters of these distributions are, respectively, equal to 394 s, 2.27 s, and 315 s for ICs and 235 s, 3 s, and 186 s for SPs.

3.1.4. Simulated Train Failure. To study how failures affect the train operations, the study simulates a train failure that a local train, SP6927, lost half of its engine. Such a train failure affects several following train operations and causes secondary delays.

3.1.5. Scenarios. The experiment aims to examine the performances of the integrated TMS and ATO under both punctual and delayed circumstances. In addition, it also aims to evaluate the impacts of full/partial knowledge of real-time train parameters. Three sets of scenarios are built to achieve our targets; see Table 3.

Scenario I: the first scenario set simulates the train operation under punctual circumstances. No TMS rescheduling function is needed. The traffic is controlled by OpenTrack internal FCFS (first come first serve) dispatching rule. Scenario I includes two cases: Scenario I(a), as a benchmark, assumes all trains are controlled by human drivers (OpenTrack simulated human behavior). Scenario I(b), instead, assumes all trains are operated with ATOs.

Scenario II: the second scenario set simulates the traffic in case of delay disruptions by assuming simulated trains get random entrance delays at their initial stations. Scenario II includes four cases: in Scenario II(a), trains are controlled by human drivers and the train dispatching follows the original timetable and the FCFS dispatching rule. Scenarios II(b) and II (c) have only ATOs or TMS into use. Scenario II(d) assumes both TMS and ATOs come into use.

Scenario III: the third scenario set is to examine the impact of full/partial knowledge of real-time train parameters. This scenario set considers a real-time train failure of SP6927, which might be known or not known by TMS and ATOs. Both entrance delays and the train failure are considered in Scenario III. Scenario III(a), as a benchmark scenario, assumes all trains are not equipped with ATOs and there is no TMS rescheduled timetable. Scenarios III(b)-III(d) assume all trains are operated with ATOs and follow the TMS rescheduled timetables. The differences between the three scenarios include the following: in Scenarios III(b) and III (c), only ATOs or the TMS are/is aware of the train failure. In Scenario III(d), both TMS and ATOs have knowledge of the train failure. In Scenarios III(b) and III(d), it is assumed that new lookup-tables were computed offline considering different types of train failures. Those lookup tables can be used accordingly in real time for the train speed control.

3.1.6. Monte Carlo Simulation Settings. For every scenario, the analysis is performed over 20 different cases in a typical Monte Carlo setup. Each case is generated by randomly sampling entrance delays (Scenarios II-III) and disturbances to dwell times at stations (Scenarios I-III). The Monte Carlo simulation is performed on the basis of the framework presented in Figure 2.

3.1.7. Metrics. Paper [38] provides a comprehensive review on different metrics for the evaluation of rail service. This paper selects the following metrics from [38] to evaluate the simulation results, respectively, from punctuality, energy-saving, and conflict-resolving perspectives:

- (i) *Average delay* is the average of the total arrival delay (i.e., $\max\{\text{actual arrival time} -$

TABLE 3: Scenarios (HUM: simulated human driver; FCFS: first come first serve; WEIB: Weibull distribution).

Scenario	Train control	Traffic control	Delay	Train failure	Train failure knowledge	
I	(a)	HUM	FCFS	0	×	—
	(b)	ATO	FCFS	0	×	—
II	(a)	HUM	FCFS	WEIB	×	—
	(b)	ATO	FCFS	WEIB	×	—
	(c)	HUM	TMS	WEIB	×	—
	(d)	ATO	TMS	WEIB	×	—
III	(a)	HUM	FCFS	WEIB	✓	—
	(b)	ATO	TMS	WEIB	✓	ATO
	(c)	ATO	TMS	WEIB	✓	TMS
	(d)	ATO	TMS	WEIB	✓	ATO&TMS

scheduled arrival time, 0}), over all planned stops of all trains.

- (ii) *Average early* is the average of the total early arrivals (i.e., $\max\{\text{scheduled arrival time} - \text{actual arrival time}, 0\}$), over all planned stops of all trains.
- (iii) *Punctuality* with respect to a threshold of 1 min (P1min) and 3 min (P3min). This is the percentage of trains whose arrival delay at a planned stop is less than 1 min or 3 min.
- (iv) *Energy* is the overall traction energy cost over the simulation time period (2 h).
- (v) *Red signals* refers to the number of red signals of all trains within the simulation time period (2 h).
- (vi) *Yellow signals* refers to the number of yellow signals of all trains within the simulation time period (2 h). Yellow and red signals request the train to reduce its speed or stop in unplanned locations, which may lead to an extra running time loss as well as less riding comfort for passengers. Therefore, the less yellow and red signals are, the better the traffic is.

The first three metrics indicate the deviation in time and space between the planned and simulated train paths, which are closely related to the capability of the control framework to keep acceptable levels of traffic performance when stochastic disturbances affect operations. The fourth metric is closely related to operational costs. The last two metrics give a measure of how vulnerable the traffic is in terms of unexpected yellow and red signals (the planned timetable has no yellow or red signal).

4. Results

4.1. Scenario I: Undelayed Traffic. The metric values of the scenario are reported in Figure 4. Those variations are reported in average; note that variation might occur, due to the different delay samples, and within each delay sample, due to the heterogeneity of traffic (which we analyze in detail in Figure 5).

Concerning *punctuality*, both Scenarios I(a) and I(b) contain some delays. Those delays are caused by the random dwell time settings we used: in order to simulate the uncertain passenger boarding/alighting time, the simulation assumes random dwell times. Therefore, the simulated dwell

times might be longer than the scheduled dwell times in some cases, causing affected trains depart/arrive late at following stations. The average early and late arrival times by using ATOs are much less than the cases with simulated human drivers. Besides, the values of punctuality indicators, P1min and P3min, of Scenario I(b) are higher than Scenario I(a). We can conclude that the ATO algorithm performs better in maintaining the given timetable in case of no entrance delay. The early arrivals in Scenario I(a) have two reasons. First, the original timetable includes some time supplements to be robust against small disturbances. Second, the OpenTrack speed controller does not make full use of those time supplements and this leads to early arrivals.

Concerning *energy*, the total energy consumption of all trains is reported in Figure 4. Scenario I(b) saves around 487 kWh (11.1%) energy costs, in 2 hours. The energy-saving can be explained as the ATO makes better use of time supplements (less early arrivals) and adopts energy-efficient driving commands (coasting, low-speed holding). Scenario I(a) uses reduced train performance parameters (90%–96%) to simulate human behaviors; i.e., trains are driven with reduced speeds and traction power, instead of moving at their maximum power and speed. Therefore, this is a relatively fair simulation of human drivers, and thus we can conclude that implementation of ATO has an essential effect on reducing energy consumption in case of no big disruptions.

Concerning *conflict-resolving*, red signals are not observed in both Scenarios I(a) and I(b), as Scenario I simulates no-delay circumstances. But, around 10 yellow signals on average are observed in both scenarios. Those yellow signals are caused by two reasons: first, some trains get late departures due to unexpected long dwell times. Late departures make the actual train paths differ from their scheduled paths and cause small headways and yellow signals. Second, some SPs are overtaken by ICs at station Gdm, and the headways between ICs and SPs are quite small around Gdm, therefore, causing yellow signals.

We further consider the heterogeneity of traffic and analyze in Figure 5 the probability distribution of variations in energy (left) and delay (right), split by the two train categories. Positive values along the x -axis identify better performance of Scenario I(b) compared to Scenario I(a): reduced delays, or reduced energy consumption. The energy variation is typically slightly positive; in other terms Scenario

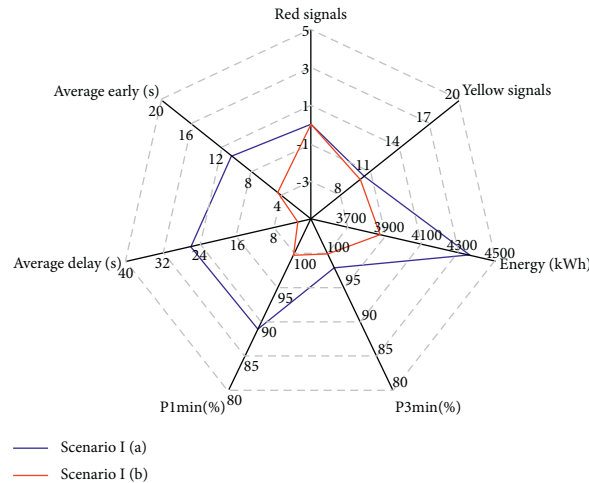


FIGURE 4: Results of Scenario I.

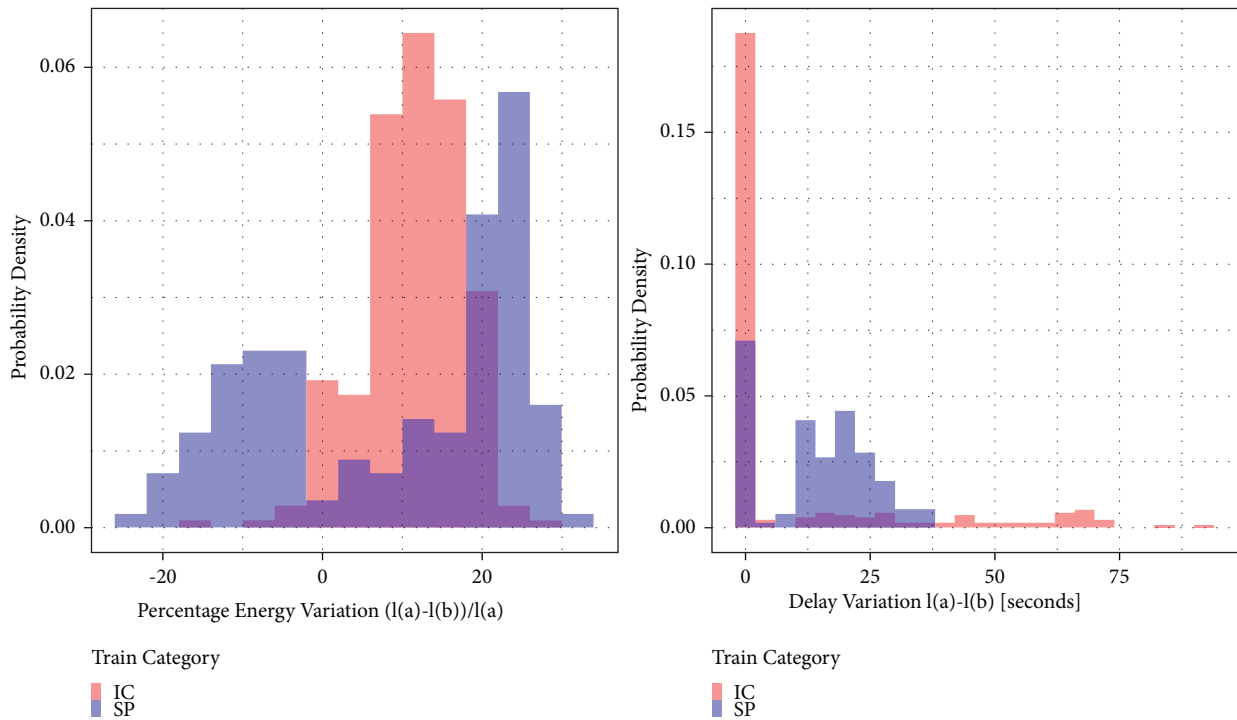


FIGURE 5: Analysis of Scenario I by energy/delays, based on train categories.

I(b) reduces energy compared to Scenario I(a), for IC trains. For the SP trains, a bimodal case arises, with some trains which have 25% less energy, and some which have 10% more energy. The delay variation is always positive; in other terms, Scenario I(b) never increases delay compared to Scenario I(a), with a large peak at 0. The delay variation for the IC trains is larger, spreading all the way to more than a minute. For SP trains, it is more concentrated into a range of about 10–30 seconds. The diversity in outcomes between IC and SP shows that heterogeneous traffic is affected differently from TMS-ATO, already when limited random factors are considered.

4.2. *Scenario II: Delayed Traffic.* The metric values of the scenario are reported in Figure 6. Those variations are reported in average; note that variation might occur, due to the different delay samples, and within each delay sample, due to the heterogeneity of traffic (which we analyze in detail in Figure 7).

Concerning *punctuality*, Scenario II simulates the traffic under delay circumstances in four cases: in Scenario II(a), trains are controlled by human drivers and the train dispatching follows the original timetable and the FCFS dispatching rule. Scenarios II(b) and II(c) have only ATOs or TMS into use. Scenario II(d) assumes both TMS and ATOs

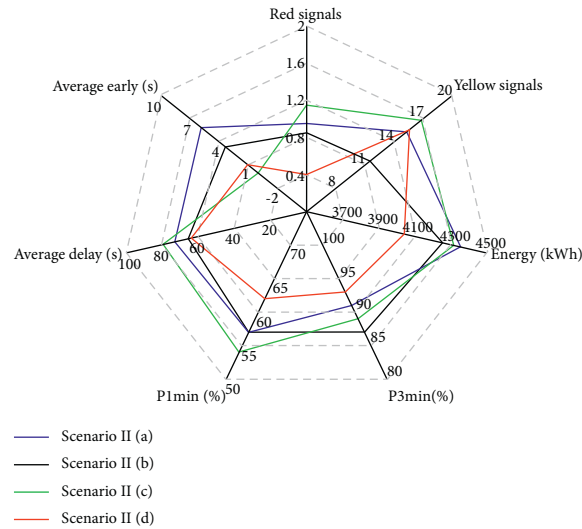


FIGURE 6: Results of Scenario II.

are put into use. In Scenario II(b), although the average delay is lower than the benchmark scenario, the overall punctuality, especially P3min, is not improved. This can be explained by the fact that Scenario II(b) does not have the TMS rescheduling function, so that ATOs can only follow a non-optimal timetable. Scenario II(c) simulates the cases equipped with only the TMS. Its P1min, P3min, and average delay are even worse than Scenario II(a). This indicates that the TMS rescheduled timetable might not be able to achieve its delay reduction function if the rescheduled timetable is not well executed by drivers. Compared to Scenario II(c), the punctuality is much improved in Scenario II(d), which replaces human drivers with ATOs. In other terms, the ATO algorithm is better at following TMS rescheduled timetables than the simulated human drivers.

Concerning *energy*, the energy consumption in Scenario II is higher than the one in Scenario I. Scenario II simulates all trains operated under delay circumstances. The ATO driving command generation function prioritizes delayed trains in catching up timetables instead of energy-saving operation (see Section 2.5). The OpenTrack simulated drivers move delayed trains with performance values between 98% and 100%, which basically means to run with nearly maximum power and speed. Therefore, the energy consumption in Scenario II is higher than Scenario I. Meanwhile, by comparing Scenarios II(b)-II(d) to II(a), we see that the energy costs are reduced by using TMS or/and ATOs. It is observed that, with only TMS or ATOs (Scenarios II(b)-II(c)), the energy costs are reduced by 0.86%–2.29%. With both TMS and ATOs (Scenario II(d)), the energy costs decrease by as much as 8%. It indicates ATOs' ability in energy-saving can be better achieved only with conflict-free schedules from TMS. In addition, the energy reduction in Scenarios II(b)-II(d) can be explained for two reasons: first, the average early arrivals in Scenarios II(b)-II(d) are less than that in Scenario II(a), so that punctual trains in Scenarios II(b)-II(d) make better use of running time supplements and therefore require less energy

consumption than in Scenario II(a). Second, the number of conflicts (red signals) in Scenarios II(b) and II(d) is lower than in Scenario II(a). As the conflicts may lead to energy-consuming braking and re-accelerating processes, fewer red signals basically means less energy wastes due to the braking and reaccelerating processes.

Concerning *conflict-resolving*, Scenario II leads to more red and yellow signals than Scenario I, because Scenario I simulates traffic without delay impacts whereas Scenario II simulates traffic with entrance delays. Delayed trains deviate from their conflict-free scheduled time and space paths, thus resulting in more conflicts. By comparing Scenarios II(b)-II(d) to the benchmark Scenario II(a), it is seen that the number of yellow signals is not much reduced by using ATOs or/and TMS, and the number of red signals can be only significantly reduced with the combination of TMS and ATOs. The results reveal that the TMS rescheduling has its full effect in improving punctuality and conflict-resolving only if the TMS rescheduled timetable is well executed.

We further consider the heterogeneity of traffic. Figure 7 analyzes the probability distribution of variations in energy (left) and delay (right), split by the two train categories; and comparing Scenarios II(b), II(c) and II(d) against Scenario II(a), respectively in the rows. Positive values identify better performance compared to Scenario II(a): reduced delays, or reduced energy consumption. Scenario II(b) shows a certain amount of variation between the different trains, when compared to Scenario II(a), with the IC trains achieving a general 10% reduction in energy. Instead, the SP trains have a larger variation, linked to a small (around 10%) increase in energy. There are individual trains which have a large increase in energy, up to 50%. In any case, the two train classes have distinct behavior: IC are generally saving energy, while SP do not. Scenario II(c) has a very sharp distribution of variation (with the value at 0 out of vertical scale); i.e., energy is pretty much the same as in Scenario II(a). Scenario II(d) shows a rather homogeneous picture with both train classes having a similar variation, covering the range $\pm 25\%$.

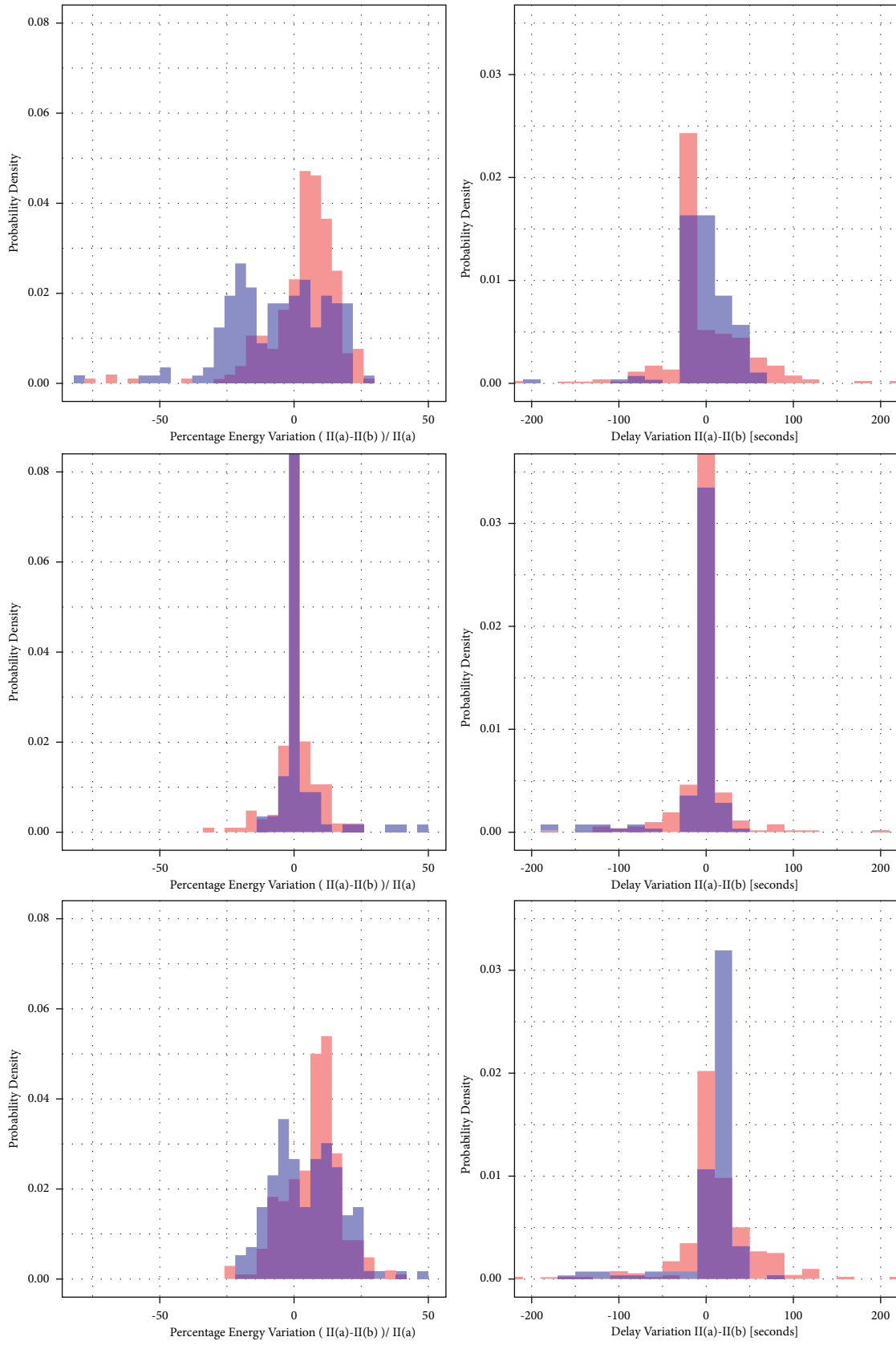


FIGURE 7: Analysis of Scenario II by energy/delays, based on train categories.

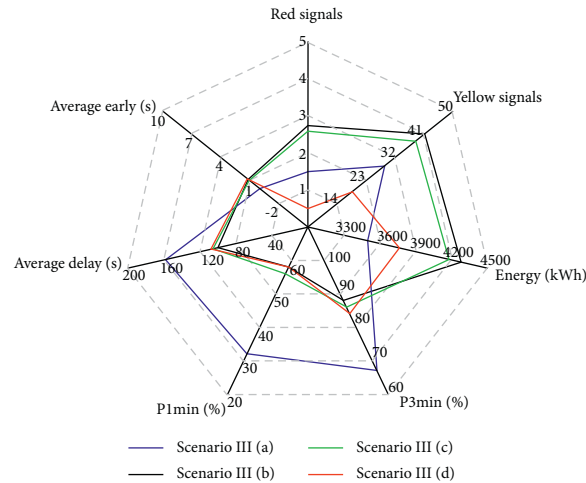


FIGURE 8: Results of Scenario III.

Summarizing, the usage of ATO affects more the energy consumption, compared to the TMS. IC new trains can achieve larger saving in energy compared to SP trains, due to their longer run and more freedom in trajectory.

Looking at delays, Scenario II(b) shows a relatively broader distribution for IC trains, ranging positive and negative values. SP trains have limited variations, in the range of $[-30, +60]$ seconds. This is possibly due to the intermediate stops which regularize the delays. Scenario II(c) shows again a very sharp distribution (with the value at 0 out of vertical scale); i.e., TMS alone is able to identify which few specific trains need to be controlled, to result in less delays. Again, IC have a slightly larger spread; some SP have a negative variation of delay (i.e., an increase), following the order optimized by TMS. Scenario II(d) shows a sharper peak for SP trains, while IC trains suffer more variation of delay. For all scenarios, the average delay variation is not very large, and only for the scenarios using TMS this increases. Overall, the two train categories result in different characteristics, when exposed to the four scenarios, with a larger spreading of effect compared to the benchmark II(a), especially when ATO is used.

4.3. Scenario III: Disrupted Traffic. Scenario III is designed to examine the impact of full/partial knowledge of the train failure by TMS/ATOs. The metric values of the scenario are reported in Figure 8.

Concerning *Punctuality*, Scenarios III(b)-III(d) lead to similar delays and punctuality: the average delays are around 50 s (32%) less than Scenario III(a), and the P1min and P3min are, respectively, improved by 25% and 18%. Such an improvement can be explained for two reasons: First, the ATO function prioritizes following rescheduled timetables. Second, the TMS changes train orders (rerouting is not considered in our case study) thus avoids delay propagation. The knowledge of train failure by TMS/ATOs seems not to affect punctuality much. In other words, in the case of the tested train failure, equipping with TMS/ATOs is beneficial

for the improvement of punctuality, no matter if the train failure is known by TMS/ATOs or not.

Concerning *energy*, Scenario III(d) consumes 11.97% and 9.85% less energy than Scenarios III(b) and III(c), respectively; and 7.60% more energy than Scenario III(a). In Scenarios III(b)-III(c), the train failure is known by only TMS or ATOs. Therefore, TMS produces nonoptimal timetables, or ATOs come up with improper driving commands, which may cause conflicts (red and yellow signals, most numerous for those two scenarios) and frequent braking-acceleration behaviors. Such frequent braking-acceleration behaviors lead to high energy costs in Scenarios III(b)-III(c). In Scenario III(d), on the contrary, both TMS and ATOs are aware of the train failure. Hence, the TMS rescheduling function and ATO driving command generation function have taken into account the train failure. This contributes to a reduction of conflicts (red signals) as well as a reduction of energy consumption. Scenario III(d) consumes 7.60% more energy than Scenario III(a). The additional energy costs come from the concurrent effect of a few factors. In Scenario III(a), TMS is not considered; therefore the traffic runs in such a way that multiple trains are hindered by the broken train and cannot drive as fast as they are supposed to do. This is an effect reducing energy for Scenario III(a). Scenario III(d) has all trains controlled by ATOs. ATOs prescribe delayed trains to run as fast as possible. Instead, in Scenario III(a) the assumed reduced performance value (modeling a human driver) is fixed to a maximum of 98%; thus trains will go at a lower maximum speed, reducing energy. Finally, the traffic control in Scenario III(a) does not take into account traffic properties and results in trains running too close, seeing repetitively yellow signal and experience multiple acceleration and braking processes to keep up with the speed of the preceding (broken, or delayed) trains. This can be evaluated by the number of yellow and red signals in Scenario III(d), which is the lowest, about half compared to Scenario III(a). This is an effect increasing energy consumption for Scenario III(a). The final balance of those three effects is what empirically evaluated.

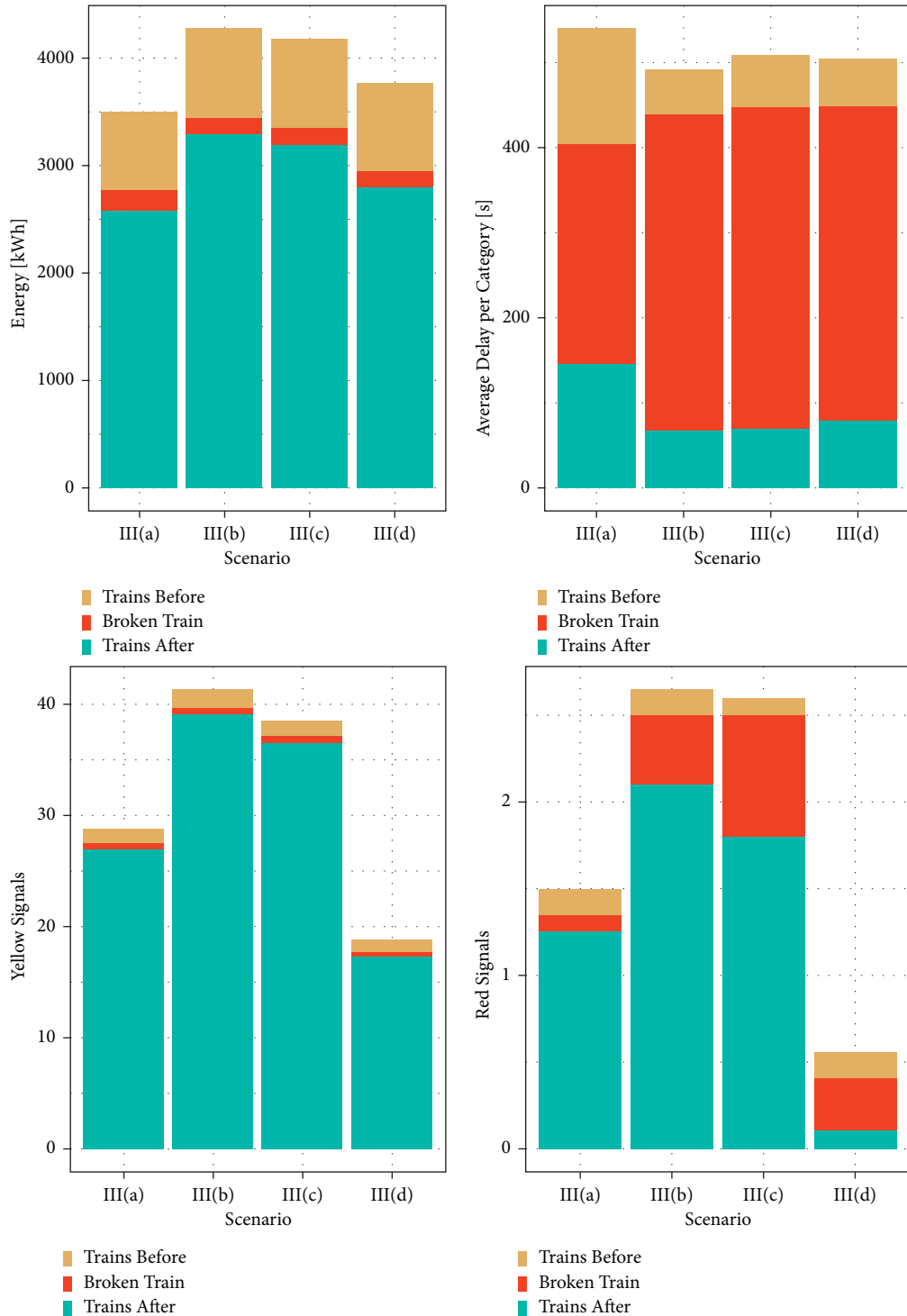


FIGURE 9: Analysis of Scenario III by energy/average delays (top), and yellow/red signals (bottom), for trains before or after the disrupted train.

Concerning *conflict-resolving*, the number of yellow and red signals in Scenario III(d) is the lowest, about half compared to Scenario III(a). That indicates the full knowledge of the train failure by TMS/ATOs is beneficial for conflict-resolving.

We again quantitatively study the heterogeneity of traffic, this time by analyzing the performance of trains

happening to run before the broken train, and those afterwards, also identifying the single contribution of the broken train itself. We report the results in Figure 9, with stacked bars identifying energy, average delays, and yellow and red signals (respectively, top and bottom row). In those cases, we do not separate into train categories IC/SP, as the data samples might be too small. We specifically focus on the

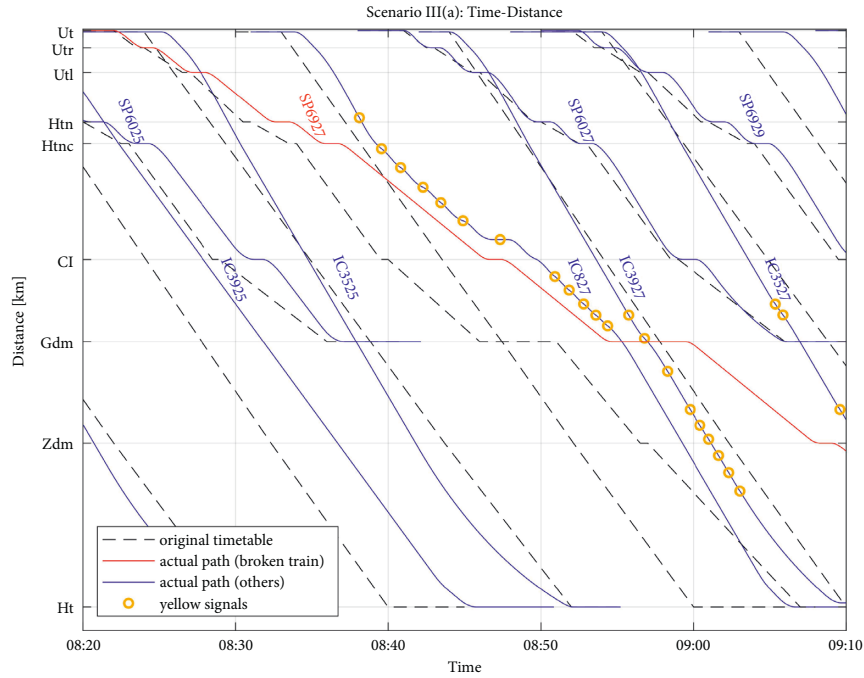


FIGURE 10: An example of time-distance paths in Scenario III (a).

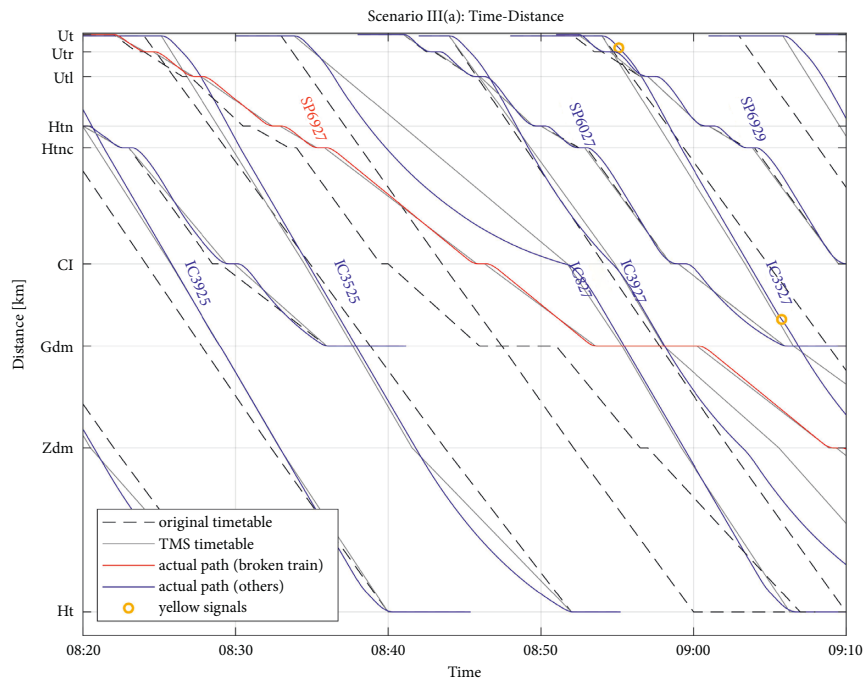


FIGURE 11: Scenario III(d): time-distance.

variation before/after the disruption (respectively, top/bottom of the stacked bar), which shows the capabilities to limit propagation of non-performance in the network.

Concerning *energy*, Scenario III(a) achieves the lowest values, and for all trains (before, and after): this is a systematic condition due to the default train dynamic in OpenTrack. The variations in the other scenarios pertain only traffic after the disruption. Scenario III(b) results in the

largest energy for the trains running after the disruption, due to multiple downstream speed adjustments (TMS aims to reduce delays, at the cost of low fluidity of traffic, being unaware of the disruption).

We study *average delays*, reporting in Figure 9 the average per each category, in order to compare the categories, which have different amounts of trains. In all cases, the broken train has the largest average delay, as it runs at

TABLE 4: Metrics of the examples in Figures 10 and 11.

	Example of Scenario III (a)	Example of Scenario III (d)
Average early (s)	0	1.45
Average delay (s)	155.32	91.04
P1min (%)	35.21	67.57
P3min (%)	66.20	85.14
Energy (kWh)	3631	3796
Yellow signals	30	13
Red signals	0	0

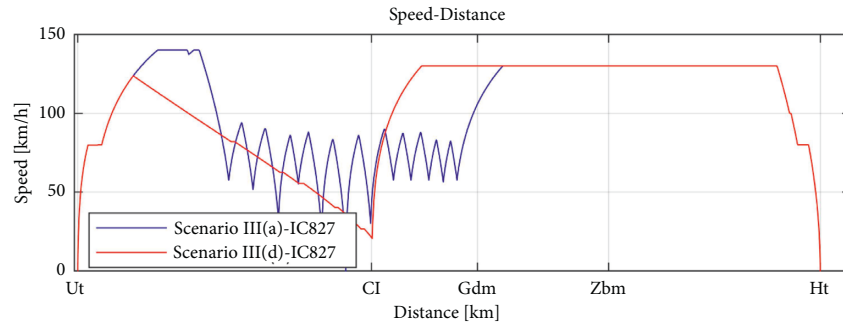


FIGURE 12: Speed profiles of IC827 for Scenarios III (a) and III (d).

reduced performance. In Scenario III(a), the traffic before and after the broken train faces a larger average delay, compared to Scenarios III(b), III(c), and III(d). In those latter cases, the traffic is reorganized by means of TMS or ATO actions, which results in increasing the average delay of the broken train, but achieving better performance for the rest of the traffic. In Scenario III(b), the delays are the lowest, but the traffic fluidity is low: the trains after (i.e., those subject to rescheduling) face many yellow and red signals, as they follow each other very closely. In other terms, a bad model of traffic characteristics used in TMS results in the ATO reactively (and not proactively) managing the speed changes. Scenario III(d) results in the smallest amount of yellow and red signals shown, mostly due to much better management of trains after the broken train, limiting propagation of non-performance. Scenarios III(b) and III(c) are relatively similar with most differences in the trains following the broken train. Scenario III(b) gives fewer red signals to the broken train, compared to Scenario III(c) at the cost of having more red signals for the downstream traffic. In other terms, some actions computed happen to be myopic with regard to their consequences.

Finally, we select a case respectively from Scenarios III(a) and III(d) for a further illustrative analysis as below. The two example cases have same entrance delays, dwell times, and train failure. Their time-distance paths are presented in Figures 10 and 11. The metric values of the two cases are shown in Table 4. Overall, the example case of Scenario III(d) has fewer delays, higher punctuality, more energy costs and fewer conflicts than Scenario III(a). This finding is the same as the results presented in Figure 8.

Again, the average energy cost of Scenario III(d) is a bit higher (4.5%) than Scenario III(a), and this further differs if we look into individual trains. Take the intercity, IC827, in

Figures 10 and 11 for instance. We also draw the speed profiles of IC827 of the two scenarios in Figure 12. In Scenario III(a), IC827 is directly hindered by the broken train SP6727 and meets many yellow signals in the section between Htnc and Gdm. The speed profile shows indeed no less than 12 braking-acceleration actions. In Scenario III(d), instead, the TMS rescheduled timetable suggests IC827 to run slow, with a long coasting between Ut and CI, and ultimately arriving at Gdm at the same time as in Scenario III(a). Deciding to perform such a long coasting is beneficial for conflict-resolving as well as for energy-saving and requires the cooperation between ATOs (to actually implement the coasting) and TMS (to know how long the coasting should be). Failure to include those both aspects explains why Scenario III(d) consumes much less energy costs than Scenarios III(b) and III(c).

5. Conclusions

This paper investigates the benefits and limitations of an integrated framework of connecting TMS and ATO in terms of punctuality, energy efficiency, and conflict-resolving with simulation methods. The simulation is built by interfacing a rescheduling tool [26, 27] and an ATO tool [28] with the realistic traffic simulation environment OpenTrack [29], and applying the proposed methodology to a real case-study in the Netherlands: the railway corridor between Utrecht and Den Bosch. The main findings are summarized below:

- (i) In Scenario I, we see that the energy costs are reduced by around 11.1% and the punctuality (P1min) is improved by 12.36% by using ATOs. It reveals that, in case of no entrance delays, the implementation of ATO systems is beneficial for

maintaining timetables and saving energy costs. The effects of ATO/TMS to the traffic are different, depending on the train type; in other terms, heterogeneity of traffic adds complexity to the evaluation and setup of the systems.

(ii) Scenario II simulates the traffic under delay circumstances. The results show that the energy reduction by equipping both TMS and ATOs (8%) is more obvious than by equipping only the TMS or ATOs (0.86%–2.29%). In other words, the energy-saving by using ATOs can be better achieved with conflict-free schedules. On the other hand, from the punctuality and conflict-resolving point of view, the punctuality at 1 min level and 3 min level of Scenario II(d) is, respectively, improved by 8.77% and 2.20%. The red signals of Scenario II(d) are reduced by 57.89%. In other words, by equipping both TMS and ATOs, the punctuality is much improved with less early arrivals and less delays. On the contrary, in case that only TMS or ATOs are equipped (Scenarios II(b)–II(c)), either the punctuality is not improved, or the conflict-resolving is not improved. We can conclude that the TMS rescheduling has its full effect in improving punctuality and conflict-resolving only if trains are able to follow TMS rescheduled timetables.

(iii) Scenario III examines the impact of full/partial knowledge of the train failure by TMS/ATO. In the case of our tested train failure (half engine loss of one single train), equipping with TMS/ATO is beneficial for the improvement of punctuality (the average delay is reduced by 32%), no matter if the train failure is known by TMS/ATO or not. But the knowledge of train failure by TMS/ATO is beneficial for conflict-resolving and energy-saving. The benefit is more obvious if both TMS and ATOs have knowledge of train failure. Compared to the scenarios in which only TMS or ATOs know the train failure, the energy cost is reduced by around 9.85%–11.97% and the number of red signals is reduced by 68.25%–81.82% in the scenario that both TMS and ATOs have knowledge of train failure. However, different train dynamics and failures (mass, or failure, or driving strategies, etc.) may have different impacts on the performance of the integrated TMS/ATO framework. Therefore, it is also an interesting topic to explore the impacts of different knowledge of train dynamics in the future.

In addition, the results also prove the usefulness of the proposed simulation methodology to analyze the benefits and limitations of the integrated traffic and train control in a stochastic and dynamic environment. The simulations and evaluations done in this work can be further developed and improved in several aspects. To start with, the simulation framework can be extended to study other failures that affect the train operations; the simulation framework can be extended with the implementation of an online multiple open-

loop or a closed-loop TMS, sending updated rescheduled timetables based on the current traffic state. Further, the speed optimization function could be moved to a central unit, so that multiple trains' speed profiles are optimized together, which is expected to reduce the amount of yellow and red signals. Both TMS and ATO could be extended to include also other performance indicators, such as passenger travel time, or riding comfort. The more precise estimation of impact on traffic flow from the interaction between ATO and TMS could support other design choices, like infrastructure (block layout), or timetable (buffer time) [39]. Finally, this simulation approach proved how the microscopic timetable has benefits in solving local conflicts and reducing delays. This assumes though that some processes and barriers in real life can be overcome, which include availability of offline and real time data, with a good microscopic quality for timetable and actual operations, availability of train dynamic parameters, specific reaction time and delay of the ATO system, and description of the specific rules used by the signalling system. Moreover, it assumes a fast-enough loop to disseminate information between the involved users.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the NWO (Netherlands Organisation for Scientific Research) Rubicon Grant (grant no. 019.172EN.010); and Rubicon Grant (grant no. 019.192EN.014); the Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, with Project 501100001711; the Swiss Innovation Agency Innosuisse; and the Swiss Competence Center for Energy Research SCCER Mobility; and Shanghai Sailing Program (grant no. 21YF1450200).

References

- [1] A. D'Ariano, "Innovative decision support system for railway traffic control," *IEEE Intelligent Transportation Systems Magazine*, vol. 1, no. 4, pp. 8–16, 2009.
- [2] J. Yin, T. Tang, L. Yang, J. Xun, Y. Huang, and Z. Gao, "Research and development of automatic train operation for railway transportation systems: a survey," *Transportation Research Part C: Emerging Technologies*, vol. 85, pp. 548–572, 2017.
- [3] M. Luethi, W. Ulrich, F. Laube, and G. Medeossi, "Rescheduling and train control: a new framework for railroad traffic control in heavily used networks," in *Proceedings of the 86th Transportation Research Board Annual Meeting*, Switzerland, 2007.

- [4] M. Mazzarello and E. Ottaviani, "A traffic management system for real-time traffic optimisation in railways," *Transportation Research Part B: Methodological*, vol. 41, no. 2, pp. 246–274, 2007.
- [5] X. Rao, M. Montigel, and U. Weidmann, "A new rail optimisation model by integration of traffic management and train automation," *Transportation Research Part C: Emerging Technologies*, vol. 71, pp. 382–405, 2016.
- [6] F. Corman and E. Quaglietta, "Closing the loop in real-time railway control: framework design and impacts on operations," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 15–39, 2015.
- [7] V. Cacchiani, D. Huisman, M. Kidd et al., "An overview of recovery models and algorithms for real-time railway rescheduling," *Transportation Research Part B: Methodological*, vol. 63, pp. 15–37, 2014.
- [8] F. Corman and L. Meng, "A review of online dynamic models and algorithms for railway traffic management," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1274–1284, 2014.
- [9] X. Luan, Y. Wang, B. De Schutter, L. Meng, G. Lodewijks, and F. Corman, "Integration of real-time traffic management and train control for rail networks - Part 1: optimization problems and solution approaches," *Transportation Research Part B: Methodological*, vol. 115, pp. 41–71, 2018.
- [10] X. Luan, Y. Wang, B. De Schutter, L. Meng, G. Lodewijks, and F. Corman, "Integration of real-time traffic management and train control for rail networks - Part 2: extensions towards energy-efficient train operations," *Transportation Research Part B: Methodological*, vol. 115, pp. 72–94, 2018.
- [11] F. Corman, "Interactions and equilibrium between rescheduling train traffic and routing passengers in microscopic delay management: a game theoretical study," *Transportation Science*, vol. 54, no. 3, pp. 785–822, 2020.
- [12] S. Zhan, S. C. Wong, P. Shang, Q. Peng, J. Xie, and S. M. Lo, "Integrated railway timetable rescheduling and dynamic passenger routing during a complete blockage," *Transportation Research Part B: Methodological*, vol. 143, pp. 86–123, 2021.
- [13] S. Zhan, P. Wang, S. C. Wong, and S. M. Lo, "Energy-efficient high-speed train rescheduling during a major disruption," *Transportation Research Part E: Logistics and Transportation Review*, vol. 157, Article ID 102492, 2022.
- [14] On-TIME, *Best Practice, Recommendations and Standardisation*, Macomb, IL, USA, 2014.
- [15] E. Quaglietta, P. Pellegrini, R. M. P. Goverde et al., "The ON-TIME real-time railway traffic management framework: a proof-of-concept using a scalable standardised data communication architecture," *Transportation Research Part C: Emerging Technologies*, vol. 63, pp. 23–50, 2016.
- [16] A. Fernández-Rodríguez, A. Fernández-Cardador, and A. P. Cucala, "Real time eco-driving of high speed trains by simulation-based dynamic multi-objective optimization," *Simulation Modelling Practice and Theory*, vol. 84, pp. 50–68, 2018.
- [17] L. Zhu, L. Chen, C. Roberts, and N. Zhao, "Dynamic trajectory optimization design for railway driver advisory system," *IEEE Intelligent Transportation Systems Magazine*, vol. 10, no. 1, pp. 121–132, 2018.
- [18] O. Yalcinkaya and G. M. Bayhan, "A feasible timetable generator simulation modelling framework for train scheduling problem," *Simulation Modelling Practice and Theory*, vol. 20, no. 1, pp. 124–141, 2012.
- [19] K. Panou, P. Tzieropoulos, and D. Emery, "Railway driver advice systems: evaluation of methods, tools and systems," *Journal of Rail Transport Planning & Management*, vol. 3, no. 4, pp. 150–162, 2013.
- [20] P. Wang, R. M. P. Goverde, and J. Van Luipen, "A connected driver advisory system framework for merging freight trains," *Transportation Research Part C: Emerging Technologies*, vol. 105, pp. 203–221, 2019.
- [21] M. Montigel, "Operations control system in the lotschberg base tunnel," *European Rail Technology Review*, vol. 49, 2009.
- [22] A. Cubris, "What Makes Greenspeed Unique," 2013, <http://www.cubris.dk/what-is-greenspeed/what-makes-greenspeed-unique/>.
- [23] Li Yang, L. Tomas, and P. Leander, "Achieving energy-efficiency and on-time performance with driver advisory systems," in *Proceedings of the 2013 IEEE International Conference on Intelligent Rail Transportation Proceedings*, pp. 13–18, IEEE, Beijing, China, September 2013.
- [24] J. Yin, T. Tang, L. Yang, Z. Gao, and B. Ran, "Energy-efficient metro train rescheduling with uncertain time-variant passenger demands: an approximate dynamic programming approach," *Transportation Research Part B: Methodological*, vol. 91, pp. 178–210, 2016.
- [25] X. Yang, A. Chen, B. Ning, and T. Tang, "A stochastic model for the integrated optimization on metro timetable and speed profile with uncertain train mass," *Transportation Research Part B: Methodological*, vol. 91, pp. 424–445, 2016.
- [26] X. Luan, F. Corman, and L. Meng, "Non-discriminatory train dispatching in a rail transport market with multiple competing and collaborative train operating companies," *Transportation Research Part C: Emerging Technologies*, vol. 80, pp. 148–174, 2017.
- [27] X. Luan, B. De Schutter, L. Meng, and F. Corman, "Decomposition and distributed optimization of real-time traffic management for large-scale railway networks," *Transportation Research Part B: Methodological*, vol. 141, pp. 72–97, 2020.
- [28] P. Wang and R. M. P. Goverde, "Multiple-phase train trajectory optimization with signalling and operational constraints," *Transportation Research Part C: Emerging Technologies*, vol. 69, pp. 255–275, 2016.
- [29] A. Nash and D. Huerlimann, "Railroad simulation using opentrack," *WIT Transactions on The Built Environment*, vol. 74, 2004.
- [30] ICT-53-2020, "5GRail Consortium. 5GRail: 5G for Future RAILway mobile Communication System," p. 2021, 2021, <https://cordis.europa.eu/project/id/951725>.
- [31] P. Wang, A. Trivella, R. M. P. Goverde, and F. Corman, "Train trajectory optimization for improved on-time arrival under parametric uncertainty," *Transportation Research Part C: Emerging Technologies*, vol. 119, Article ID 102680, 2020.
- [32] W. Raimond Matthias, "The opentrack api: case study using opentrack for benchmarking automatically generated timetable scenarios specified by the transport service intention," in *Proceedings of the nextRail Conference in EPFL*, Lausanne, September 2017.
- [33] P. Peter and P. Howlett, "Optimal driving strategies for a train journey with speed limits," *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics*, vol. 36, no. 1, pp. 38–49, 1994.
- [34] A. Albrecht, P. Howlett, P. Pudney, X. Vu, and P. Zhou, "The key principles of optimal train control-Part 1: formulation of the model, strategies of optimal type, evolutionary lines,

- location of optimal switching points,” *Transportation Research Part B: Methodological*, vol. 94, pp. 482–508, 2016.
- [35] A. Albrecht, P. Howlett, P. Pudney, X. Vu, and P. Zhou, “The key principles of optimal train control-Part 2: e,” *Transportation Research Part B: Methodological*, vol. 94, pp. 509–538, 2016.
- [36] J. P. Powell and R. Palacín, “Passenger stability within moving railway vehicles: limits on maximum longitudinal acceleration,” *Urban Rail Transit*, vol. 1, no. 2, pp. 95–103, 2015.
- [37] E. Quaglietta, F. Corman, and R. M. P. Goverde, “Stability analysis of railway dispatching plans in a stochastic and dynamic environment,” *Journal of Rail Transport Planning & Management*, vol. 3, no. 4, pp. 137–149, 2013.
- [38] F. Corman, E. Quaglietta, and R. M. P. Goverde, “Automated real-time railway traffic control: an experimental analysis of reliability, resilience and robustness,” *Transportation Planning and Technology*, vol. 41, no. 4, pp. 421–447, 2018.
- [39] E. Quaglietta, “A simulation-based approach for the optimal design of signalling block layout in railway networks,” *Simulation Modelling Practice and Theory*, vol. 46, no. 4–24, 2014.