



Integrated speed modeling and traffic management to precisely model the effect and dynamics of temporary speed restrictions to high-speed railway traffic

Sihui Long^a, Lingyun Meng^{b,*}, Yihui Wang^c, Jianrui Miao^c, Xiaojie Luan^b,
Francesco Corman^d

^a School of Transportation Engineering, Kunming University of Science and Technology, 650504, Yunnan, China

^b School of traffic and transportation, Beijing Jiaotong University, 100044, Beijing, China

^c State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, 100044, Beijing, China

^d Institute for Transport Planning and Systems (IVT), ETH Zürich, Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland

ARTICLE INFO

Keywords:

Train rescheduling
Train control
Integrated optimization
Mixed-integer linear programming (MILP)
High-speed railway
Temporary Speed Restriction (TSR)

ABSTRACT

This paper investigates the integrated optimization problem of train rescheduling and train control for high-speed railway lines, where perturbations occur and cause Temporary Speed Restriction (TSR) to trains.

We consider microscopic details (i.e., block sections) for ensuring the feasibility of the solution from a train signaling point of view, and an even higher level of detail, to accurately represent the train speed profiles. Running time and headway time are variable, at the same time depending from, and affecting, traffic. We optimize train speed profiles by considering vehicle and infrastructure constraints (e.g., traction, slopes). The model naturally considers the transition along the normal, disturbed and the recovery operation periods.

A mixed-integer nonlinear programming (MINLP) model is first developed to simultaneously optimize train orders, routes, and departure and arrival times, as well as train speed profiles, aiming at reducing total train deviation time. The MINLP model is difficult to solve; thus we further reformulate it into a mixed-integer linear programming (MILP) model by means of piecewise linear approximation. A two-step approach is designed to speed up the solving procedure of the MILP model: first estimate the upper/lower bounds of train speeds and then solve the MILP model based on the estimated bounds of train speeds.

Three instances (i.e., a small-scale line, a medium-scale line, and a large-scale network) are used to highlight the performance of the approach, verify the benefit of the integration, and its dependence on the parameters used. According to the experimental results, our integrated optimization method leads to an average improvement of 3%-36% in solution quality, compared with the integrated approach without train rerouting measure. Moreover, the integrated optimization method outperforms the sequential approach, achieving 6%-9% improvement in solution quality.

1. Introduction

In the past decade, high-speed railway has been well-developed, and its network scale has been significantly expanded in China. High-speed railway plays an important role in the future sustainable transport by offering safe, fast, punctual, reliable, and

* Corresponding author.

E-mail addresses: LongSH@kust.edu.cn (S. Long), lymeng@bjtu.edu.cn (L. Meng), yihui.wang@bjtu.edu.cn (Y. Wang), jrmiao@bjtu.edu.cn (J. Miao), xjluan@bjtu.edu.cn (X. Luan), francesco.corman@ivt.baug.ethz.ch (F. Corman).

<https://doi.org/10.1016/j.trc.2023.104148>

Received 11 October 2022; Received in revised form 11 February 2023; Accepted 19 April 2023

Available online 11 May 2023

0968-090X/© 2023 Elsevier Ltd. All rights reserved.

comfortable services to passengers. Its high-speed feature shortens the travel time of passengers, making the travel of passengers efficient and convenient; at the same time, it also brings challenges of keeping safety and high punctuality and reliability of the train operations.

In daily operations, various perturbations (e.g., bad weather condition, infrastructure failure, and temporary speed restriction) often occur, causing delays to the planned train timetable. Among the perturbation sources, temporary speed restriction (TSR) occurs more commonly in high-speed railway operations, taking up 85% among all the disruptions (Xu et al., 2016). It is usually applied when the track or infrastructure is not in a proper condition for normal high-speed operations. To ensure the operation safety, the trains, passing through some specific block sections or segments (i.e., the TSR areas) at a specific time period (i.e., when the TSR effects), are required to reduce their speed (i.e., respecting the speed limited by the TSR). TSR is a time-limited perturbation, the train speed dynamically depends on the train arrival and departure times at the TSR area. Thus, it is necessary to consider the train rescheduling and train control at the same time for the real-time train operation under TSR.

In the presence of disruptions, the train timetable rescheduling and train control are the keys to improve the efficiency of real-time traffic management. However, the problems of train rescheduling and train control were typically studied separately. For resolving train conflicts (safety) as quickly as possible, recovering delays (punctuality), and avoiding cancellations (reliability), relevant actions (of dispatching and driving) are desired to be optimized in an integrated way and performed in an efficient manner. This is particularly crucial for high-speed railway due to its features and market positioning (Ning et al., 2019).

Integrated optimization of train rescheduling and train control is important to ensure efficient high-speed rail operation with good quality, which has attracted much attention (e.g., Xu et al., 2017; Luan et al., 2018a,b; Zhou et al., 2020; Zhan et al., 2022). The train control problem contains lots of non-linear constraints, thus the studies simplified the optimization of train speed trajectories by means of managing the speed level (e.g., Xu et al., 2017) or discrete speed (e.g., Luan et al., 2018a,b; Zhan et al., 2022), picking the speed profiles from a set of candidate options (e.g., Zhou et al., 2020). However, the non-linear effects on train control (i.e., train speeds) are ignored. Therefore, the rescheduled timetable and/or the train speed profile may not be implementable in the real-world. Moreover, the previous approaches avoid the possibility to vary speed at any moment in time and space. In addition, most literature deals with disruptions in a all or nothing approach, i.e., the approach focuses on the disturbed operations only (e.g., by giving a-priori affected trains). However, under TSR perturbation, we need to model the speed change according to the rescheduled timetable and TSR conditions. Further, we need to model the transition from normal to disturbed operations, and finally to the full recovery of normal train operations. Thus, TSR perturbation requires the approach to represent a wider range of dynamical behavior, concerning traffic and speeds which are not known a-priori.

This paper therefore focuses on the integrated optimization approach of train rescheduling and train control. Our method follows a much more detailed description of the two problems than most of the literature. We use optimization methods, for dealing with TSR disruptions occurring during high-speed railway operations. A mixed-integer nonlinear programming (MINLP) model is proposed to optimize the train orders, routes, departure and arrival times, and the speed trajectories at the same time. The model considers to minimize the total deviation time, in which deviation time is defined as the absolute value of deviation between the planned arrival/departure time and the rescheduled time value. In the model, we describe the problem with microscopic infrastructure details (i.e., block sections) for ensuring the operational feasibility of the train rescheduling solutions, and we further represent each block section by a set of discrete-space intervals with an equal length for calculating the train speed profiles more precisely. Moreover, the train running time and the headway time are considered to be variables with a high level of detail, depending on the real train speeds and the blocking time. The train speed trajectories are optimized with the consideration of the real train characteristics and infrastructure characteristics. Since the MINLP model is difficult to solve, we further reformulate it into a mixed-integer linear programming (MILP) model by means of piecewise linear approximation. To speed up the solving procedure of the MILP model, a two-step approach is further designed: first estimate the upper and lower bounds of the speed for each train on each block section (for reducing solution space) and then solve the MILP model based on the estimated bounds of train speeds (exploring the reduced solution space). We consider two instances to examine the proposed methods: a small-scale instance and a realistic instance.

Experiments on the small-scale line are conducted to evaluate the settings of the parameters related to the piecewise approximation and the length of the space intervals (i.e., balancing the approximation error, the problem size, and the computational efficiency), to show the impacts of the TSR-related parameters (e.g., the duration and the TSR speed limit) on the results, as well as to demonstrate the benefit of the train rerouting and reordering measures. With the parameter settings verified in the case study of the small-scale instance, we further conduct experiments on the realistic medium-scale instance, to comparatively assess the performance of the proposed model, the two-step approach, and the sequential adjustment approach, showing the benefit of integrating train rescheduling and train control. Moreover, a large-scale network is built to compare the performances of the proposed integrated optimization method on instances with different scales. According to the experimental results, the proposed integrated MILP method can obtain feasible solutions within 180 s, showing it possible to effectively integrate those two problems that are most often tackled separately. Meanwhile, the goal of deviation reduction by integrating train rescheduling and train control can be achieved.

The rest of the paper is organized as follows. A comprehensive literature review is proposed in Section 2, followed by the problem statement and assumptions (Section 3). Section 4 presents a mixed-integer linear programming (MILP) model and Section 5 proposes the solution approaches. Section 6 shows the experimental results and discussions and Section 7 ends the paper with conclusions and topics for further research.

2. Literature review

The train rescheduling problem has been extensively studied in the past few decades. It aims to generate a feasible train schedule with good quality in a reasonable time. In the literature, there are some optimization models and algorithms available for addressing the train rescheduling problem, see the surveys and review papers by [Cacchiani et al. \(2014\)](#), [Corman and Meng \(2014\)](#), [Fang et al. \(2015\)](#) and [Hansen and Pachl \(2014\)](#) for more information. The dispatching measures commonly used in the train rescheduling models include: re-timing by changing the train arrival and departure times; and re-ordering by changing the arrival and/or departure orders of trains ([Hansen, 2010](#); [Cacchiani et al., 2014](#)). Some studies consider also the measure of rerouting, i.e., adapting the train routes (see e.g., [Meng and Zhou, 2014](#)). Re-routing measure is particularly interesting in case of disruption, because it allows to enlarge the solution space considerably, enabling much more reordering options, which can reduce the delays and their propagation. The train rescheduling problem has been tackled with a large variety of objective functions, ranging from delay recovery, to minimizing the total train delay, the maximum train delay, the consecutive train delays, passenger travel time or delay, and priority-weighted timetable deviations ([Hansen and Meng, 2019](#)).

Here, we review some representative studies on the train rescheduling problem in the literature, where train rerouting is considered in the rescheduling process. For the following literature on the train rescheduling problem, all of them only considers fixed train speed and linear constraints. Based on cumulative flow variables, [Meng and Zhou \(2014\)](#) developed an integer programming model for re-timing and re-routing trains on an N -track network to minimize the total deviation time of trains from the planned timetable. [D'Ariano et al. \(2008\)](#) proposed an alternative-graph based model for addressing the train rescheduling problem. The authors also designed a local search algorithm for rerouting trains to improve train punctuality. [Pellegrini et al. \(2014\)](#) proposed a mixed-integer linear programming (MILP) model to reroute and reschedule trains, with the aim of minimizing delays in case of perturbations. [Corman et al. \(2010\)](#) developed a novel tabu search algorithm to solve the train routing and rescheduling problem and showed the good performance of rerouting to minimize consecutive delays. [Lusby et al. \(2013\)](#) designed a resource-based set-packing formulation for real-time junction train routing problem and proposed a branch-and-price based solution approach to solve the problem. [Caimi et al. \(2012\)](#) proposed a predictive control method based on a bi-level tree conflict graph to reschedule trains (and locally reroute trains) in complex central railway station areas.

The train control problem is to determine the train speed profiles based on the arrival and departure time constraints (specified by train schedules), with the aim of minimizing energy consumption. The train speed is typically represented as a fine-grained vector function of space or time, resulting in discrete-space or discrete-time model, respectively. The train control studies mostly focus on the single-train speed trajectory optimization problem based on a pre-defined train schedule, i.e., no adaptation and optimization to the train schedule (e.g., [Howlett, 2000](#); [Franke et al., 2002](#)). [Wang et al. \(2013\)](#) proposed a pseudospectral method to solve the train speed profile optimization problem and reformulated the nonlinear constraints by approximating with piecewise affine functions. [Lu et al. \(2013\)](#) proposed genetic algorithm, dynamic programming, and ant colony optimization algorithms to solve the energy-efficient train speed trajectories optimization problem. In order to deal with real-world complicated running conditions, [Ko et al. \(2004\)](#) improved Bellman's dynamic programming method to optimize train running trajectory, which could obtain the optimal speed trajectory in an acceptable computation time. Based on the given initial speed and time duration for each train on each subsection, [Ye and Liu \(2017\)](#) proposed a non-linear programming formulation to solve the optimal control problem of multiple trains. [Goverde et al. \(2020\)](#) developed a computational framework based on pseudospectral method and Pontryagin's Maximum Principle to improve the solution accuracy for the optimal train control problem.

The train rescheduling problem and the train control problem are well-studied separately, but a gap still exists with regard to their integration. The parameters used in the timetable rescheduling and train control models are often determined based on the normal operations and without considering the interactions between them. Most existing train rescheduling models use the train-related parameters (e.g., train running times on block sections and safety headway times) based on experiences or estimated values. The train speed is completely ignored in the model formulation, since the speed-related parameters/variables are all represented by time parameters/variables. For example, the train running time on each block section is usually determined based on the minimum train running time and the certain time supplements, and the minimum running time further depends on the length of the block section and the maximum train speed on the block section ([Meng and Zhou, 2014](#)). These train-related parameters are pre-defined and considered to be constant in the train rescheduling models.

However, the train running time and the safety headway time are in fact variables, closely related to the real train speed. The train speeds depend on many factors, which in turn depend on many factors from infrastructure and vehicle, respectively line characteristics (slopes) and train characteristics (resistance, traction power). Moreover, speed depends on traffic by means of headways, which are dependent on the train positions, and are regulated by means of the signaling system. Finally, train speeds are also adjusted in real time in case of disruptions. The most common case is that a temporary restriction on speeds is issued (due to bad weather, maintenance, etc.) for ensuring safety. TSR causes extra running times, and further cascading effects ([van der Kooij et al., 2017](#)), thus causing delays with high probability. The headway between two trains is the time interval between them. The headway cannot be reduced below a minimum headway time, for safety reasons. To precisely compute headways, the blocking time theory is used ([Pachl, 2009](#)). The blocking time is the time interval in which a block section is exclusively allocated to a train and therefore blocked for other trains; it lasts from the moment of issuing a train movement authority to the moment that it becomes possible to issue a movement authorization to another train to enter the same section. The blocking time, and therefore the headway, again depends on speed. One typical issue here is that the calculated minimum running time (i.e., maximum running speed) for normal situations does not apply to the situations where temporary speed restrictions are implemented; moreover while the speed restriction is fixed in time and space, traffic is not, as it subject to delays and delay propagation. It is therefore hard to

know in advance which train on which block section will be affected by the temporary speed restriction. This results in a strong interconnection between traffic control: reordering and rerouting decisions (which train goes when, typically assuming some speed profiles); and train control: speed decisions (how fast can a train go, typically assuming a given order and route). The vast majority of the literature, though, considers those two problems separately.

Most existing train control studies consider a pre-determined train schedule, i.e., given the time points of each train departing from and arriving at each station, to optimize a specific train speed profiles, for instance minimizing the energy consumption. The train-related parameters used in the train rescheduling model and the pre-determined train schedules used in the train control model would largely affect the results of the two problems. Inappropriately pre-defined parameters or schedules may cause a mismatch between what is supposed to happen, and what can actually happen in real life when implementing. Possible drawbacks are large delays, the impossibility to follow a planned speed profile, resulting in further delays and energy losses, and need for further downstream adjustment as the planned schedule is not feasible, given the prevailing conditions. This hides the potential performance improvement of train operations in both service quality and energy efficiency. Therefore, integration of the train rescheduling problem and the train control problem is needed, rather than tackling the two problems separately, in order to achieve better performance of train operations.

In recent years, some studies have paid attention to the interaction of the two problems. Most of them adopt a sequential or iterative optimization approach. The sequential approach is a two-stage method to study the two problems, i.e., first solving one problem to obtain the solutions (i.e., train speed trajectories or train schedules) and then using the obtained solutions as input for the other problem (e.g., [Albrecht, 2009](#); [D'Ariano et al., 2010](#)). The iterative method is also a two-stage approach for the two problems but with feedback control between them (e.g., [Mazzarello and Ottaviani, 2007](#); [Lüthi, 2009](#)). Some other studies solve the train rescheduling problem by fixing a-prior set of speed trajectories. For example, with the given timetable, by selecting the train speed profiles from a finite set, [Corman et al. \(2009\)](#) proposed a model based on the alternative graph for the conflict detection and resolution problem with subject to the constraints related to the energy consumption. Based on discrete-time model predictive control, [Caimi et al. \(2012\)](#) first considered a set of alternative blocking-stairways, in which the blocking-stairways are combined with alternative routes (each route corresponding to a speed trajectory) and arrival/departure times, and then selected one blocking-stairway among the alternatives for each train. [Wang and Goverde \(2016\)](#) developed a multiple-phase optimal control model to calculate the optimal train trajectories to minimize the train delay time and energy consumption. Such sequential or iterative methods can hardly generate the globally optimized solutions of the two problems, or guarantee the optimality in either train rescheduling or train control problem ([Luan et al., 2018a](#)).

The integrated optimization method has been widely used in the train operation management problems, which could achieve superior performance compared with sequential optimization approach ([Schöbel, 2017](#)). Many studies tried to integrated multiple operation strategies to achieve the benefits of system optimization for metro lines. For example, there are some researches for integrating optimization of the following operation strategies: energy-efficient train timetabling and train control (e.g., [Li and Lo, 2014a,b](#); [Mo et al., 2019](#)); train timetabling and rolling stock circulation planning ([Wang et al., 2021](#)); service frequency planning, timetabling, rolling stock planning, and train speed profiles for a subway line ([Mo et al., 2021](#)). Compared with the railway train operation, the metro train operation has a structure of high homogeneity, which does not need to consider complex dispatching measures like rerouting and reordering. Moreover, most metros are operated underground, thus, the TSR case is uncommon in metro operation. Recently, the integrated optimization method for railway operations management has been attracting much attention. For example, some researchers proposed integrated optimization methods for the following railway operation phases: timetabling and track maintenance scheduling (e.g., [Luan et al., 2017](#); [D'Ariano et al., 2019](#); [Zhang et al., 2019](#)); passenger demand, line planning and train timetabling ([Meng and Zhou, 2019](#)); train rescheduling and passenger reassignment ([Hong et al., 2021](#)).

Recently, the integration of the train rescheduling and train control decisions for railway traffic has received much attention. [Luan et al. \(2018a,b\)](#) developed three optimization approaches to solve the integrated problems of train rescheduling and train control. The authors considered re-timing and re-ordering measures of train rescheduling and selected the train speed from a set of discrete speed values. For a set of random delays, their proposed method can obtain feasible solutions for a single direction along a 50 km corridor with 15 trains within 3 min. [Xu et al. \(2017\)](#) proposed a MILP model based on the alternative graph, incorporating train rescheduling with speed management, for high-speed railway lines under TSRs. A two-step approach was designed to solve the proposed model: the approach first generates an initial solution by adopting the full train speed and fixing it; then, the approach improves the obtained initial solution by considering more speed options. [Zhou et al. \(2020\)](#) proposed an integrated optimization model of train timetable rescheduling and train speed profiles optimizing under disturbance. The decisions of re-timing, re-ordering, and speed trajectory selection are considered in the model. Compared with them, our method gives a more detailed description of the train speed trajectories, which allows to model the speed change according to the rescheduled timetable and TSR conditions.

Just a few real-time rescheduling models in the literature directly showed to be able to optimize the train operations under TSR. [Xu et al. \(2017\)](#) is one of them that propose optimization methods to tackle the train rescheduling problem incorporating the speed management for high-speed railways under TSRs. However, the train speed trajectories were not calculated precisely, but represented by different speed levels which further affect the running times of trains. [Dong et al. \(2021\)](#) proposed an integration system of operation control and train rescheduling. Moreover, they employed the case of temporary speed restriction to analyze the principle of why and how the integration system can promote the recovery ability of carrying capacity for a high speed railway. Nevertheless, there was barely any study modeling speed variations with a sufficient level of detail (i.e., resolution of meters and seconds).

[Table 1](#) summarizes some relevant studies on the train rescheduling problem, in terms of dispatching measures, mathematical formulations (including objective and model structure), dispatching scenario, and model efficiency (i.e., scale of the problem

Table 1

Summary of the research on the train timetable rescheduling.

Publications	Dispatching measures	Objective function	Model structure	Dispatching scenario	Calculation of speed trajectories	Infrastructure details	Scale of the problem solved/reported
Pellegrini et al. (2014)	RT, RO, RR	Minimize the maximum delay of any train; Minimize the total delay	MILP	PD	No, but the speed variation dynamics considered	Track-circuits details of the block sections	A control area, 589 trains
D'Ariano et al. (2008)	RT, RO, RR	Minimize the maximum consecutive delay	MILP	PD	NO	NO	A network of 50 km, 191 block sections, and 52 trains
Corman et al. (2010)	RT, RO, RR	Minimize the maximum and average consecutive delays	MILP	PD	NO	NO	A network of 50 km, 191 block sections, and 40 trains
Meng and Zhou (2014)	RT, RO, RR	Minimize the total completion time of all involved trains	MILP	PD	NO	NO	A network of 287.7 km, 97 block sections, and 40 trains
Luan et al. (2018a)	RT, RO	Minimize the total delay time of all trains at all visited stations	MINLP, MILP	PD	YES	Grade of the block sections	A railway network of 50 km, 42 block sections, and 15 trains
Xu et al. (2017)	RT, RO	Minimize the secondary delay	MILP	TSR	NO, but the speed level considered	NO	A high-speed railway line of 500 km, 374 block sections, and 20 trains
Zhou et al. (2020)	RT, RO	Minimum weighted sum of delay time and energy consumption	MILP	PD	Not, but the speed trajectories selection considered	Grade of the block sections	A high-speed railway line with 6 stations, and 7 trains
Long et al. (this work)	RT, RO, RR	Minimize the total deviation times	MILP	TSR	YES	Grade, curve, and tunnel of the block sections	A high-speed railway line of 431 km, 247 block sections, and 15 trains

* Symbol descriptions for [Table 1](#): Dispatching measures: Retiming (RT), Reordering (RO), Rerouting (RR); Model structure: Mixed-Integer Linear Programming (MILP), Mixed-Integer Non-Linear Programming (MINLP); Dispatching scenario: Primary Delay is given for the trains (PD), Temporary Speed Restriction (TSR).

solved/reported), particularly focusing on whether the train speed trajectory is optimized while dispatching trains and which level of infrastructure detail is considered in the calculation/optimization. Infrastructure detail is most often very rough: space is assumed featureless and homogeneous. In reality, curves, slopes, and tunnels affect the maximum acceleration and/or the maximum speed a train can reach. Only when those infrastructure characteristics are precisely identified in their positions and effects, the speed profile is actually precise enough to be implemented.

As the table shows, the studies mostly focus on the perturbations existing caused by primary delays on some trains, rarely tackling the disruptions caused by TSR. Thus, the existing approaches could only optimize the rescheduled timetable with a predetermined disturbance (i.e., primary delay), rather than rescheduling train timetable in a disrupted state (i.e., beginning of TSR) and the transition from a disrupted state (i.e., ending of TSR). Moreover, these studies mostly calculate speed trajectories based on a static offline consideration of infrastructure characteristics and disruption, without integrating rescheduled timetable in the calculation and adjustment of speed trajectories. Thus, some dynamic-related train control constraints, such as precise calculation of running time and safety headway time, cannot be included. When incorporating speed management into train rescheduling, the available studies adopt only the dispatching measures of re-timing and re-ordering to achieve good computational results; none of those integrated optimization models considers train rerouting, which is proven to be effective and can lead to a considerable reduction in train delays (see [Meng and Zhou, 2014](#)). Thus, rerouting measure is needed as it allows to manage traffic on a wider solution space and accommodate larger-scale changes required by the speed reduction and its effect to traffic (including additional possibilities for reordering and usage of resources).

This paper makes the following contributions:

- We investigate the integrated optimization problem of train rescheduling and train control, in order to deal with TSR disruptions that occur in the operation of high-speed railways. The method describes the two problems in a detailed way. Further, our method considers the operation transitions among the normal operation (i.e., train operation following the planned timetable), disturbed state (i.e., duration of the TSR), and the recovery period (i.e., from end of the TSR and back to normal operation). Thus, the integrated method enables us to better coordinate the two problems, reducing train deviations, preventing infeasibility and unnecessary deviations caused by the lack of interaction between them.

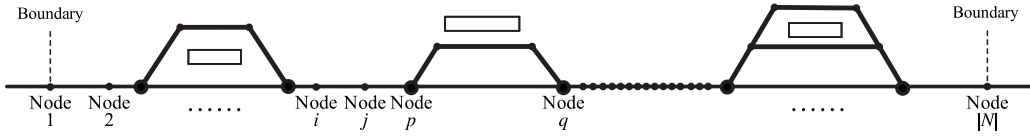


Fig. 1. A high-speed railway line.

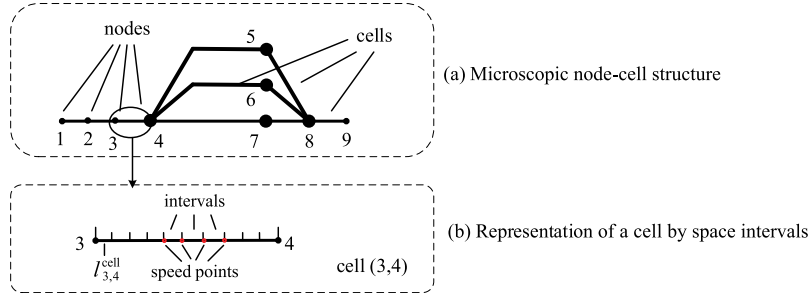


Fig. 2. Illustration of the discrete-space railway line.

- A novel mixed-integer nonlinear programming (MINLP) model is proposed, to simultaneously identify the train orders, routes, departure and arrival times, and speed profiles. The measure of train rerouting is considered, which is otherwise neglected in the available integrated approaches. In the model, the running time and the headway time are considered to be variables, dynamically depending on the real train speed and the blocking time theory. The train speed trajectories are calculated at a high spatial resolution based on the traction/braking force, the basic resistance, and the infrastructure grade and characteristics. The model considers realistic train characteristics (e.g., the train length, mass, the performance of acceleration or deceleration) and infrastructure characteristics (i.e., the grade, curve, and tunnel characteristics of the line); so the obtained train speed trajectories are feasible and close to the real-world operation.
- A piecewise linear approximation method is applied to linearize the nonlinear terms of the MINLP model, resulting in a mixed-integer linear programming (MILP) model. The method has the flexibility in selecting the number of the piecewise segments: with sufficiently many segments, a very high precision is achieved in modeling. Moreover, our method enables us to evaluate and balance the approximation errors, problem size, and computation time by setting different numbers of the piecewise segments. A two-step approach is proposed to speed up the solving process of the MILP problem.
- Comprehensive experiments are conducted, to show the impacts of the TSR-related parameters, to demonstrate the benefit of the rerouting measure, and to assess the performance of the proposed approaches. According to the experimental results, our integrated optimization method leads to an average improvement of 3–36% in solution quality (compared with the integrated approach without considering train rerouting), as well as the integrated approach outperforms the sequential approach, achieving 6%–9% improvement in solution quality.

3. Problem statement and model assumptions

3.1. Explanations of relevant concepts

We describe a high-speed rail network with microscopic details, see an illustrative example in Fig. 1. We further discretize each cell into a series of space intervals, as shown in Fig. 2(b), in order to model the train movement and calculate the train speed trajectories on a cell more precisely (Lu et al., 2016; Tan et al., 2017; Wu et al., 2018). Relevant concepts are:

- A **node** represents the position of a signal (i.e., a block signal, a home signal, or a starting signal) on high-speed railway lines.
- A **cell** describes a block section of track between two signals (i.e., two nodes).
- A **speed point** is any discrete point of a cell, where speed can change.
- An **interval** is a space between two speed points, where the trains could have only a uniform acceleration or deceleration motion.

3.2. Problem statement

Due to the differences in train characteristics, infrastructure characteristics and grade, and also the disturbances such as TSR, the running times and speeds of trains on block sections can vary a lot. Most existing microscopic models of the train rescheduling problem consider constant speed or offline-determined speed profiles, which do not include the actual TSR effects. This is a rather

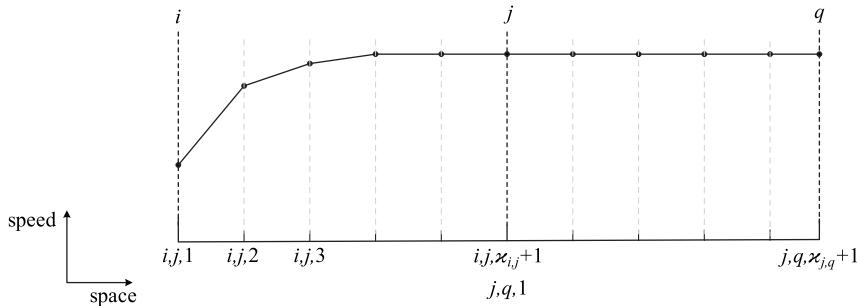


Fig. 3. Illustration of train speed trajectory on two adjacent cells.

restrictive assumption which hides the improvements of the solutions. Even worse, it may also result in infeasible solutions to the train control problem, i.e., the estimated train running times are improper, thus causing that trains need to run at a speed over the required limit (which is not allowed in reality, thus trains may be delayed again due to the improper running times). To this end, it is desired to adjust train timetables by considering train dynamics, where train running times are variables, depending on the train speeds.

Regarding the high-speed railway network, the set of nodes is denoted as N , and $|N|$ indicates the number of nodes. We consider the traffic of one direction, e.g., from node 1 to node $|N|$ in the example of Fig. 1. The infrastructure characteristics (i.e., length, grade, curve, tunnel, and speed limit) of each cell are given as input and considered piecewise constant. Moreover, we consider a set of trains scheduled on the railway line. Each train departs from an origin station, visits/passes a sequence of stations with/without a scheduled stop, and arrives at a destination, regulated by the planned timetable. The train characteristics, e.g., the length, the weight, the traction, and braking force, are all pre-determined. Specifically, we consider the maximum traction force to be a piecewise linear function of train speed. The maximum braking force is assumed to be constant (Hansen, 2008).

The headway time depends on speed, as prescribed by the Blocking Time Theory (Pachl, 2009). The blocking time is defined as a time interval in which a block section is exclusively allocated to a train and therefore blocked for other trains. Therefore, the blocking time lasts from the moment of issuing a train movement authority to the moment that it becomes possible to issue a movement authorization to another train to enter the same section. The blocking time of a block section is always longer than the time the train occupies the block section. The blocking time includes of the following time intervals: the setup time, the sight and reaction time, the approach time, the running time, clearing time, and release time. We refer to the book by Hansen (2008) for more detailed explanation of these time intervals. The approach time, running time, and the clearing time depend on train speed and length, as well as the speed limit and length of block section. Thus, we consider them as decision variables, and the others (i.e., the setup time, the sight and reaction time, and release time) are given as input parameters.

We can divide a cell into space intervals as many as wished, as shown in Fig. 2(b). The shorter the space interval is, the more precise the train speed trajectory solution becomes. However, a large number of space intervals can increase the model scale, leading to computation burden. To construct a discrete-space railway line, we first describe the railway line by a set of nodes $N = \{i, j, p, q, \dots\}$ and a set of cells $E = \{(i, j), (j, p), (p, q), \dots\}$, see Fig. 2(a) for instance. Then, each cell (i, j) is divided into $x_{i,j}$ intervals with the same length of $l_{i,j}^{\text{cell}}$, e.g., $l_{3,4}^{\text{cell}}$ for cell $(3, 4)$ in Fig. 2(b). Theoretically, there are $x_{i,j} + 1$ speed points for each train on cell (i, j) , and the position of each speed point is denoted by (i, j, k) , where integer $k \in [1, x_{i,j} + 1]$. For clarity, Fig. 3 shows the speed trajectory of a train on two adjacent cells (i, j) and (j, q) . As illustrated, the last point of cell (i, j) is in fact identical to the first point of cell (j, q) , where the speed of a traversing train f must be consistent, i.e., $v_{f,i,j,x_{i,j}+1} = v_{f,j,q,1}$. Specifically, in the model formulation, we employ a block section (i.e., cell) discretization for the train rescheduling problem, and we adopt a space discretization for the train control problem.

During exceptional events and disruptions, TSR might be issued, which constrain railway operations to ensure safety. In such cases, the operation control office would provide detailed information on the TSR, indicating the start and end times, the block sections affected, and the required limits to train speeds. The trains that traverse the affected block sections during the specific time period of the TSR must respect the maximum allowed speed limited by the TSR. Specifically, we here consider the TSR caused by for instance bad weather, maintenance works conducted on the tracks in the opposite direction, and malfunction in signaling systems, rather than by the technical failure in trains. For this reason, the TSR requires only a reduction on the maximum allowed train speeds, rather than on the train acceleration and deceleration. However, the change of the maximum allowed train speed may consequently influence the train movements in terms of acceleration and deceleration. A time-limit TSR perturbation results in an inter-dependency between the speed limits and the actual train entrance/exit times at the TSR area. To determine whether a train out of a set of running trains will suffer a speed limit, one needs to know the exact entrance and exit times in the disrupted area; those depend on the train orders by means of the headway; headway directly depends on the speed, while orders indirectly depend on speed as it is better to schedule fast trains before slow trains, to avoid delay propagation. Thus ultimately everything depends on whether the speed limit applies or not. As a result of this circular set of constraints, we do not know in advance which train should follow the TSR limit, which causes the main difficulty for the problems of train rescheduling and train control. We tackle the integrated problem of the train rescheduling and the train control for the case of TSR. The problem is to optimize for each train

the route, train orders, the arrival and departure times at/from each cell, and the speed at each speed point, with respect to the TSR, aiming at reducing the negative consequences caused by the TSR (i.e., train deviations). As the major outputs, a rescheduled timetable of all trains and a speed profile for each train along its route can be delivered.

3.3. Model assumptions

In our model, we make the following assumptions:

- (1). Each train is assumed to follow a uniform acceleration or deceleration motion in each space interval with a constant acceleration or deceleration rate. Thus, the train speeds in each space interval increase or decrease linearly. Also, the acceleration and deceleration of a train can vary in different space intervals.
- (2). We use the average value of the speed limit, track grade, curve, and tunnel on a cell to represent the infrastructure details for each cell. Thus, the speed limit, track grade, curve, and tunnel are assumed to be the same on each cell, but differs among cells.
- (3). The granularity of time is one second.

4. Mathematical formulations

In this section, we formulate the optimization model to integrate the train rescheduling problem and the train control problem for high-speed railway operations with Temporary Speed Restriction (TSR) on trains. We introduce the notations in Section 4.1, followed by the mathematical formulations of the integrated optimization model in Section 4.2, which is a mixed-integer nonlinear programming (MINLP) model. Section 5.1 further reformulates the nonlinear terms of the MINLP model into a linear form by mean of an approximation method.

4.1. Notations

Table 2 gives the general subscripts, sets, input parameters, and decision variables.

4.2. Integrated optimization method

When the train operation is disturbed by some issue, the aim of the dispatcher is typically to let the train operation recover to the planned train timetable. Thus, we consider the objective criterion to be the total absolute deviation time of all trains at all visited stations (Luan et al., 2018a). The objective function can be written as follows:

$$Z_{\text{deviation}} = \min \sum_{f \in F} \left| \sum_{(i,j):(i,j) \in E_f^{\text{Stop}}} (a_{f,i,j} - \bar{a}_{f,i,j}) \right|. \quad (1)$$

• Train motion constraints

We here propose some constraints to enforce modeling of the train motion, i.e., describing the train motion on discrete-space intervals, restricting the traction and braking force, calculating the basic and line resistances, as well as representing the motion with a varying speed, but uniform acceleration or deceleration on each space interval.

The mass-point method is often used in the studies of the train control problem (Howlett and Pudney, 1995). Based on the discrete-space method (see the description in Section 3.2), the motion of a train can be described as the followings:

$$\rho_f m_f \varpi_{f,i,j,k} = u_{f,i,j,k} - R_{f,i,j,k}^{\text{basic}} - R_{f,i,j,k}^{\text{line}}, \quad \forall f \in F, (i,j) \in E_f, k \in [1, x_{i,j}]. \quad (2)$$

Eq. (2) shows that, at the position (i,j,k) , the acceleration/deceleration rate $\varpi_{f,i,j,k}$ of train f is related to the traction or braking force $u_{f,i,j,k}$, the basic resistance $R_{f,i,j,k}^{\text{basic}}$ (including roll resistance and air resistance), and the line resistance $R_{f,i,j,k}^{\text{line}}$ caused by track grade, curves, and tunnels, as well as the mass of train f .

The following constraint

$$u_{f,i,j,k}^{\text{brk}} \leq u_{f,i,j,k} \leq u_{f,i,j,k}^{\text{trc}}, \quad \forall f \in F, (i,j) \in E_f, k \in [1, x_{i,j}] \quad (3a)$$

ensures that the traction or braking force $u_{f,i,j,k}$ is bounded by the maximum traction force $u_{f,i,j,k}^{\text{trc}}$ and the maximum braking force $u_{f,i,j,k}^{\text{brk}}$. The maximum braking force $u_{f,i,j,k}^{\text{brk}}$ is considered to be constant (Hansen, 2008). The maximum traction force is typically a function of the speed, which is shown in Fig. 4. This diagram is described as a group of hyperbolic or parabolic formulas, where each formula approximates the actual traction force for a certain speed interval, e.g., Hansen (2008) and Wang (2014). For example, if the train speed belongs to interval $[\bar{v}_{f,r}, \bar{v}_{f,r+1}]$, then the maximum traction force can be written as

$$u_{f,i,j,k}^{\text{trc}} = c_{1,f,r} + c_{2,f,r} \cdot v_{f,i,j,k} + c_{3,f,r} \cdot v_{f,i,j,k}^2, \quad \forall f \in F, r \in R, (i,j) \in E_f, k \in [1, x_{i,j}]. \quad (3b)$$

The movement of high-speed trains (we consider electrical multiple units, or EMUs, in the following) is influenced by many factors, such as the train characteristics, the technical status, the infrastructure grade, the train speed, etc. These factors are

Table 2

General subscripts, sets, input parameters, and decision variables.

Symbol	Description
Subscripts and sets	
N	Set of nodes
E	Set of cells (i.e., block sections), E^{sta} denotes the set of cells in station areas, and E^{opt} indicates the set of cells in open track areas.
F	Set of trains
R	Set of speed intervals for calculating the maximum traction force
E^{TSR}	Set of cells affected by TSR, $E^{\text{TSR}} \subseteq E$
i, j, p, q	Node index, $i, j, p, q \in N$
r	Index of speed intervals for calculating the maximum traction force, $r \in R$
f	Train index, $f \in F$
E_f	Set of cells that train f may use, $E_f \subseteq E$
E_f^{Stop}	Set of cells in which train f could stop, $E_f^{\text{Stop}} \subseteq E^{\text{sta}}$
Input parameters	
$\bar{a}_{f,i,j}$	Planned arrival time of train f of cell (i, j)
$L_{i,j}^{\text{cell}}$	Length of cell (i, j)
$l_{i,j}^{\text{cell}}$	Length of each space interval of cell (i, j) , i.e., $l_{i,j}^{\text{cell}} = \frac{L_{i,j}^{\text{cell}}}{x_{i,j}}$, $x_{i,j}$ is the number of the space intervals of cell (i, j)
$v_{i,j}^{\text{max}}$	The maximum allowed train speed on cell (i, j) under normal operations (unaffected by TSR)
g	The gravitational acceleration
$\varphi_{i,j}^{\text{grade}}, \varphi_{i,j}^{\text{cur}}, \varphi_{i,j}^{\text{tun}}$	Average track grade, curve radius, and tunnel length on cell (i, j)
$\Delta \varpi^{\text{max}}$	Threshold of the change of the acceleration/deceleration rate between two adjacent space intervals, for ensuring passenger-riding comfort
$t_{i,j}^{\text{sta}}, t_{i,j}^{\text{end}}$	Start time and end time of TSR on cell $(i, j) \in E^{\text{TSR}}$
$v_{i,j}^{\text{TSR}}$	The maximum allowed train speed limited by TSR for cell $(i, j) \in E^{\text{TSR}}$
$u_{f,i,j}^{\text{min}}, u_{f,i,j}^{\text{max}}$	The minimum and maximum dwell time of train f on cell (i, j)
d_f^{car}	The earliest departure time of train f at its origin node
o_f, s_f	Origin node and destination node of train f
v_{o_f}	Initial speed of train f at its origin node
ρ_f	Rotating mass of train f
m_f, L_f^{tra}	Mass and length of train f
$u_{f,i,j,k}^{\text{brk}}$	The maximum braking force of train f at the k th speed point on cell (i, j)
$c_{1,f,r}, c_{2,f,r}, c_{3,f,r}$	Coefficients of the maximum traction force for train f within the given train speed interval of $[\bar{v}_{f,r}, \bar{v}_{f,r+1}]$, where $r = 1$ or 2
A_f, B_f, C_f	Coefficients of the basic resistance for train f
$\tau_{f,i,j}^{\text{set}}$	Setup time of train f on cell (i, j)
$\tau_{f,i,j}^{\text{sr}}$	Sight and reaction time of train f on cell (i, j)
$\tau_{f,i,j}^{\text{release}}$	Release time for releasing cell (i, j) after the clearance of train f
$\phi_{f,i,j,p,q}$	Train approach indicator, representing whether the running time of train f on cell (p, q) should be included as one component of the approach time for train f on cell (i, j)
τ^{appsta}	Approach time for trains running in station areas (i.e., on siding tracks).
M, ε	A sufficiently large or small positive number
Decision variables	
$a_{f,i,j}, d_{f,i,j}$	Arrival time, departure time of train f at cell (i, j)
$w_{f,i,j}$	dwell time of train f on cell (i, j)
$\theta_{f,f',i,j}$	Binary train order variable, $\theta_{f,f',i,j} = 1$ if train f' arrives at cell (i, j) after train f , and otherwise $\theta_{f,f',i,j} = 0$
$x_{f,i,j}$	Binary train route variable, $x_{f,i,j} = 1$ if train f runs on cell (i, j) , and otherwise $x_{f,i,j} = 0$
$v_{f,i,j,k}$	Speed of train f at the k th discretized speed point of cell (i, j) , where $k \in [1, x_{i,j} + 1]$
$\varpi_{f,i,j,k}$	Acceleration or deceleration rate of train f at the k th discretized speed point of cell (i, j) , where $k \in [1, x_{i,j} + 1]$
$\xi_{f,i,j}$	Binary variable, $\xi_{f,i,j} = 1$ if train f is affected by TSR on cell (i, j) , and otherwise, $\xi_{f,i,j} = 0$
$\tau_{f,i,j}^{\text{approach}}$	Approach time of train f on cell (i, j)
$\tau_{f,i,j}^{\text{clear}}, \tau_{f,i,j}^{\text{run}}$	Clearing time, running time of train f on cell (i, j)
$\psi_{f,i,j}$	Safety time interval between occupancy of cell (i, j) and the arrival of train f

(continued on next page)

Table 2 (continued).

$h_{f,j}$	Safety time interval between the departure of train f and the release of cell (i, j)
$u_{f,i,j,k}^{\text{trc}}, u_{f,i,j,k}, R_{f,i,j,k}^{\text{basic}}, R_{f,i,j,k}^{\text{line}}$	The maximum traction, traction or braking force, basic resistance, and line resistance of train f at the k th discretized speed point of cell (i, j) , where $k \in [1, \kappa_{i,j} + 1]$
$\zeta_{f,i,j}$	Logical variable, $\zeta_{f,i,j} = 1$ if train f stops on cell (i, j) , and otherwise, $\zeta_{f,i,j} = 0$
$\mu_{f,i,j,k}$	Time duration that train f stays in the k th space interval of cell (i, j) , i.e., the duration between the arrival and the departure of train f at/from the k th space interval of cell (i, j)
$\alpha_{f,i,j}, \beta_{f,i,j}$	Occupancy time, release time of train f on cell (i, j)

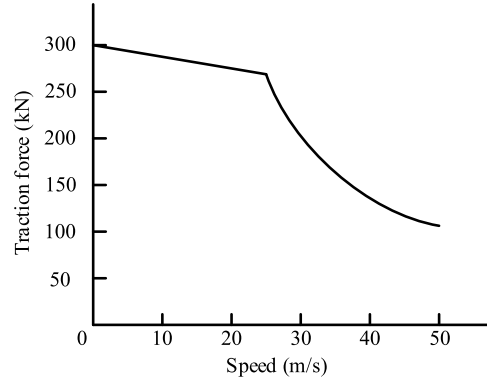


Fig. 4. Illustration of maximum traction force as a function of train speed (Hansen, 2008).

rather complicated and difficult to model with theoretical formulas. In practice, an empirical formula (Davis equation) given by a large number of train experiments is typically used for calculating the basic resistance. We here apply the Davis formula

$$R_{f,i,j,k}^{\text{basic}} = m_f \cdot g \cdot (A_f + B_f \cdot v_{f,i,j,k} + C_f \cdot v_{f,i,j,k}^2), \quad \forall f \in F, (i, j) \in E_f, k \in [1, \kappa_{i,j}], \quad (4)$$

to calculate the basic resistance of train f at position (i, j, k) , which is a quadratic function of the real train speed. The line resistance caused by track grade, curves, and tunnels can be formulated by

$$R_{f,i,j,k}^{\text{line}} = \frac{m_f \cdot g \cdot \varphi_{i,j}^{\text{grade}}}{1000} + \varphi_{i,j}^{\text{cur}} + R_{f,i,j,k}^{\text{tun}}, \quad \forall f, (i, j) \in E_f, k \in [1, \kappa_{i,j}]. \quad (5a)$$

In (5a), the first two terms (i.e., the resistances caused by track grade and curves) do not depend on the train speed. When the train is running in tunnels, the train faces a different air resistance that depends on the tunnel form, the smoothness on the tunnel wall, the exterior surface of the train, and so on Wang (2014). Typically, an empirical formula can be used to calculate the tunnel resistance of a EMU (Zhu et al., 2015), which is written as:

$$R_{f,i,j,k}^{\text{tun}} = \frac{m_f \cdot g \cdot \varphi_{i,j}^{\text{tun}} \cdot v_{f,i,j,k}^2}{10^7}, \quad \forall f, (i, j) \in E_f, k \in [1, \kappa_{i,j}]. \quad (5b)$$

Recall that there are $\kappa_{i,j} + 1$ speed values $v_{f,i,j,k}$ for train f on cell (i, j) , where integer $k \in [1, \kappa_{i,j} + 1]$. We assume that each train follows a uniform acceleration or deceleration motion in each space interval. The acceleration or deceleration rate $\varpi_{f,i,j,k}$ of train f at the k th speed point of cell (i, j) can be calculated by

$$\varpi_{f,i,j,k} = \frac{v_{f,i,j,k+1}^2 - v_{f,i,j,k}^2}{2 \cdot l_{i,j}^{\text{cell}}}, \quad \forall f, (i, j) \in E_f, k \in [1, \kappa_{i,j}]. \quad (6a)$$

The time duration that train f stays in the k th space interval of cell (i, j) can be calculated by

$$\mu_{f,i,j,k} = \frac{l_{i,j}^{\text{cell}}}{(v_{f,i,j,k} + v_{f,i,j,k+1})/2}, \quad \forall f, (i, j) \in E_f, k \in [1, \kappa_{i,j}]. \quad (6b)$$

In addition, we can enforce the riding comfort by the following constraint

$$|\varpi_{f,i,j,k+1} - \varpi_{f,i,j,k}| \leq \Delta \varpi^{\max}, \quad \forall f, (i, j) \in E_f, k \in [1, \kappa_{i,j}]. \quad (6c)$$

Constraint (6c) represents that the change of the acceleration/deceleration rate variable $\varpi_{f,i,j,k}$ between two adjacent space intervals cannot exceed the given threshold $\Delta \varpi^{\max}$, since passengers may feel sick if the train speed changes heavily in a short time, i.e., heavy acceleration or brake of a sudden.

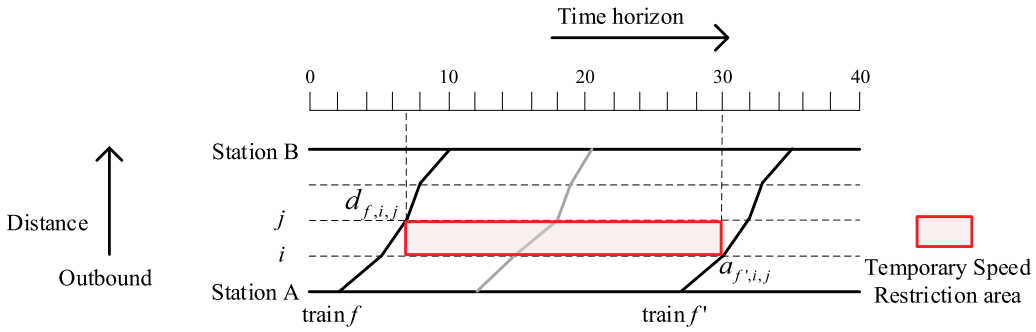


Fig. 5. Illustration of the train traversing the TSR area.

Constraints (3b), (4), (5b), and (6a) are nonlinear, since they all include a quadratic term, i.e., the square of train speed v^2 . Section 5.1 introduces a reformulation method to transform the nonlinear term v^2 into a linear form by means of piecewise linear approximation. Constraint (6b) is also a nonlinear constraint; we describe its reformulation in Section 5.1.

• Train speed constraints

We here formulate the train speed constraints, including the setup of the initial train speeds at the origins, the link of train speeds with train dwell activities, the cell-to-cell transition of train speed, and the train speed limits under a normal situation. The following constraint

$$v_{f,o_f,j,1} = v_{o_f}, \quad \forall f \in F, (o_f, j) \in E_f, \quad (7a)$$

ensures that train f departs from its origin node o_f at the given initial speed v_{o_f} . If train f is scheduled to stop at its origin, then the corresponding initial speed v_{o_f} is set to be zero; otherwise, we set v_{o_f} to be non-zero, i.e., the actual speed when train f enters the railway line under consideration. Moreover, the train speed cannot exceed the given speed limit on each cell, which is forced by the followings:

$$0 \leq v_{f,i,j,k} \leq v_{i,j}^{\max} \cdot x_{f,i,j}, \quad \forall f \in F, (i, j) \in E_f, k \in [1, x_{i,j}]. \quad (7b)$$

The cell-to-cell speed transition constraint

$$\sum_{i:(i,j) \in E_f} v_{f,i,j,x_{i,j}+1} = \sum_{q:(j,q) \in E_f} v_{f,j,q,1}, \quad \forall f \in F, j \in N \setminus \{o_f, s_f\}. \quad (8)$$

enforces the consistency of the train speed between two adjacent cells, i.e., the speed of train f at the last speed point of cell (i, j) equals to its speed at the first speed point of cell (j, q) .

A set of constraints is considered to link the train dwell and the train speed, in which

$$(\zeta_{f,i,j} - 1) \cdot w_{f,i,j}^{\max} \leq w_{f,i,j} - \varepsilon, \quad \forall f \in F, (i, j) \in E_f, \quad (9a)$$

$$\zeta_{f,i,j} \cdot w_{f,i,j}^{\max} \geq w_{f,i,j}, \quad \forall f \in F, (i, j) \in E_f, \quad (9b)$$

use a logical variable $\zeta_{f,i,j}$ to satisfy the condition $[w_{f,i,j} > 0 \Leftrightarrow \zeta_{f,i,j} = 1]$. The constraint

$$(\zeta_{f,i,j} - 1) \cdot v_{i,j}^{\max} + v_{f,i,j,x_{i,j}+1} \leq 0, \quad \forall f \in F, (i, j) \in E_f, \quad (9c)$$

further ensures that the speed of train f at the last speed point $x_{i,j} + 1$ must be zero if train f stops on cell (i, j) , i.e., $\zeta_{f,i,j} = 1$. In inequality (9c), the train could stop at the planned stations or stop due to train conflicts. Specifically, if $\zeta_{f,i,j} = 1$, then the inequality (9c) turns into $v_{f,i,j,x_{i,j}+1} \leq 0$; as the train speed is non-negative, the speed of train f at the last speed point on cell (i, j) must equal zero (i.e., $v_{f,i,j,x_{i,j}+1} = 0$). If $\zeta_{f,i,j} = 0$, the inequality (9c) turns into $v_{f,i,j,x_{i,j}+1} \leq v_{i,j}^{\max}$, which has equal effect to constraint (7b).

• Temporary speed restriction (TSR) constraints

A set of constraints is used to formulate the TSR. As shown in Fig. 5, train f will not be affected by Temporary Speed Restriction (TSR) on cell (i, j) , i.e., $\xi_{f,i,j} = 0$, if train f has left the cell before the start of the TSR (i.e., $d_{f,i,j} \leq t_{i,j}^{\text{sta}}$) or only enter the cell after the end of the TSR (i.e., $a_{f,i,j} \geq t_{i,j}^{\text{end}}$). Otherwise, the train will be affected by TSR ($\xi_{f,i,j} = 1$, if $d_{f,i,j} > t_{i,j}^{\text{sta}} \cap a_{f,i,j} < t_{i,j}^{\text{end}}$).

The TSR of the trains could be modeled by using the following steps. First, let us consider two logical variables $\lambda_{1,f,i,j}$ and $\lambda_{2,f,i,j}$ to satisfy the conditions: $[d_{f,i,j} > t_{i,j}^{\text{sta}} \Leftrightarrow \lambda_{1,f,i,j} = 1]$ and $[a_{f,i,j} < t_{i,j}^{\text{end}} \Leftrightarrow \lambda_{2,f,i,j} = 1]$. We can write the former condition as the following set of linear constraints

$$(\lambda_{1,f,i,j} - 1) \cdot M \leq d_{f,i,j} - t_{i,j}^{\text{sta}} - \varepsilon, \quad \forall f \in F, (i, j) \in E_f, \quad (10a)$$

$$\lambda_{1,f,i,j} \cdot M \geq d_{f,i,j} - t_{i,j}^{\text{sta}}, \quad \forall f \in F, (i, j) \in E_f. \quad (10b)$$

Similarly, the latter condition $\left[a_{f,i,j} < t_{i,j}^{\text{end}} \Leftrightarrow \lambda_{2,f,i,j} = 1 \right]$ can be represented as

$$(\lambda_{2,f,i,j} - 1) \cdot M \leq t_{i,j}^{\text{end}} - a_{f,i,j} - \varepsilon, \quad \forall f \in F, (i,j) \in E_f, \quad (10c)$$

$$\lambda_{2,f,i,j} \cdot M \geq t_{i,j}^{\text{end}} - a_{f,i,j}, \quad \forall f \in F, (i,j) \in E_f. \quad (10d)$$

We further consider another logical variable $\xi_{f,i,j}$ to satisfy the following condition: $[\lambda_{1,f,i,j} = 1 \cap \lambda_{2,f,i,j} = 1 \Leftrightarrow \xi_{f,i,j} = 1]$, which could be equivalently transformed into the following set of constraints (Luan et al., 2018a; Wang, 2014):

$$-\lambda_{1,f,i,j} + \xi_{f,i,j} \leq 0, \quad \forall f \in F, (i,j) \in E_f, \quad (10e)$$

$$-\lambda_{2,f,i,j} + \xi_{f,i,j} \leq 0, \quad \forall f \in F, (i,j) \in E_f, \quad (10f)$$

$$\lambda_{1,f,i,j} + \lambda_{2,f,i,j} - \xi_{f,i,j} \leq 1, \quad \forall f \in F, (i,j) \in E_f. \quad (10g)$$

As such, we can use the logical variable $\xi_{f,i,j}$ to indicate whether train f is affected by the TSR on cell (i,j) . If the train is affected, i.e., $\xi_{f,i,j} = 1$, then the speed of train f on cell (i,j) cannot exceed the maximum allowed train speed $v_{i,j}^{\text{TSR}}$ limited by the TSR, which can be represented by

$$(1 - \xi_{f,i,j}) \cdot v_{i,j}^{\text{max}} \geq v_{f,i,j,k} - v_{i,j}^{\text{TSR}}, \quad \forall f \in F, (i,j) \in E_f \cap E^{\text{TSR}}, k \in [1, x_{i,j}]. \quad (10h)$$

• Train transition time constraints

We here identify the train arrival, departure, and dwell times by presenting a set of constraints, which ensures the feasible time transitions of trains within each cell and between each pair of adjacent cells. The train transition within a cell can be written as follows:

$$d_{f,i,j} - w_{f,i,j} - a_{f,i,j} = \tau_{f,i,j}^{\text{run}}, \quad \forall f \in F, (i,j) \in E_f, \quad (11a)$$

i.e., the running time of train f on cell (i,j) equals the departure time of train f from cell (i,j) minus its corresponding dwell time and arrival time. The running time of train f on cell (i,j) is calculated by

$$\tau_{f,i,j}^{\text{run}} = \sum_{k=1}^{x_{i,j}} \mu_{f,i,j,k}, \quad \forall f \in F, (i,j) \in E_f, \quad (11b)$$

i.e., the sum of the running times of train f over all the space intervals of cell (i,j) .

The cell-to-cell transition constraint

$$\sum_{i:(i,j) \in E_f} d_{f,i,j} = \sum_{q:(j,q) \in E_f} a_{f,j,q}, \quad \forall f \in F, j \in N, \quad (11c)$$

enforces the transition time between two adjacent cells.

A train cannot leave its origin node before the earliest departure time (i.e., the planned departure time at its origin), ensured by

$$\sum_{j:(o_f,j) \in E_f} a_{f,o_f,j} \geq d_f^{\text{ear}}, \quad \forall f \in F. \quad (11d)$$

The train dwell time constraint

$$w_{f,i,j}^{\text{min}} \cdot x_{f,i,j} \leq w_{f,i,j} \leq w_{f,i,j}^{\text{max}} \cdot x_{f,i,j}, \quad \forall f \in F, (i,j) \in E_f, \quad (11e)$$

guarantees the required minimum and maximum dwell times at stations. The maximum dwell time typically includes a sufficient large value to account for conflicts and readjustments of train traffic.

• Safety headway constraints

As mentioned in Section 3.2, we use blocking time theory to ensure the safety headway time among trains. Next, we illustrate the calculation of the approach time and the clearing time.

The approach time indicates the time duration for train running from the signal that provides the approach indication to the signal at the entrance of the block section (Pachl, 2009).

In this paper, we separate the calculation of the approach time in open track and station area to simplify the model. The approach time is calculated by

$$\tau_{f,i,j}^{\text{approach}} = a_{f,i,j} - \sum_{(p,q) \in E_f} a_{f,p,q} \cdot \phi_{f,i,j,p,q}, \quad \forall f \in F, (i,j) \in E_f \cap E^{\text{opt}}, \quad (12a)$$

for open track area, and by

$$\tau_{f,i,j}^{\text{approach}} = \tau_{f,i,j}^{\text{appsta}} \cdot x_{f,i,j}, \quad \forall f \in F, (i,j) \in E_f \cap E^{\text{sta}}, \quad (12b)$$

for station area.

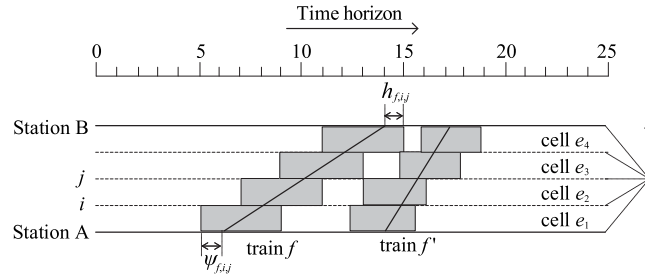


Fig. 6. Illustration of blocking times of two trains.

The clearing time is defined as the time duration to clear the block section and the overlap with the full length of the train. We formulate the clearing time of train f on cell (i, j) as

$$\tau_{f,i,j}^{\text{clear}} = \sum_{q:(j,q) \in E_f} \frac{L_f}{L_{\text{cell}}} \cdot \tau_{f,j,q}^{\text{run}}, \quad \forall f \in F, (i, j) \in E_f, \quad (13)$$

according to the train length, the running time, and the dwell time on the successive cell (j, q) .

Fig. 6 depicts the blocking times of two trains on a sequence of cells. As illustrated in Fig. 6, a cell is pre-blocked for a train before its arrival and post-released after its departure. The safety time interval $\psi_{f,i,j}$ between cell occupancy and train arrival is defined as

$$\psi_{f,i,j} = \tau_{f,i,j}^{\text{set}} + \tau_{f,i,j}^{\text{sr}} + \tau_{f,i,j}^{\text{approach}}, \quad \forall f \in F, (i, j) \in E_f, \quad (14)$$

including the setup time, the sight and reaction time, and the approach time. It is used for pre-blocking a block section before the arrival of a train. The safety time interval $h_{f,i,j}$ between train departure and cell release is used for post-releasing a block section after the departure of a train, which can be written as follows:

$$h_{f,i,j} = \tau_{f,i,j}^{\text{release}} + \tau_{f,i,j}^{\text{clear}}, \quad \forall f \in F, (i, j) \in E_f, \quad (15)$$

including the clearing time and the release time.

Based on the above safety headway times, the blocking time of train f traversing cell (i, j) , i.e., the cell occupancy and cell release time can be respectively written as

$$\alpha_{f,i,j} = a_{f,i,j} - \psi_{f,i,j}, \quad \forall f \in F, (i, j) \in E_f, \quad (16a)$$

$$\beta_{f,i,j} = d_{f,i,j} + h_{f,i,j}, \quad \forall f \in F, (i, j) \in E_f. \quad (16b)$$

• Train rerouting and reordering constraints

Here we formulate the constraints for traffic regulations, which are often included in the train rescheduling models. These include balancing the traffic flow on the rail network, ensuring one and only one route applied to each train, describing the orders of trains, and enforcing the cell capacity.

A set of flow balance constraints

$$\sum_{j:(o_f,j) \in E_f} x_{f,o_f,j} = 1, \quad \forall f \in F, \quad (17a)$$

$$\sum_{i:(i,j) \in E_f} x_{f,i,j} = \sum_{q:(j,q) \in E_f} x_{f,j,q}, \quad \forall f \in F, j \in N \setminus \{o_f, s_f\}, \quad (17b)$$

$$\sum_{i:(i,s_f) \in E_f} x_{f,i,s_f} = 1, \quad \forall f \in F, \quad (17c)$$

is used to balance the traffic flow from the origin node, via intermediate nodes, and to destination node, respectively.

The following constraints

$$x_{f,i,j} - 1 \leq a_{f,i,j} \leq x_{f,i,j} \cdot M, \quad \forall f \in F, (i, j) \in E_f, \quad (18a)$$

$$x_{f,i,j} - 1 \leq d_{f,i,j} \leq x_{f,i,j} \cdot M, \quad \forall f \in F, (i, j) \in E_f, \quad (18b)$$

link the train arrival and departure variables $a_{f,i,j}$ and $d_{f,i,j}$ with the train route variable $x_{f,i,j}$, so as to describe whether cell (i, j) is selected by train f for traversing the network from its origin to the destination.

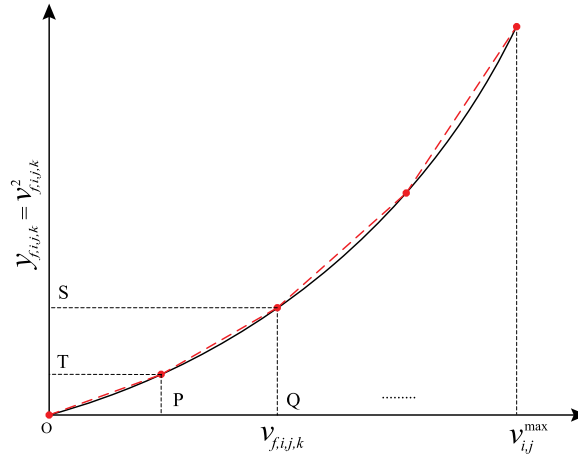


Fig. 7. Illustration of piecewise linear approximation.

The train order constraint

$$x_{f,i,j} + x_{f',i,j} - 1 \leq \theta_{f,f',i,j} + \theta_{f',f,i,j} \leq 3 - x_{f,i,j} - x_{f',i,j}, \quad \forall f, f' \in F, f \neq f', (i, j) \in E_f \cap E_{f'} \quad (19a)$$

will reduce to $\theta_{f,f',i,j} + \theta_{f',f,i,j} = 1$, if and only if both train f and train f' use the cell (i, j) , i.e., $x_{f,i,j} = x_{f',i,j} = 1$. The equality $\theta_{f,f',i,j} + \theta_{f',f,i,j} = 1$ indicates that either train f arrives at cell (i, j) after train f' or vice versa. The constraints

$$\theta_{f,f',i,j} \leq x_{f,i,j}, \quad \forall f, f' \in F, f \neq f', (i, j) \in E_f \cap E_{f'}, \quad (19b)$$

$$\theta_{f',f,i,j} \leq x_{f',i,j}, \quad \forall f, f' \in F, f \neq f', (i, j) \in E_f \cap E_{f'}, \quad (19c)$$

link the train order variable and the train route variable. Specifically, if either train f or train f' does not use cell (i, j) , i.e., $x_{f,i,j} = 0$ or $x_{f',i,j} = 0$, then we have $\theta_{f,f',i,j} = 0$ and $\theta_{f',f,i,j} = 0$.

The cell capacity constraints

$$\alpha_{f,i,j} + (3 - x_{f,i,j} - x_{f',i,j} - \theta_{f,f',i,j}) \cdot M \geq \beta_{f,i,j}, \quad \forall f, f' \in F, f \neq f', (i, j) \in E_f \cap E_{f'}, \quad (20a)$$

$$\alpha_{f',j,i} + (3 - x_{f,i,j} - x_{f',j,i} - \theta_{f,f',i,j}) \cdot M \geq \beta_{f',j,i}, \quad \forall f, f' \in F, f \neq f', (i, j) \in E_f, (j, i) \in E_{f'}, \quad (20b)$$

ensure that a cell is accessible by a train if and only if the cell has been released from the use of the preceding train. In other words, any pair of trains can only access one cell sequentially, rather than at the same time.

The integrated optimization model including the objective function (1) and constraints (2)–(20) is a non-linear model, due to non-linear constraints (3b), (4), (5b), (6a), and (6b). We next introduce the reformulation of these nonlinear constraints in Section 5.1.

5. Solution approaches

5.1. Reformulation approach of the nonlinear terms

In this Section, we apply piecewise linear approximation to reformulate the nonlinear constraints (3b), (4), (5b), (6a), and (6b), and we use some extra logical variables to reformulate the if-then constraints (which appears in the approximation process).

In constraints (3b), (4), (5b), and (6a), the only non-linear term is $v_{f,i,j,k}^2$. A piecewise linear approximation to this function is illustrated in Fig. 7. We use δ -form method to linearize the quadratic term since it is usually considered to be computationally more efficient (Williams, 2013). The train speed cannot exceed $v_{i,j}^{\max}$ according to constraint (7b). The piecewise linear approximation of the quadratic term $v_{f,i,j,k}^2$ are therefore only considered for $v_{f,i,j,k} \in [0, v_{i,j}^{\max}]$. To eliminate the non-linear term, we denote $y_{f,i,j,k}$ to be the square of $v_{f,i,j,k}$, i.e., $y_{f,i,j,k} = v_{f,i,j,k}^2$. Let us assume the curve of $y_{f,i,j,k} = v_{f,i,j,k}^2$ has been equally divided into σ straight line portions. Variables $\delta_{1,f,i,j,k}$, $\delta_{2,f,i,j,k}$, ..., and $\delta_{\sigma,f,i,j,k}$ are introduced to represent single (linear) proportions of the intervals on the X-axis (e.g., OP and PQ), which are used to make up the value of $v_{f,i,j,k}$. We then get

$$v_{f,i,j,k} = \frac{v_{i,j}^{\max}}{\sigma} \cdot \sum_{i=1}^{\sigma} \delta_{i,f,i,j,k}, \quad (21a)$$

where

$$0 \leq \delta_{i,f,i,j,k} \leq 1, \quad \forall i \in [1, \sigma]. \quad (21b)$$

Since the X -axis has been divided into σ portions, the length of each interval (e.g., OP or PQ) is $\frac{v_{i,j}^{\max}}{\sigma}$. The coefficients of $\delta_{1,f,i,j,k}$, $\delta_{2,f,i,j,k}$, ..., and $\delta_{\sigma,f,i,j,k}$ are $\frac{v_{i,j}^{\max}}{\sigma}$.

Similarly,

$$y_{f,i,j,k} = \sum_{i=1}^{\sigma} \left[\left(\frac{v_{i,j}^{\max}}{\sigma} \cdot i \right)^2 - \left(\frac{v_{i,j}^{\max}}{\sigma} \cdot (i-1) \right)^2 \right] \cdot \delta_{i,f,i,j,k}, \quad (21c)$$

where the coefficients of $\delta_{1,f,i,j,k}$, $\delta_{2,f,i,j,k}$, ..., and $\delta_{\sigma,f,i,j,k}$ are now the lengths of the intervals on the Y -axis (e.g., OT and TS).

To ensure that $v_{f,i,j,k}$ and $y_{f,i,j,k}$ are the coordinates of points on the piecewise lines, we must make the following stipulation (Williams, 2013): if any $\delta_{i',f,i,j,k}$ is non-zero, all the preceding $\delta_{i',f,i,j,k}$ (where $i' < i$) must take the value of 1, and all the succeeding $\delta_{i'',f,i,j,k}$ (where $i'' > i$) must take the value of 0.

The above stipulation clearly ensures that $v_{f,i,j,k}$ and $y_{f,i,j,k}$ truly represent distances along their respective axes. The stipulation can be written as following if-then constraints:

$$\sum_{\chi=1}^{i-1} \delta_{\chi,f,i,j,k} = i-1, \text{ if } \delta_{i,f,i,j,k} > 0, \quad \forall i \in [1, \sigma], f \in F, (i, j) \in E_f, k \in [1, \kappa_{i,j}] \quad (22a)$$

$$\sum_{\chi=i}^{\sigma} \delta_{\chi,f,i,j,k} = 0, \text{ if } \delta_{i,f,i,j,k} = 0, \quad \forall i \in [1, \sigma], f \in F, (i, j) \in E_f, k \in [1, \kappa_{i,j}] \quad (22b)$$

Moreover, we introduce an auxiliary logical variable $\eta_{i,f,i,j,k}$ to satisfy the condition of $[\delta_{i,f,i,j,k} > 0 \Leftrightarrow \eta_{i,f,i,j,k} = 1]$, which can be further written as the following linear constraints:

$$\varepsilon \cdot \eta_{i,f,i,j,k} \leq \delta_{i,f,i,j,k} \leq \eta_{i,f,i,j,k}, \quad \forall i \in [1, \sigma], f \in F, (i, j) \in E_f, k \in [1, \kappa_{i,j}] \quad (22c)$$

$$\delta_{i,f,i,j,k} - \varepsilon \cdot \eta_{i,f,i,j,k} \geq 0, \quad \forall i \in [1, \sigma], f \in F, (i, j) \in E_f, k \in [1, \kappa_{i,j}] \quad (22d)$$

We can then rewrite the if-then constraints (22a) and (22b) into the following linear form:

$$\sum_{\chi=1}^{i-1} \delta_{\chi,f,i,j,k} \geq (i-1) \cdot \eta_{i,f,i,j,k}, \quad \forall i \in [1, \sigma], f \in F, (i, j) \in E_f, k \in [1, \kappa_{i,j}] \quad (23a)$$

$$\sum_{\chi=i}^{\sigma} \delta_{\chi,f,i,j,k} \leq M \cdot \eta_{i,f,i,j,k}, \quad \forall i \in [1, \sigma], f \in F, (i, j) \in E_f, k \in [1, \kappa_{i,j}] \quad (23b)$$

As a result, the nonlinear speed-related term $v_{f,i,j,k}^2$ in constraints (3b), (4), (5b), and (6a) is replaced by $y_{f,i,j,k}$ with the linear constraints (21), (22c), (22d), and (23). The variables $\delta_{i,f,i,j,k}$ for all $i \in [1, \sigma]$ are given simple upper bounds of 1, thus exploiting the bounded variable version of the simplex algorithm. By replacing the original speed-related variables and description of the TSR, with the approximated ones, the proposed method greatly decreases the complexity of the model and avoids nonlinearity while keeping high modeling precision.

To approximate the nonlinear constraint (6b), we first rewrite it as

$$l_{i,j}^{\text{cell}} = \mu_{f,i,j,k} \cdot (v_{f,i,j,k} + v_{f,i,j,k+1})/2, \quad \forall f \in F, (i, j) \in E_f, k \in [1, \kappa_{i,j}], \quad (24a)$$

which contains a nonlinear term of $\mu \cdot v$. Each nonlinear term $\mu \cdot v$ can be reformulated as $\frac{(\mu+v)^2 - (\mu-v)^2}{4}$. As such, constraint (24a) can be further rewritten as

$$8 \cdot l_{i,j}^{\text{cell}} = (v_{f,i,j,k} + v_{f,i,j,k+1} + \mu_{f,i,j,k})^2 - (v_{f,i,j,k} + v_{f,i,j,k+1} - \mu_{f,i,j,k})^2. \quad (24b)$$

This constraint includes quadratic terms, which can be reformulated by following the method similar to the one of approximating $v_{f,i,j,k}^2$, as introduced above. For the sake of compactness, we do not present the details here. It should be noted that, the curves (i.e., $\theta_{1,f,i,j,k} = (v_{f,i,j,k} + v_{f,i,j,k+1} + \mu_{f,i,j,k})^2$ and $\theta_{2,f,i,j,k} = (v_{f,i,j,k} + v_{f,i,j,k+1} - \mu_{f,i,j,k})^2$) are assumed to be divided into n straight-line portions, rather than σ portions for curve $y_{f,i,j,k} = v_{f,i,j,k}^2$. Our method allows precise modeling for the problems, and the approximation can be made arbitrarily small. The set up of the parameters σ and n has impacts on the solution accuracy and quality, as well as computational efficiency; thus, we examine their impacts on the approximation errors, problem scale, and computation time in our experiments (see the discussion in Section 6.1.1).

The model including objective function (1), constraints (2)–(3a), (5a), (5b), (6c), (7)–(20), (21), (22c), (22d), (23), and the constraints for reformulating (24b) is called the M_{TSR} problem. It is a mixed-integer linear programming (MILP) model and can be solved by an MILP solver like CPLEX.

5.2. A two-step method for solving the MILP problem

Inspired by the good performance solving the reformulation of the integrated problems in Luan et al. (2018a) and Xu et al. (2017), we develop a two-step method to speed up the solving procedure of the MILP integrated optimization model (i.e., the M_{TSR} problem). In the first step, we aim to generate the full speed profile for each train, which lets the train run as fast as possible, assuming no conflict from the operations of the other running trains. To do this, we relax the safety headway constraints (12)–(16) and rerouting and reordering constraints (17)–(20). In addition, we replace the objective function (1) by

$$Z_{\text{travel}} = \min \sum_{f \in F} \sum_{i:(i,s_f) \in E_f} d_{f,i,s_f}, \quad (25)$$

which minimizes the total travel time of all trains. In this sense, we can obtain a speed trajectory with the minimum travel time for each train. We repeat the computation for trains with multiple routes, in order to get the full speed trajectory for each train along each of its possible routes. We can apply a parallel computation method by performing a train-based and route-based decomposition, in order to obtain the speed trajectories for each train on each route in a short computation time.

In Xu et al. (2017), the full train speed profile was generated without considering the effects of TSR. Differently from that, we generate the upper bound of train speeds by including the impacts of TSR; in such a way, the solution space can be further reduced. The obtained fastest train speed at each block section is used as an upper bound of the train speed at each speed point of the cell, denoted by $\hat{v}_{f,i,j}^{\max}$. We denote the lower bound of the train speed as $\hat{v}_{f,i,j}^{\min}$ and let $\hat{v}_{f,i,j}^{\min} = \hat{v}_{f,i,j}^{\max} - 25$ m/s (consequently, we have $\frac{\hat{v}_{f,i,j}^{\min}}{\hat{v}_{f,i,j}^{\max}}$ ranging from 0.4 to 0.8). Generally, to reduce the effects of TSR and deviations, we rarely let trains run at a speed largely reduced from its possible maximum speed. In some cases, large speed reductions and unplanned stops may occur for resolving train conflicts. Such a setting of train speed lower bound can reduce the solution space, making the computation faster. It should be noted that, the integrated optimization method could dynamically determine the train speed limits based on the rescheduled timetable, as illustrated in constraints (10), thus, the lower bound of the train speed at TSR area is zero. With the obtained lower and upper bound of train speed, we then use the following constraint

$$\hat{v}_{f,i,j}^{\min} \cdot x_{f,i,j} \leq v_{f,i,j,k} \leq \hat{v}_{f,i,j}^{\max} \cdot x_{f,i,j}, \quad \forall f, (i, j) \in E_f, k \in [1, \kappa_{i,j}] \quad (26)$$

to replace (7b), restricting the range of the train speed. In the second step, we solve the M_{TSR} problem with the new constraint (26) to replace (7b).

5.3. Sequential approach

This section develops a sequential approach for the problems of train rescheduling and train control under Temporary Speed Restriction (TSR). The approach estimates the trains being affected by the TSR and then uses this information to generate the train speed trajectories firstly. With the generated train speed trajectories, we then calculate the train running times and use them for rescheduling the train schedule by re-timing, re-ordering, and re-routing. Also, in real-world train operations, the sequential approach is typically used by the dispatchers to reschedule the timetable under TSR.

The optimization of train speed trajectories under TSR is complex. The perturbation caused by TSR has a time duration, i.e., a start time and an end time. As illustrated in Section 4.2, if a train traverses the TSR area within the time duration, then the train speed in the TSR area cannot exceed the required speed limit of TSR. A speed reduction caused by TSR would cause longer running time for the trains that run in the affected area and further affect the train timetable. However, the decisions used to determine whether a train is affected by TSR are related to the actual train entrance and exit times at TSR area. Thus, there is a deep interdependency between the speed limits of trains on block sections, and the rescheduled timetable.

The sequential approach consists of two steps with two individual methods. The main output of the approach is the same as the integrated optimization approach, i.e., one solution including a rescheduled timetable and corresponding train speed graph. Regarding the method used in the first step, we first give a pre-determined set of trains that are estimated to be affected by the TSR and then use the planned train routes and timetable to optimize the train speed trajectories. The model used in this method is called the M_{STP} model, i.e., train Speed Trajectories optimization with a Pre-determined set of the affected trains. Then, we fix the train running time on each block section according to the solutions of speed trajectories in the first step; and use a M_{TRR} model (i.e., Train Rescheduling with fixed Running time of each block section) to optimize the train timetable.

Next, we describe the M_{STP} model and the M_{TRR} model for the methods in the two steps as follows:

- The M_{STP} model - train Speed Trajectories optimization with a Pre-determined amount of the affected trains: including the objective function (1), which minimizes the total absolute deviation time of all trains at all visited stations, subject to constraints (2)–(3a), (5a), (5b), (6c), (7)–(9), (10h), (11a)–(11c), (21), (22c), (22d), (23), and the constraints for linearizing (24b).

Note that we here consider the variable $\xi_{f,i,j}$ in (10h) to be fixed, given as an input parameter, and we denote its fixed value to be $\hat{\xi}_{f,i,j}$, indicating whether train f is affected by TSR on cell (i, j) . The outputs of the first step are the speed profiles for all trains. Based on the obtained speed profiles, the running time $\hat{\tau}_{f,i,j}^{\text{run}}$ of each train at each block section is fixed as the input parameter for the second step.



Fig. 8. A small high-speed railway line.

- The M_{TRR} model - Train Rescheduling with fixed Running time of each block section: including the objective function (1), subject to constraints (10a)–(10g), (11)–(20), and the following constraint:

$$\tau_{f,i,j}^{run} = \hat{\tau}_{f,i,j}^{run} \cdot x_{f,i,j,k}, \quad \forall f \in F, (i,j) \in E_f, k \in [1, x_{i,j}] \quad (27)$$

where $\hat{\tau}_{f,i,j}^{run}$ indicates the given running time for train f on cell (i,j) , which comes from the solutions of the M_{STP} model.

The correct determination of the affected trains in the optimal plan is a difficult target. Let us consider the following two situations:

- Considering a small set of affected trains will cause a short travel time of all trains, as it is recognized that fewer trains will traverse the TSR area at the lower speed required by the TSR. Thus, this situation will lead to a small (ideal) deviation time of the train rescheduling solution. However, if we use the running time obtained by this train trajectory solution, it may have more trains to be affected by the TSR in the rescheduled timetable. This timetable solution is in fact infeasible, since there are trains that violate the train speed limits in the TSR area. This infeasibility can be resolved by re-simulating the speed profile of trains, according to the actual speed constraints. Some trains will incur a larger running time in the TSR area rather than the running time obtained by the train trajectories optimization (i.e., the M_{STP} model). For example, in Fig. 13(d) of Appendix A, the affected trains are G1 to G03. If we assume that trains G01 and G02 are the affected trains, then the running time of train G03 between stations A and B is the normal value (i.e., 3); the deviation time of the rescheduled timetable is lower with the short running time. However, train G03 is actually affected by TSR; the running time of train G03 is 4 rather than 3, then the rescheduled timetable with the running time of imprecisely modeled, and it is not directly feasible for real-world train operation.
- Considering a larger set of the affected trains ensures that no train will violate the speed limit in the TSR area. However, such a solution may incur more train deviations, as it may require a train that is actually unaffected by the TSR to run at a slower speed, causing a longer travel time. As shown in Fig. 13(c) of Appendix A, it assumes that trains G01 to G04 are affected by TSR; however, train G04 runs out of the TSR area (which is not affected by TSR) in the rescheduled timetable (i.e., Fig. 13(c)). Thus, the deviation time of the rescheduled timetable with a larger set of affected trains (i.e., Fig. 13(c)) is longer than the timetable with consideration to vary the train speeds (i.e., Fig. 13(d)).

This approach determines the affected trains according to the initial timetable and the influence of the TSR case. Thus, the number of the affected trains is typically a larger one. For example, in Fig. 13(a) of Appendix A, the TSR influences trains from G1 to G4 in the initial timetable, thus, the predetermined affected trains are G1, G2, G3, and G4.

6. Case study

In this section, we use three sets of numerical examples to illustrate the efficiency and effectiveness of the proposed MILP model for the integrated optimization problem. We adopt the IBM ILOG CPLEX Optimization Studio 12.7.1.0 to solve the MILP problem. The parameters of MILP solver in CPLEX are with default settings. The following experiments are all performed on a server with Intel® Xeon® CPU E5-2660 v4 @ 2.00 GHz processor and 512 GB RAM.

6.1. Description of a small-scale instance

We consider a small case of high-speed railway line with 4 stations, with a length of about 192 km. Fig. 8 shows the detailed layout with all block sections of the line. The line is composed of 106 nodes and 113 cells, with an average length of 1707 m per cell. The number in the line indicates the position of the node, unit:km. Only the direction from Station A to Station D is considered, since the operation in different directions are independent. We consider a Temporary Speed Restriction (TSR) area, indicated by the red area shown in Fig. 8. This TSR influences 32 cells, about 57 km length in total. The tunnels of the line are indicated by the blue bars, as shown in Fig. 8. Figs. 15(a) and 15(b) in Appendix B show the grade profile and the curve radius of the line respectively. We consider to dispatch 15 trains, including two types of rolling stock, i.e., EMUs of CRH380 and CRH5 with normal speed of 83 m/s and 69 m/s respectively, the parameters of which are shown in Tables 7 and 8 in Appendix B respectively. Moreover, we set the maximum change of acceleration rate between two adjacent space intervals to be $\Delta a^{\max} = 0.5 \text{ m/s}^2$, which is used for ensuring the passenger-riding comfort (Lai et al., 2020).

6.1.1. Experimental results, performance evaluation of the three parameters σ , n , and κ

We use the small instance shown in Fig. 8 as the test bed, testing several cases with different values of the three parameters σ , n , and κ . The detailed experimental results and the discussions are presented in Appendix C, in terms of the approximation error, the problem size, and the computation time. According to the experimental results, for all the experiments reported in the remaining part of the paper, we use the following settings: $\sigma = 3$, $n = 3$, and $\kappa = 2$, due to the good balance between computational efficiency (i.e., shorter computation time) and model precision (i.e., smaller approximation error).

6.1.2. Benefits of the rerouting and reordering measures

Here we aim to demonstrate the benefits of rerouting and reordering for the integrated optimization problem. Recall that the proposed MILP model considers the control measures of Retiming (RT), Respeeding (RS), Reordering (RO) and Rerouting (RR), whereas the rerouting is mostly neglected in the integrated optimization approaches available in the literature.

We design four sets of experiments to evaluate the benefits of rerouting and reordering. Regarding the rerouting measure, we consider platform selection flexibility in each station (i.e., local-rerouting). For the small-scale network shown in Fig. 8, there are 72 candidate routes for each train. We are interested in solution quality (i.e., total train deviation reduction) for the following four scenarios:

- RT+RS+RO+RR (this work): using the MILP model with consideration of RT, RS, RO, and RR;
- RT+RS+RO: using the MILP model considering RT, RS and RO, but without RR;
- RT+RS+RR: using the MILP model considering RT, RS and RR, but without RO;
- RT+RS: using the MILP model considering only RT and RS, but without RO and RR.

There are four parameters for a TSR case: the affected block sections, the start time, the end time, and the limited speed. In this section and Appendix D, the affected block sections are shown in Fig. 8, and the start time for all the TSR cases is fixed as 19:20 pm. The end time is then (19:20 pm) + *time duration of the TSR* (unit: hour), five time duration as 0.5 h, 1 h, 1.5 h, 2 h, and 2.5 h are considered in this example. Five levels of the speed limits are considered as 22.22, 27.78, 41.67, 55.56, and 69.44, unit: m/s. By varying the values of the parameters of the end time and the limited speed, we obtain 25 sets of the experimental results. The influence of different time duration of the TSR and speed limits are reported in Appendix D. Fig. 9 comparatively presents the results of these four scenarios, as total deviation time, and the improvement in solution quality. In Fig. 9(a), the lines with different symbols and colors indicate the experimental results of each scenario, respectively. A missing line is used to distinguish the experimental results among different values of the limited speeds. In Fig. 9(b), each bar indicates the improvement in solution quality for the three scenarios, i.e., RT+RS+RO+RR, RT+RS+RO, and RT+RS+RR, when comparing the solutions obtained by RT+RS scenario, e.g., $\frac{RT+RS+RR+RO \text{ solution} - RT+RS \text{ solution}}{RT+RS \text{ solution}} \times 100\%$, and refer to the Y-axis on the left-hand side. Moreover, the average improvement of the three scenarios for each limited speed is shown by lines, and refers to the Y-axis on the right-hand side. In this example, we show the optimal solutions of all cases.

As illustrated in Fig. 9(a), we can see that the solution quality of the RT+RS scenario is the worst among the four scenarios, as the pink lines are obviously higher than the other lines. In addition, we note that, the solution quality of the four scenarios is almost same for the cases with limited speeds of 55 m/s and 69 m/s. One possible reason is that, the deviation of these cases mainly comes from the speed limits, rather than the conflicts among trains. The dispatching measures, e.g., Rerouting and Reordering, have better performances when solving the conflicts among trains, thus, the performance improvement of solutions by multiple dispatching measures is not significantly evident for the cases with high limited speed.

In addition, comparing the three approaches (i.e., RT+RS+RO+RR, RT+RS+RO, and RT+RS+RR) with approach RT+RS, as shown in Fig. 9, the improvements of the three approaches are different among cases. Regarding the TSR cases with higher severity (i.e., lower limited speed or longer time duration) the improvements of the three approaches are typically better, as the lines decreasing with the increasing limited speed in Fig. 9(b), i.e., optimization is most useful for more serious disruptions. For example, the average value of the average improvements of the three approaches for the cases of limited speed of 22 m/s, 27 m/s and 41 m/s is 12%, which is much better than the average value of the average improvements for the cases of limited speed of 55 m/s and 69 m/s (i.e., 1%). With respect to the performance of the three approaches, the average improvement is 9%, 7%, 6%, respectively for the approach of RT+RS+RO+RR, RT+RS+RO, and RT+RS+RR.

Overall, the solution quality of the RT+RS+RO+RR approach (proposed in this paper) is the best, as its total deviation time is minimal among the four approaches for the 25 TSR cases. In comparison with the RT+RS+RO approach (i.e., previous integrated optimization method without the consideration of rerouting), our method (i.e., RT+RS+RO+RR) achieves a 3% average improvement in solution quality for all the TSR cases.

6.2. Numerical experiments on a realistic high-speed railway line

In this section, we consider a real-world high-speed railway line in China. In the implementation, the railway line parameters are taken from real-world data. The high-speed railway line under consideration has 8 stations, and it has a length of about 431 km. Fig. 10 shows the layout of the line. The line is composed of 247 cells with 494 space intervals. In this example, we consider to dispatch 15 trains. The positions of some important nodes (e.g., the enter node of each station, the start and end nodes of a TSR) are shown in the figure, unit: km. The direction from Station SA to Station SH is considered, without loss of generality, as the train operations in different directions are independent.

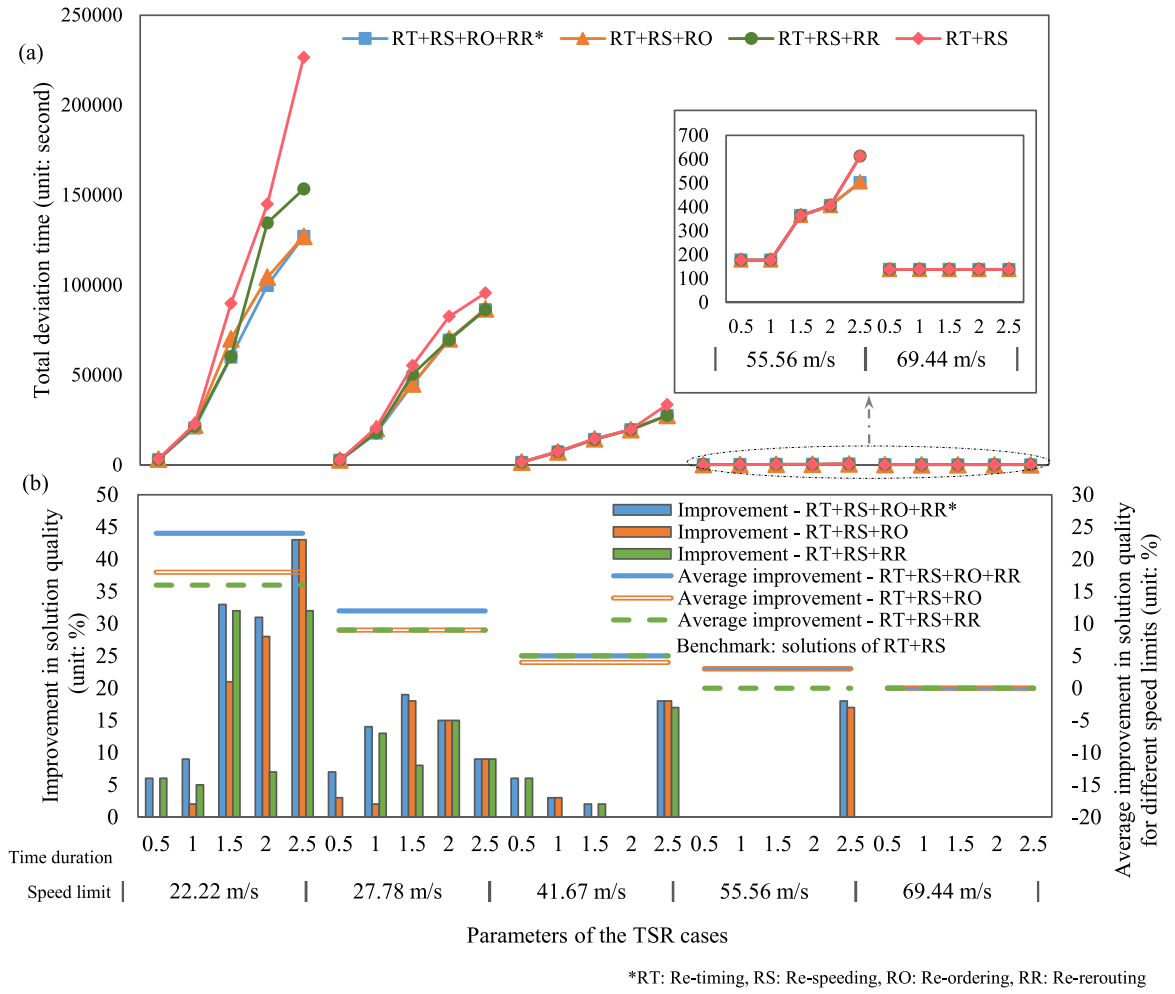


Fig. 9. Results of the scenarios of RT+RS+RO+RR, RT+RS+RO, RT+RS+RR, and RT+RS.

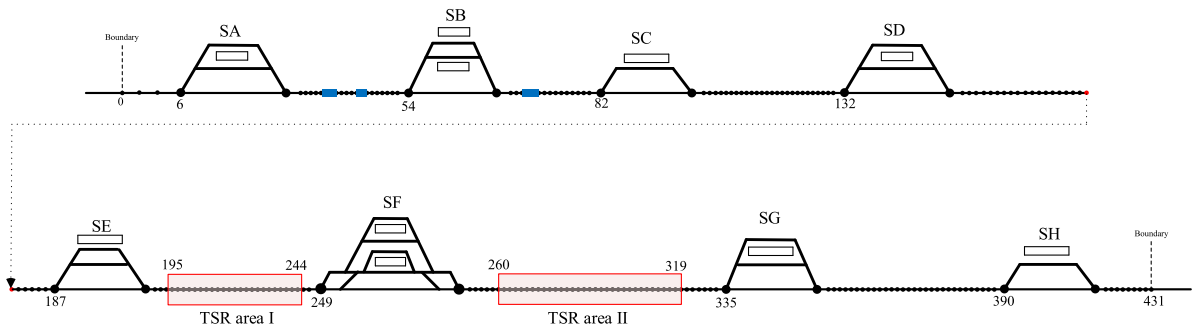


Fig. 10. Illustration of a high-speed railway line.

The blue bars in Fig. 10 indicate the tunnels of the line. Figs. 16(a) and 16(b) in Appendix B show the grade profile and the curve radius of the line respectively. Two types of rolling stock are considered, including EMUs of CRH380 and CRH5. The relevant parameters of the rolling stock are respectively shown in Tables 7 and 8 of Section 6.1. We consider two TSR areas, indicated by red areas shown in Fig. 10. Table 3 gives detailed information about the two TSRs, including the start/end position, start/end time, the number of affected block sections, and the length of the affected area, to be used in this part of the experiments.

Table 3

The relevant parameters of the two TSR areas under consideration.

TSR area	Start position (unit: km)	End position (unit: km)	Number of affected block sections	Length of the affected area (unit: km)	Start time	End time
area I	195	244	16	39.98	18:10	19:40
area II	260	319	32	62.53	18:40	20:10

Table 4

The limited speed value of the three TSR cases.

TSR case	Speed limit (unit: m/s)	
	TSR area I	TSR area II
Case A	55.56	55.56
Case B	55.56	44.44
Case C	33.33	33.33

Table 5

Total deviation time with consideration of flexible and fixed route.

TSR case	Total deviation time (s)-RT+RS+RO		Total deviation time (s)-RT+RS+RO+RR	
	180 (s)	600 (s)	180 (s)	600 (s)
Case A	12 719	12 394	12 305	11 806
Case B	15 809	14 368	14 228	13 260
Case C	22 907	21 477	20 987	19 803

6.2.1. Experimental results, benefits of the rerouting measure under different TSR

In this section, we investigate the performance of the rerouting measure under different TSR cases. Similar to Section 6.1.2, we consider to flexible platform selection in stations. On the network of Fig. 10, there are 3888 candidate routes for each train in theory. However, the amount of candidate routes for a train is typically less than the theoretical value (i.e., 3888). We could reduce the candidate routes based on the actual business rules for the track usage, to reduce the amount of candidate routes and further reducing computational times. For example, if a train passes through a station, it would prefer use the tracks in the main running lines than go to a siding track. Therefore, the total value of the candidate routes for the trains is 197 in this example.

We consider three TSR cases, where different speed limits are required for the two TSR areas as shown in Tables 3 and 4. The CPLEX solving process is terminated by considering the computation time limits of 180 s and 600 s, and we then output the best feasible solution obtained within the given time limit. Table 5 shows the experimental results of the flexible route option and fixed route option under different settings of the TSR cases. As we can see from the figure, our method could obtain feasible solutions within 180 s, and achieve a 4%, 6%, and 5% improvement for the solution quality in 10 min, respectively for Case A, B and C.

Firstly, we focus on the results with different TSR disturbances. Table 5 reports the total deviation time of the TSR cases A, B, and C respectively. As shown, the total deviation time increases with the decrease of the speed limit. Intuitively, the values of the total deviation time for Case C are bigger than the values of Cases A and B. We then analyze the benefits of rerouting. As illustrated, the quality of the solutions obtained by considering flexible train routes (i.e., RT+RS+RO+RR) is better than that with fixed train routes (i.e., RT+RS+RO). Compared to the results with fixed train routes, considering flexible train routes enables us to decrease the total deviation time by 4%, 9%, and 8% for the TSR case A, B, and C respectively. In other words, our integrated optimization method can achieve an average improvement of 7% in the solution quality by considering the control action of rerouting.

6.2.2. Benefits of the integration, compared with the sequential approach

We here compare the integrated optimization method with the sequential approach of train rescheduling and train control to investigate the benefits of the integration. The sequential adjustment approach is proposed in Section 5.3, where the two problems of train rescheduling and train control are solved in a sequential manner. In what follows, we dispatch 15 trains and still consider the two TSR areas, denoted by TSR area I and II in Table 3, with the speed limit of 55.56 m/s and 44.44 m/s respectively.

In this example, the total deviation time of the integrated optimization method and sequential approach is 13 260 s and 14 600 s, respectively. Compared with the initial train timetable, the amount of the route (order) changing in the integrated method and sequential approach is 3 (1) and 2 (0), respectively. Moreover, the integrated optimization method could consider the influences among trains and dynamically optimize the train speeds and arrival/departure times of the trains at the same time. Therefore, a rescheduled timetable with a smaller deviation time could be obtained. Compared with the integrated optimization method, the total deviation times of the sequential approach increase by $\frac{14600-13260}{14600} \times 100\% = 9\%$.

The amount of the trains running with the normal speed (i.e., the trains are not influenced by TSR) for the two approaches are both 6 (i.e., G10 to G15). However, the sequential approach should first give a set of affected trains to calculate the train speed trajectories. Some trains may have left the TSR area in the rescheduled timetable (e.g., trains G8 and G9), but the running times are calculated according to the lower train speeds, which incur a larger running time in the TSR area. As shown in Fig. 11, compared

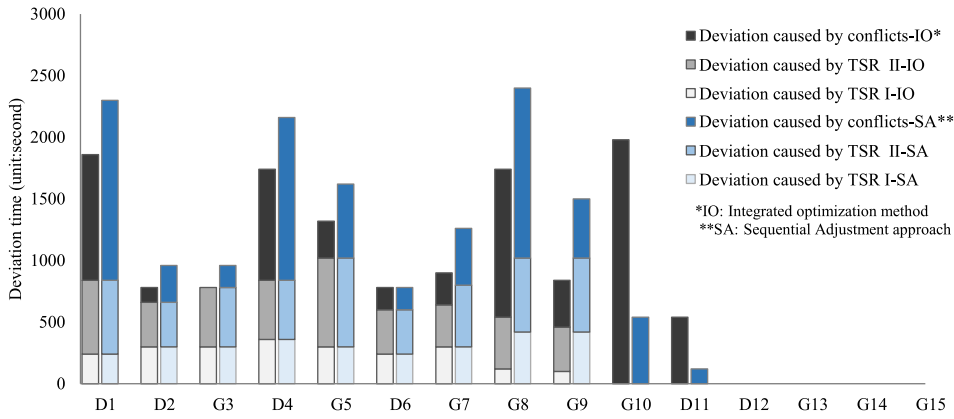


Fig. 11. Deviation time component of the rescheduled timetables.

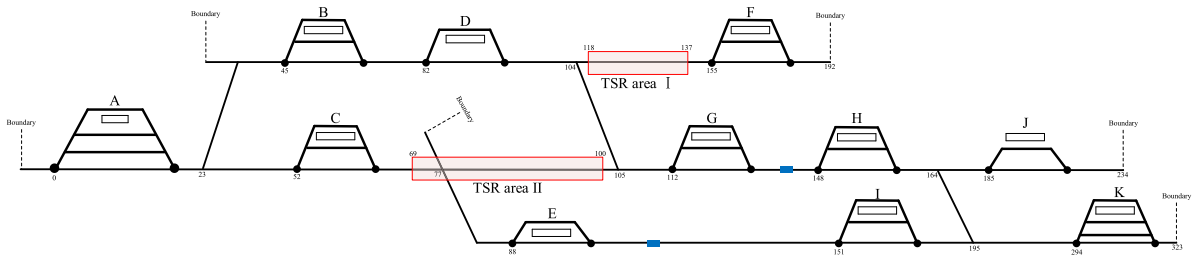


Fig. 12. Illustration of a high-speed railway network.

with the integrated optimization method, the deviation times of G8 and G9 on TSR area is higher in the sequential approach. Therefore, compared to the sequential adjustment approach, the integrated optimization method could dynamically determine the trains affected by the TSR at a microscopic level, thus the approach could avoid the unnecessary speed reduction and further reduce the train deviations. In this example, compared with the integrated optimization method, the deviation caused by TSR of the sequential approach increases by $\frac{7580-6380}{6380} \times 100\% = 18\%$.

Above all, it is evident that the integrated optimization method performs better since it can obtain a rescheduled timetable with better quality, i.e., achieving smaller deviation times. This indicates that, by the integrated optimization of the train rescheduling and the train control, the deviation time can be reduced in comparison with the sequential approach of the train rescheduling and the train control.

6.3. Performance on instances with different scales

In this section, we consider a high-speed railway network, to compare the performances of the method among the instances with small-scale line, the medium-scale line, and the large-scale network. The instances described in Sections 6.1 and 6.2 are regarded as the small-scale and medium-scale instances, respectively. We further consider a railway network with 11 stations as shown in Fig. 12. We dispatch 15 trains in this example. The stop plan of each train is reported in Fig. 17 in Appendix B. There are two TSR areas in this example. The limited speed value of TSR area I and area II is 44.44 m/s and 22.22 m/s, respectively. The start time is 10 am and the end time is 11 am, both for two areas. The earliest start time of the first train on each line at the original station is 10 am.

Table 6 shows the performances of the three instances. The first column in Table 6 reports the number of nodes and cells, and the total length of the lines for the small-scale line in Fig. 8, the medium-scale line in Fig. 10, and the large-scale network in Fig. 12. We report the experimental results of optimization methods RT+RS+RO, sequentially, and RT+RS+RR. Table 6 shows the percentage of the optimal solutions obtained in 600 s and 3600 s for each case.

As shown, the number of constraints and variables increases with the instance scale. Specifically, for method RT+RS+RO, the complexity of the large-scale network does not increase compared with medium scale line. The reason for this issue probably is: if we do not consider changing the routes of the trains in the network (i.e., without RR dispatching measure), the trains could only schedule in their planned line in the network (less passing stations compared with medium-scale instance), then the candidate orders among trains are less than the medium case. Overall, the problem size increases with the instance scale, revealing the increase of complexity in problem solving. In the cases of the small-scale instance in Section 6.1.2, the optimization method RT+RS+RO (RT+RS+RO+RR) could obtain the optimal solutions in 600 s and 3600 s for 28% (16%) and 36% (44%) cases, respectively. However, it is difficult

Table 6
Performance of instances with different scales.

Instance	Optimization method	Number of constraints	Number of variables	Percentage of optimal solutions		Objective value	Improvement of RT+RS+RO+RR
				600 s	3600 s		
Small-scale (106,113,192) ^a	RT+RS+RO	314 909	162 049	28%	36%	22 064	6%
	RT+RS+RO+RR	429 937	241 800	16%	44%	20 605	–
Medium-scale (235,247,431)	RT+RS+RO	486 459	276 574	0%	33%	14 355	8%
	Sequential	147 031	83 980	100%	–	14 600	9%
	RT+RS+RO+RR	631 779	301 790	0%	33%	13 180	–
Large-scale (345,358,658)	RT+RS+RO	430 666	240 310	0%	0%	8143	36%
	Sequential	175 653	85 461	–	100%	5550	6%
	RT+RS+RO+RR	807 891	354 379	0%	0%	5189	–

^a(a,b,c): a is the number of nodes, b is the number of cells, c is the total length of the lines (unit: km).

to obtain the optimal solution for the medium and large in 3600 s. Regarding the sequential approach, the optimal solution could be obtained quickly, i.e., 600 s and 3600 s for medium and large instance, respectively.

Regarding the objective value term, we output the best feasible solution obtained in 3600 s for the three instances. For the small-scale instance, the TSR case is with limit speed value of 22.22 m/s and a duration time of 1 h. For the medium-scale instance, we show the results of Case B. In addition, we give a comparison in the solution quality of the three approaches among instances. As shown in Table 6, it is clear that the solution quality of method RT+RS+RO+RR is best among cases. Then, we give a comparison on the solution quality among the methods. The value in the last column is calculated by $\frac{[RT+RS+RR+RO \text{ solution} - RT+RS+RO \text{ solution}]}{RT+RS+RO \text{ solution}} \times 100\%$, and $\frac{[RT+RS+RR+RO \text{ solution} - \text{Sequential solution}]}{\text{Sequential solution}} \times 100\%$, for approaches RT+RS+RO and sequentially, respectively. We can see that, compared with approaches RT+RS+RO and sequential, the method RT+RO+RS+RR (i.e., our proposed integrated optimization method) could achieve an improvement of 6%–36% in solution quality.

7. Conclusions and future research

In this paper, we have studied the problem of integrating train rescheduling and train control in the case that Temporary Speed Restriction (TSR) occurs during high-speed railway operations. We have developed a mixed-integer nonlinear programming (MINLP) model, which is further reformulated into a mixed-integer linear programming model by means of piecewise linear approximation. The MILP model simultaneously delivers the train schedules and the train speed trajectories, with the minimization of train deviations at the stations. In the model, the train speed trajectory is calculated based on the traction/braking force, the basic resistance, and the line resistance caused by grades, curves, and tunnels. A set of constraints for train rescheduling and train control under TSR area is designed to further reduce the train deviations. Moreover, the running time of a train on a block section and the safety headway time between two adjacent trains dynamically depend on the real speed. Further, the model considers the operation transitions among the operation periods, i.e., normal, disturbed, and recovery periods. Passenger riding comfort is ensured when generating the train speed profiles. A two-step approach has been designed to speed up the solving procedure of the MILP model. Moreover, we have proposed a sequential approach for interacting the two problems of train rescheduling and train control, to be compared with the integrated optimization method. Experiments have been conducted based on three instances: a small-scale line, a medium-scale instance adopted from the real-world high-speed railway line in China, and a large-scale network, to evaluate the effectiveness of the proposed methods. According to the experimental results, we find that our integrated optimization method leads to an average improvement of 3–36% in solution quality, compared with the integrated approach without considering train rerouting. Additionally, we find that our integrated optimization method outperforms the sequential approach by 6–9% in solution quality.

Future research includes the following extensions of the problem. First, although the two-step method is efficient to obtain feasible solutions within a suitable computation time, we can still consider better optimization algorithm to further improve the computational efficiency. Second, we focus more on the investigation of the train rescheduling aspect; thus, we consider only the objective related to the traffic operations (i.e., deviation reduction). In the future study, we can include the objective of the train trajectory optimization problem, i.e., calculating the energy consumption and improving energy efficiency of train operation. Finally, the proposed formulation method can be potentially used for modeling moving block, with further modifications and extensions.

CRediT authorship contribution statement

Shihui Long: Methodology, Software, Investigation, Visualization, Writing – original draft. **Lingyun Meng:** Conceptualization, Supervision, Project administration, Methodology, Writing – review & editing. **Yihui Wang:** Methodology, Software, Validation, Writing – review & editing. **Jianrui Miao:** Methodology, Software, Writing – review & editing. **Xiaojie Luan:** Methodology, Resources, Software, Validation, Writing – review & editing. **Francesco Corman:** Conceptualization, Resources, Methodology, Supervision, Writing – review & editing.

Acknowledgments

This work is supported by the National Nature Science Foundation of China (72022003, 61790573), Yunnan Xing Dian Talents Plan young, China (KKRD202202112, 2022), State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, China (RCS2019ZJ001), and the Swiss National Science Foundation under Project 1481210/DADA.

Appendix A. Illustrative example

We describe an example to illustrate the benefits of integrated optimization of multiple dispatching decisions of train rescheduling and train control in reducing the total deviation times, including the decisions of Retiming (RT), Respeeding (RS), Reordering (RO), and Rerouting (RR), see an illustrative example in Fig. 13. In the example, five trains, named G01 to G05, travel from station A to station C. The example illustrates the train routes at station B and the train timetable. As illustrated, there are three tracks at station B, named P0, P1, and P2. In Fig. 13, we use the lines with arrow to represent the different operation periods of the considered time horizon: green line, red line, and yellow line indicates the normal, disturbed, and recovery operation periods, respectively. It should be noted that, we use microscopic level of the railway network to establish the model, however, a macroscopic-level example is shown here to better illustrate the benefits of considering the multiple dispatching decisions. Moreover, the microscopic train speed trajectories could show the detailed train operation process among stations.

Fig. 13(a) shows an initial train timetable and a TSR case. The planned running times of trains G01 to G03 and G05 between two stations are the same (i.e., 3), which are shorter than the running time of train G04 (i.e., 5). A longer running time corresponds to a lower train speed of train G04, as shown in Fig. 14(a). The TSR occurs between stations A and B, with a time duration that starts at time 1 and ends at time 12. Therefore, Fig. 14 only shows the speed trajectories for all the trains between stations A and B, corresponding to the timetables of Fig. 13.

According to the initial train timetable and the parameters of the TSR case, the TSR influences four trains, i.e., G01 to G04. The running time of the influenced trains is increased to 6, also the headway time is increased, according to the principles of train separation. If we just consider to straightforwardly extend the running times and adjust the timetable of affected trains, train operation conflicts may occur due to the mutual influence between trains, as illustrated in Fig. 13(b). The conflicts lead to an infeasible timetable. Therefore, we need to reschedule the train timetable requiring adjustments or re-computation, before implementation. The model proposed directly tackles this problem, i.e. determines the best way of adjusting traffic, and train speeds, for a specified given objective.

This example considers four approaches to dispatch trains under TSR.

(1). Retiming, RT

This approach only changes the planned arrival and departure times at the station. For the trains influenced by TSR, i.e., G01 to G04, this approach recalculates the running time of the trains traversing the TSR area with the limited speeds. The rescheduled timetable satisfies the constraints of train operation, e.g., train running time constraint and safety headway time constraint. A feasible rescheduled timetable and the corresponding train speed profiles are illustrated in Figs. 13(c) and 14(b), respectively. As shown, the total deviation times of the affected trains (i.e., from G01 to G04) are mostly coming from the increased running times and headway times. Then, train G05 is delayed due to the delay propagation. In total, the deviation time of all trains is 51.5 for approach RT.

However, approach RT may lead to redundant deviation times due to unnecessary speed reduction. We can see from Fig. 13(c), train G03 traverses out the TSR area after time 12, as well as train G04 does not go through the TSR area, whereas the running times used in this approach are assumed that they will be affected by TSR. Thus, the following approaches are proposed to avoid the unnecessary speed reduction.

(2). Integrated Optimization of Retiming and Respeeding, IO-RT&RS

This approach considers the integrated optimization of RT and RS. In detail, the decisions include the arrival time, the departure time and the speed of trains at each block section. Thus, the train speed limits at TSR area dynamically depend on the rescheduled arrival and departure times, rather than being a fixed and predetermined value. A feasible timetable is shown in Fig. 13(d). Compared with Fig. 13(c), the deviation time of Fig. 13(d) (i.e., 48) decreases due to the increased speeds and the corresponding shorter headway times of trains G03 and G04. In Fig. 14(c), train G03 accelerates from the temporary speed limit to a higher speed limit, as well as train G04 runs at a normal speed. Therefore, the train speed profiles show that train G03 runs out of the TSR area and train G04 will not be influenced by the TSR, which are consistent with the timetable of Fig. 13(d).

As shown, we could obtain a timetable with shorter deviation times by integrated optimization of RT and RS, compared with a single decision of RT. The paper supports the claim that it is possible to get a rescheduled timetable with fewer deviations by integrated optimization of more dispatching measures, based on a sophisticated mathematical programming approach. Therefore, the following methods are proposed to obtain a timetable with lower deviation times.

(3). Integrated Optimization of Retiming, Respeeding, and Reordering, IO-RT&RS&RO

This approach considers changing the train orders between trains, compared with approach IO-RT&RS. For TSR case, one of the advantages of changing train orders is to reduce the speed difference between the planned train speed and the limited speed. As shown in Fig. 13(e) of this example, the order of train G03 and G04 is changed, because the planned speed of train G04 is lower than the speed of train G03. In the rescheduled timetable, train G03 is not subject to the TSR, and therefore traverses the TSR area in a shorter running time; actually, train G03 can run at a normal speed as shown in Fig. 14(d). Overall, the total deviation time reduces to 46 due to integrating RT, RS, and RO.

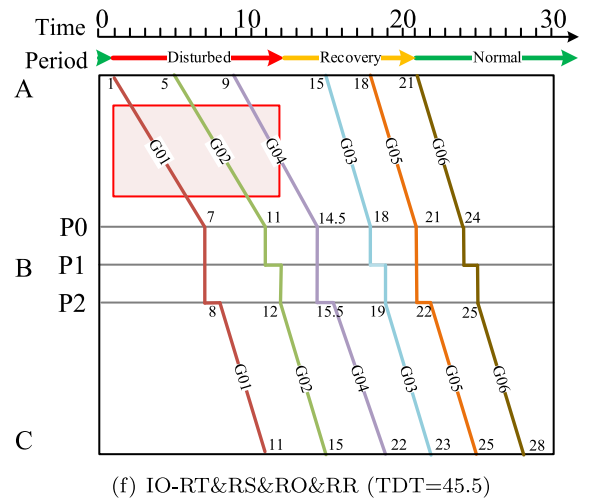
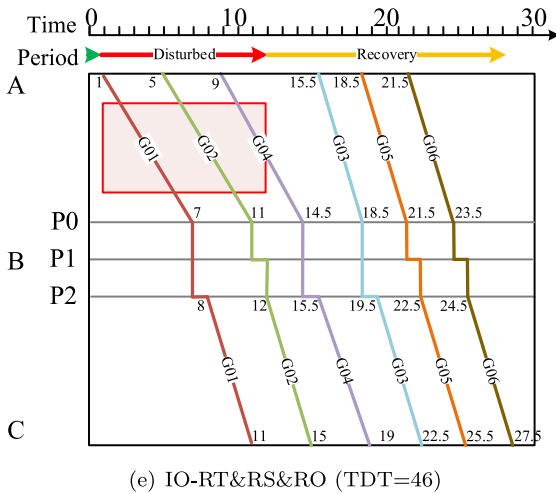
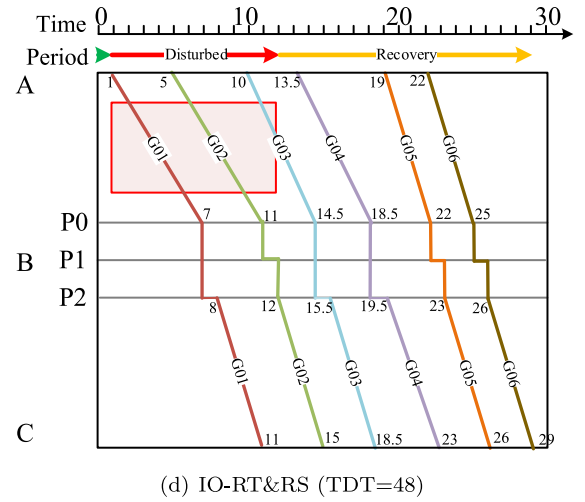
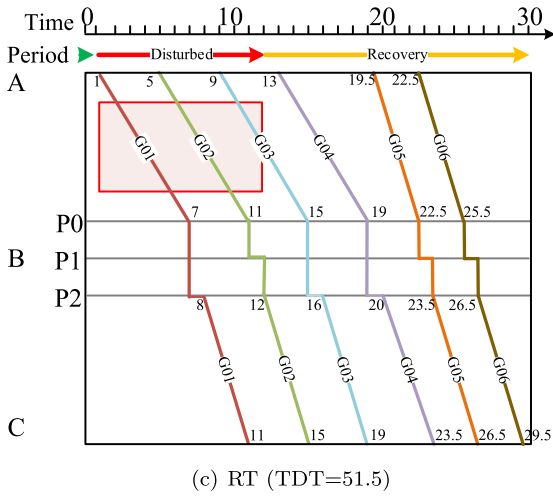
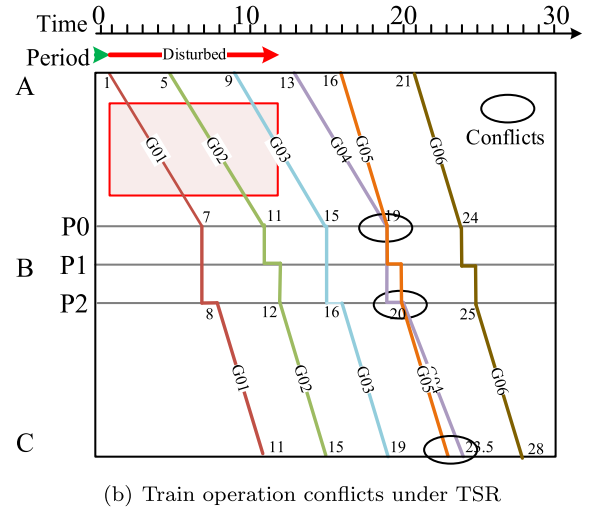
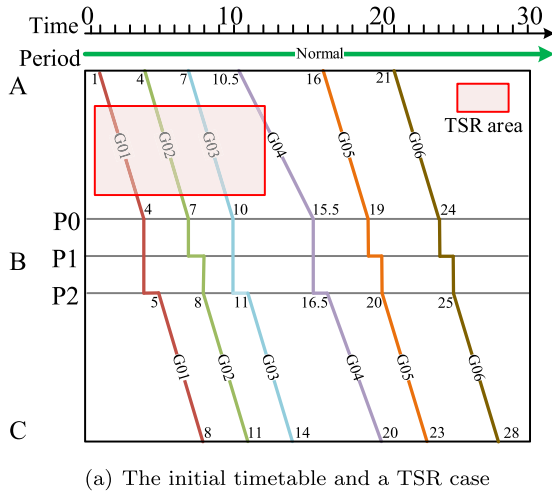
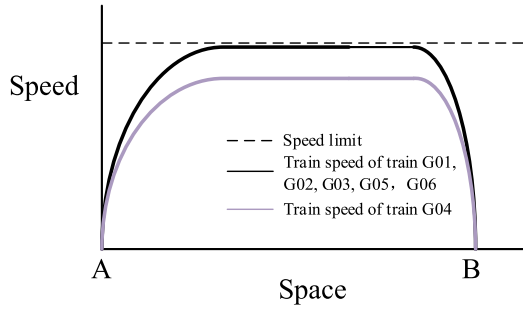
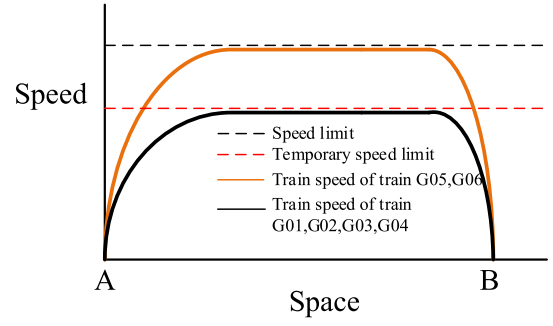


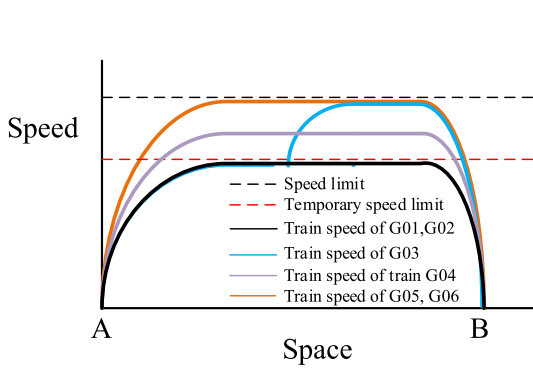
Fig. 13. Illustration of integrated optimization of RT, RS, RO, and RR under TSR (TDT: Total Deviation Time).



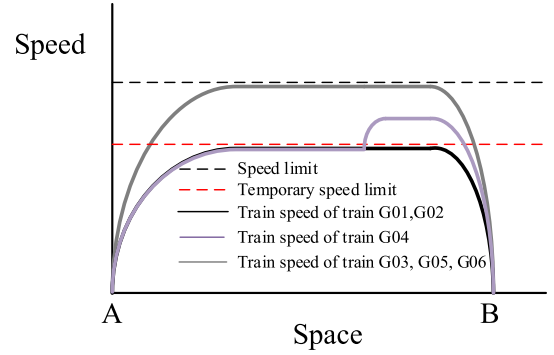
(a) The train speed trajectories for Figure 13(a)



(b) The train speed trajectories for Figure 13(c)



(c) The train speed trajectories for Figure 13(d)



(d) The train speed trajectories for Figure 13(e) and Figure 13(f)

Fig. 14. The train speed trajectories between stations A and B for the timetables illustrated in Fig. 13.

(4). Integrated Optimization of Retiming, Respeeding, Reordering and Rerouting, IO-RT&RS&RO&RR

A safety time is needed between two adjacent trains to use the same platform. Thus, by changing the platforms that the trains require, the deviation times of the obtained timetable may decrease. Moreover, to change the order of trains as one train could use a siding platform when it can be overtaken, which allows more possibility for reordering compared to the approach without considering the reordering decision. In this example, train G03 changes the platform at station B from P2 to P1, as well as train G05 uses platform P2 in the rescheduled timetable. As a result, the deviation time decreases to 45.5 by considering integrated optimization of Retiming, Respeeding, Reordering, and Rerouting, due to better usage of the available.

In conclusion, our integrated optimization method could not only avoid unnecessary speed reduction under TSR, but also decrease train deviation times by using multiple dispatching decisions at the same time. Further, our method could let the train operation go back to the normal operation period quickly (i.e., less recovery time of backing to the normal operation). Our method of simultaneously optimizing the train orders, speed profiles, train arrival and departure times, and train routes could obtain a timetable with the shortest deviation times and recovery times.

Appendix B. The input parameters of the line in Section 6

Figs. 15(a) and 15(b) show the grade profile and the curve radius of the line illustrated in Fig. 8 respectively. Figs. 16(a) and 16(b) show the grade profile (i.e., slope) and the curve radius of the line illustrated in Fig. 10 respectively. Tables 7 and 8 show the parameters of EMU CRH380 and EMU CRH5, respectively.

Regarding the train operation in the railway network of Fig. 12, trains could use different routes to reach their destination. Fig. 17 shows the stop plan of trains for railway network in Fig. 12.

Appendix C. The experimental results of the three parameters σ , n , and κ

Recall that the piecewise linear approximation method is used to linearize the nonlinear terms in constraints (3b), (4), (5b), (6a), and (24b) with different piecewise linear portions (i.e., σ for constraints (3b), (4), (5b), and (6a), and n for constraint (24b)). As illustrated in Section 3.2, each cell (i, j) is divided into $\kappa_{i,j}$ intervals to model the train movement, and here we also study the

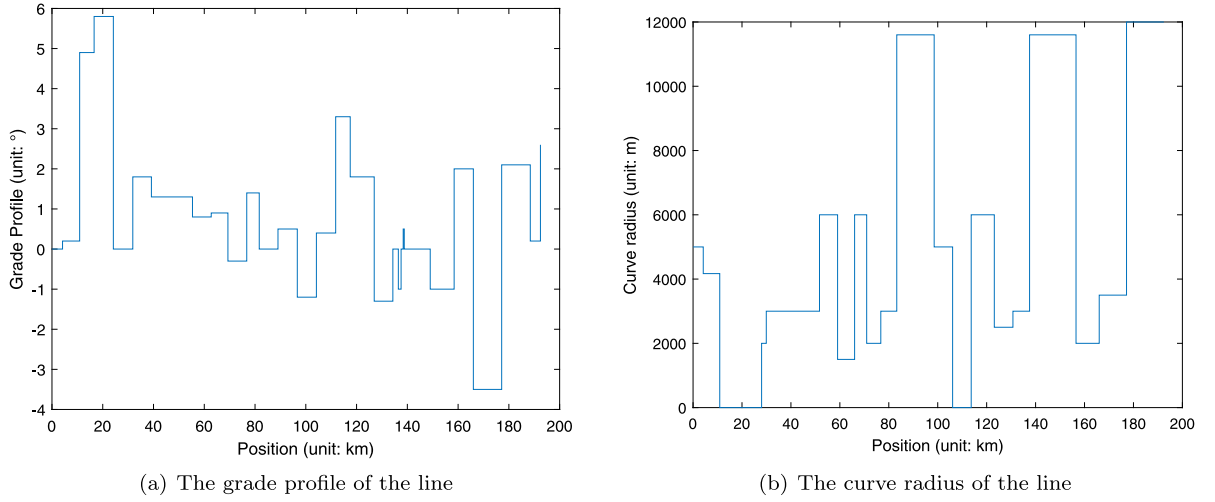


Fig. 15. The grade profile and curve radius of the line in Section 6.

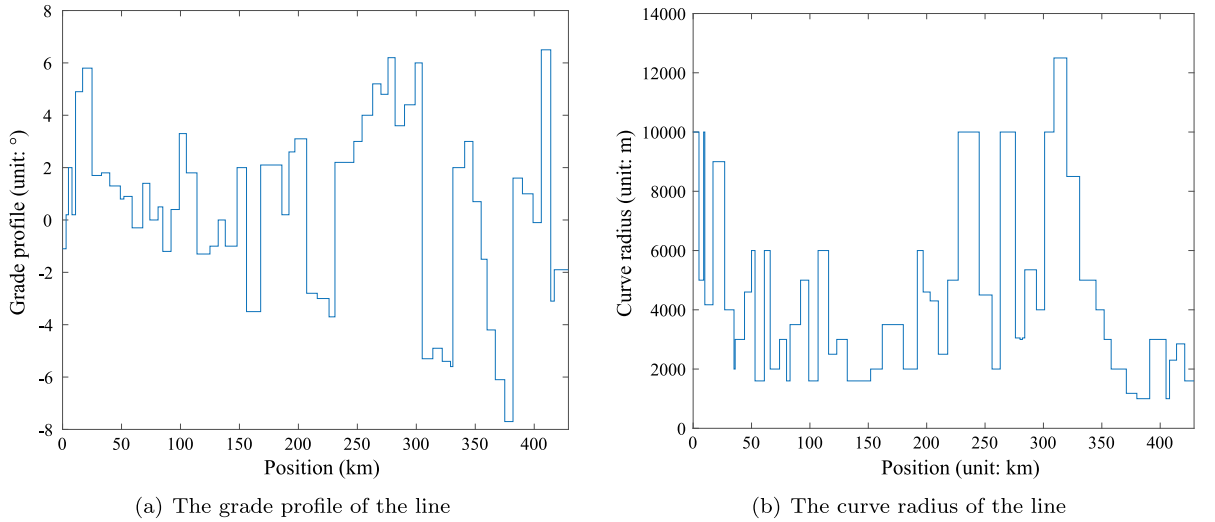


Fig. 16. The grade profile and curve radius of the line in Section 6.2.

Table 7

Parameters of EMU CRH380.

Property	Symbol	Value
Train mass [unit: kg]	m_f	429 200
Train length [unit: m]	L_f^{tra}	203.5
Basic resistance [N/kg]	$R_{f,i,j,k}^{basic}$	$A = 0.55, B = 0.004, C = 0.000109$
Mass factor [–]	ρ_f	1.08
Maximum traction force [N]	u_f^{trc}	$c_1 = 280\,000, c_2 = -245, c_3 = 0$, if $v_{f,i,j,k} \in [0, 36.11]$ $c_1 = 405\,968.3, c_2 = -1434, c_3 = 1.693$, if $v_{f,i,j,k} \in [36.11, 83.33]$
Maximum braking force [N]	$u_{f,i,j,k}^{brk}$	–555 000

effects of different values of $\kappa_{i,j}$ on the results. We here consider one train traversing the line and one-hour TSR starting at 10 am with the maximum speed limit of 22.22 m/s. When we vary the values of a parameter for the experiments, the values of the other two parameters are set to be 3.

We present in Figs. 18 to 20, respectively the investigation results of the three parameters, in terms of the approximation error, the problem size, and the computation time. The approximation errors are shown as percentage, calculated by $\frac{|\text{approximated value} - \text{actual value}|}{\text{actual value}} \times 100\%$.

Table 8
Parameters of EMU CRH5.

Property	Symbol	Value
Train mass [unit: kg]	m_f	500 000
Train length [unit: m]	L_f^{tra}	211.5
Basic resistance [N/kg]	$R_{f,i,j,k}^{basic}$	$A = 0.69, B = 0.006, C = 0.000146$
Mass factor [–]	ρ_f	1.08
Maximum traction force [N]	u_f^{trc}	$c_1 = 337\,250, c_2 = 0, c_3 = 0$, if $v_{f,i,j,k} \in [0, 13.89]$ $c_1 = 453\,020, c_2 = -2564, c_3 = 4.972$, if $v_{f,i,j,k} \in [13.89, 69.44]$
Maximum braking force [N]	$u_{f,i,j,k}^{brk}$	–500 000

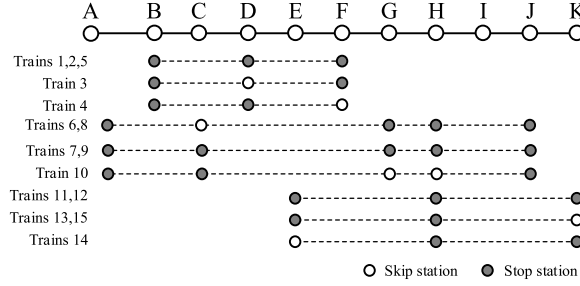


Fig. 17. The train stop plan of the railway network in Fig. 12.

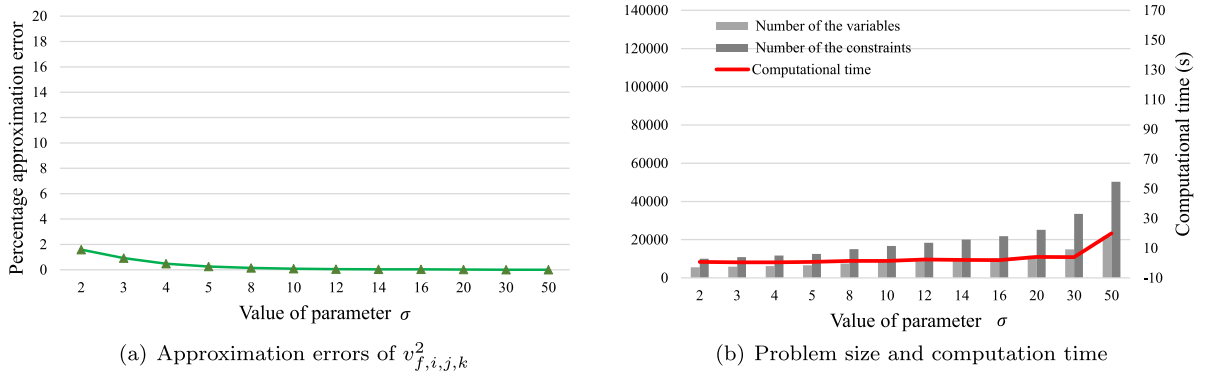


Fig. 18. Performance comparison with different values of σ .

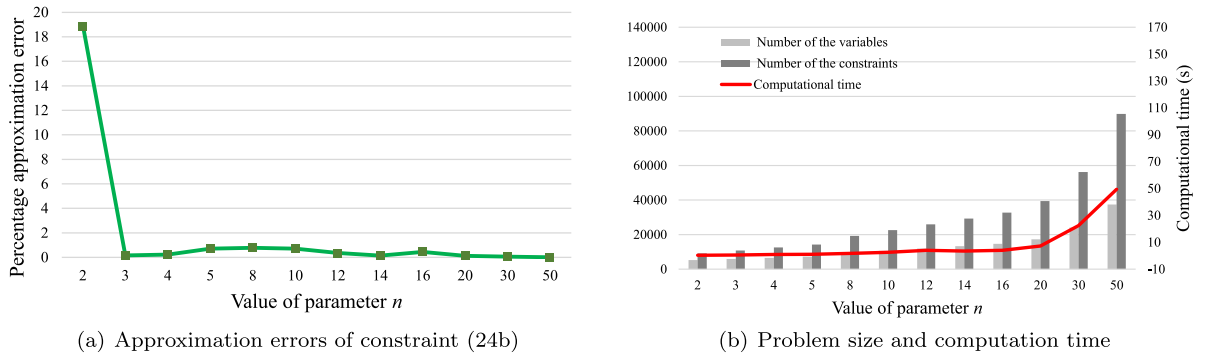


Fig. 19. Performance comparison with different values of n .

100%. In Fig. 19, we analyze the deviation of $l_{i,j}^{cell}$, since the errors caused by approximating constraint (6b) lead to a deviation for calculating $l_{i,j}^{cell}$ in (24b). The setting of parameter x affects the precision of the train speed trajectories. Fig. 20 analyzes the deviation of $\varpi_{f,i,j,k}$ in constraints (6a).

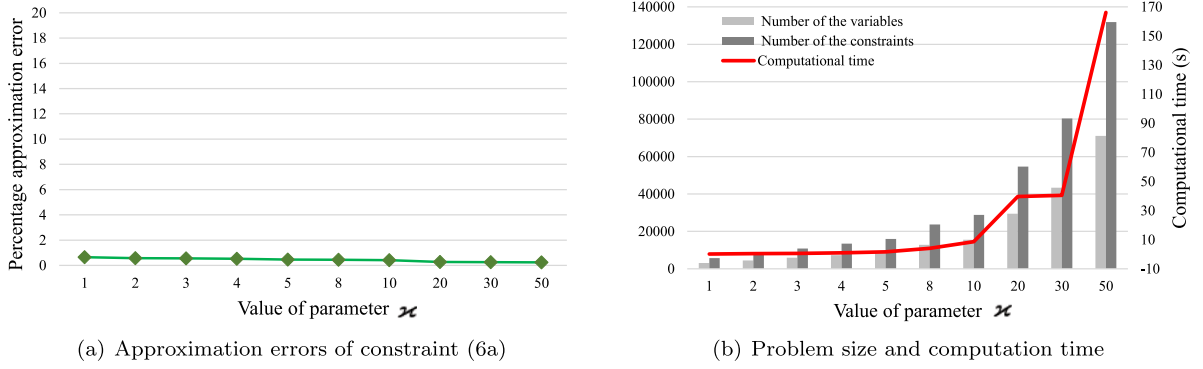
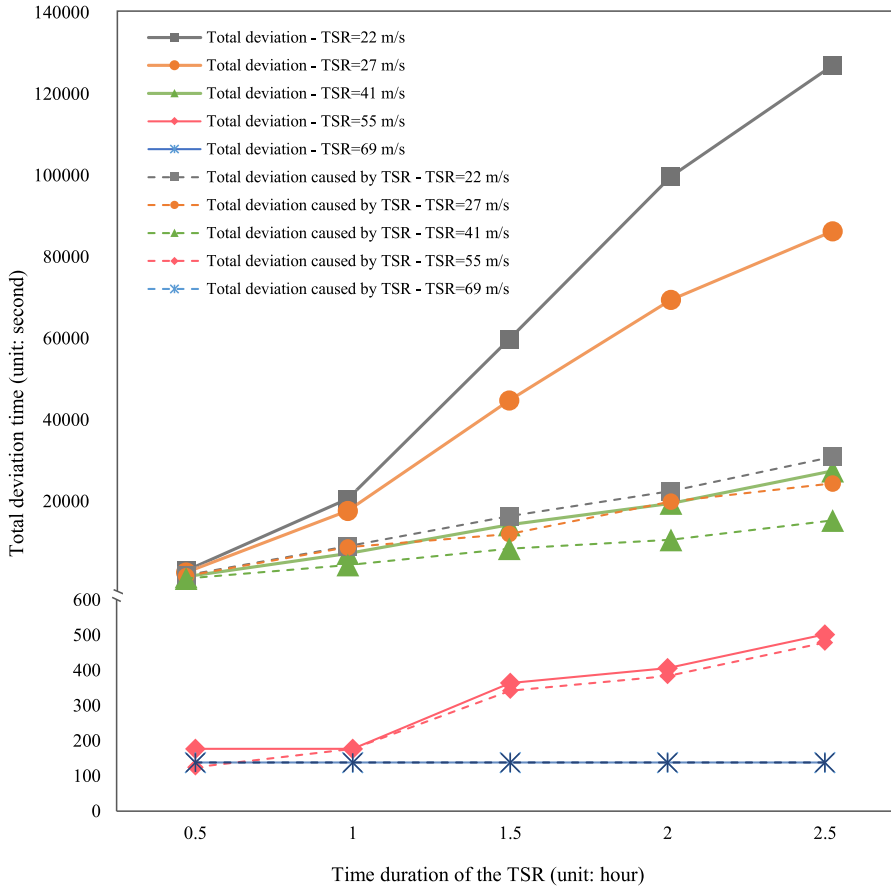
Fig. 20. Performance comparison with different values of κ .

Fig. 21. Total deviation time caused by a different set of speed limits and time duration of the Temporary Speed Restriction (TSR).

As illustrated in Fig. 18, the approximation errors tend to decrease when the parameter σ increases from 2 to 50, and the approximation errors become smaller than 1% when the parameter σ is larger than 2. Fig. 19 shows the results of parameter n . We can see a sharp reduction (from 18% to 0.16%) when the parameter n increases from 2 to 3. When we further enlarge the value of n , i.e., from 3 to 50, the change of the approximation errors is not significant, while the computation time largely increases due to the increase of the problem scale (in terms of constraints and variables). When the parameter n equals 50, the approximation error is small than 10^{-5} (i.e., $\frac{0.25}{63594}$). The approximation error of (2) and (6c) is relatively small, as shown in Fig. 20, ranging from 0.24% to 0.65%. The approximation errors slightly decrease when increasing the value of κ from 1 to 50; however, the computational time increases from 0.16 s to 166.16 s.

Appendix D. Analysis of the influence for parameters of the TSR

In this section, we use the same input parameters of the TSR in Section 6.1.2. Fig. 21 shows the experimental results for the different time durations of the TSR (i.e., represented on the X-axis), and speed limits (i.e., indicated by the curves in different colors, e.g., the gray line indicates the speed limit of 22.22 m/s). In addition, Fig. 21 reports the deviation time caused by TSR using the dotted lines with marked points. The Y-axis represents the deviation time.

From the experimental results, we can see that, under the same time duration of the TSR, the total deviation time and the deviation time caused by TSR evidently increase by decreasing the TSR speed limits (which becomes more severe, e.g., from 69.44 m/s to 22.22 m/s). Moreover, it is evident that the deviation time increases with the increase of the TSR time duration (e.g., from 0.5 h to 2.5 h). This fact shows that, the TSR with higher severity has a more serious impact on the train operation (longer deviation time).

In addition, comparing the two sets of the deviation time (i.e., total deviation times and the deviation caused by TSR), we can see that, the value of the deviation time caused by TSR is typically lower than the total deviation times. It is clear that the deviation time caused by TSR is much lower than the deviation time for the cases of the TSR with higher severity e.g., limited speeds are 22.22 m/s and 27 m/s, and the time duration is longer than 1.5 h. This indicates that, the TSR with higher severity increase the conflicts among trains, and further increase the deviation time of the timetable. Also, we note that, the difference between two kinds of deviation times is small, when the severity of the TSR is low, e.g., limited speeds of 55.56 m/s and 69.44 m/s.

References

- Albrecht, T., 2009. The influence of anticipating train driving on the dispatching process in railway conflict situations. *Netw. Spat. Econ.* 9 (1), 85–101.
- Cacchiani, V., Huisman, D., Kidd, M., Kroon, L., Toth, P., Veelenturf, L., Wagenaar, J., 2014. An overview of recovery models and algorithms for real-time railway rescheduling. *Transp. Res. B* 63, 15–37.
- Caimi, G., Fuchsberger, M., Laumanns, M., Lüthi, M., 2012. A model predictive control approach for discrete-time rescheduling in complex central railway station areas. *Comput. Oper. Res.* 39 (11), 2578–2593.
- Corman, F., D'Ariano, A., Pacciarelli, D., Pranzo, M., 2009. Evaluation of green wave policy in real-time railway traffic management. *Transp. Res. C* 17 (6), 607–616.
- Corman, F., D'Ariano, A., Pacciarelli, D., Pranzo, M., 2010. A tabu search algorithm for rerouting trains during rail operations. *Transp. Res. B* 44 (1), 175–192.
- Corman, F., Meng, L., 2014. A review of online dynamic models and algorithms for railway traffic management. *IEEE Trans. Intell. Transp. Syst.* 16 (3), 1274–1284.
- D'Ariano, A., Albrecht, T., Allan, J., Brebbia, C., Rumsey, A., 2010. Running time re-optimization during real-time timetable perturbations. In: *Timetable Planning and Information Quality*, Vol. 1. WIT Press, pp. 147–156.
- D'Ariano, A., Corman, F., Pacciarelli, D., Pranzo, M., 2008. Reordering and local rerouting strategies to manage train traffic in real time. *Transp. Sci.* 42 (4), 405–419.
- D'Ariano, A., Meng, L., Centulio, G., Corman, F., 2019. Integrated stochastic optimization approaches for tactical scheduling of trains and railway infrastructure maintenance. *Comput. Ind. Eng.* 127, 1315–1335.
- Dong, H., Liu, X., Zhou, M., Zheng, W., Xun, J., Gao, S., Song, H., Li, Y., Wang, F.-Y., 2021. Integration of train control and online rescheduling for high-speed railways in case of emergencies. *IEEE Trans. Comput. Soc. Syst.*
- Fang, W., Yang, S., Yao, X., 2015. A survey on problem models and solution approaches to rescheduling in railway networks. *IEEE Trans. Intell. Transp. Syst.* 16 (6), 2997–3016.
- Franke, R., Meyer, M., Peter, T., 2002. Optimal control of the driving of trains. *Automatisierungstechnik* 50, 606.
- Goverde, R.M., Scheepmaker, G.M., Wang, P., 2020. Pseudospectral optimal train control. *European J. Oper. Res.*
- Hansen, I.A., 2008. *Railway Timetable & Traffic: Analysis, Modelling, Simulation*. Eurail Press.
- Hansen, I.A., 2010. Railway network timetabling and dynamic traffic management. *Int. J. Civ. Eng.* 8 (1), 19–32.
- Hansen, I.A., Meng, L., 2019. Closing the loop between data mining and fast decision support for intelligent train scheduling and traffic control. *J. Beijing Jiaotong Univ.* 43 (1), 23–35.
- Hansen, I.A., Pachl, J., 2014. *Railway Timetabling & Operations: Analysis, Modelling, Optimisation, Simulation, Performance Evaluation*. Eurailpress, Hamburg.
- Hong, X., Meng, L., D'Ariano, A., Veelenturf, L.P., Long, S., Corman, F., 2021. Integrated optimization of capacitated train rescheduling and passenger reassignment under disruptions. *Transp. Res. C* 125, 103025.
- Howlett, P., 2000. The optimal control of a train. *Ann. Oper. Res.* 98 (1–4), 65–87.
- Howlett, P., Pudney, P., 1995. *Energy-Efficient Train Control*. Springer-Verlag, London.
- Ko, H., Koseki, T., Miyatake, M., 2004. Application of dynamic programming to the optimization of the running profile of a train. *WIT Trans. Built Environ.* 74.
- van der Kooij, R.B., Landmark, A.D., Seim, A.A., Olsson, N.O., 2017. The effect of temporary speed restrictions, analyzed by using real train traffic data. *Transp. Res. Procedia* 22, 580–587.
- Lai, Q., Liu, J., Haghani, A., Meng, L., Wang, Y., 2020. Energy-efficient speed profile optimization for medium-speed maglev trains. *Transp. Res. E* 141, 102007.
- Li, X., Lo, H.K., 2014a. An energy-efficient scheduling and speed control approach for metro rail operations. *Transp. Res. B* 64, 73–89.
- Li, X., Lo, H.K., 2014b. Energy minimization in dynamic train scheduling and control for metro rail operations. *Transp. Res. B* 70, 269–284.
- Lu, S., Hillmansen, S., Ho, T.K., Roberts, C., 2013. Single-train trajectory optimization. *IEEE Trans. Intell. Transp. Syst.* 14 (2), 743–750.
- Lu, S., Wang, M.Q., Weston, P., Chen, S., Yang, J., 2016. Partial train speed trajectory optimization using mixed-integer linear programming. *IEEE Trans. Intell. Transp. Syst.* 17 (10), 2911–2920.
- Luan, X., Miao, J., Meng, L., Corman, F., Lodewijks, G., 2017. Integrated optimization on train scheduling and preventive maintenance time slots planning. *Transp. Res. C* 80, 329–359.
- Luan, X., Wang, Y., De Schutter, B., Meng, L., Lodewijks, G., Corman, F., 2018a. Integration of real-time traffic management and train control for rail networks-Part 1: Optimization problems and solution approaches. *Transp. Res. B* 115, 41–71.
- Luan, X., Wang, Y., De Schutter, B., Meng, L., Lodewijks, G., Corman, F., 2018b. Integration of real-time traffic management and train control for rail networks-Part 2: Extensions towards energy-efficient train operations. *Transp. Res. B* 115, 72–94.
- Lusby, R.M., Larsen, J., Ehrgott, M., Ryan, D.M., 2013. A set packing inspired method for real-time junction train routing. *Comput. Oper. Res.* 40 (3), 713–724.
- Lüthi, M., 2009. *Improving the Efficiency of Heavily Used Railway Networks through Integrated Real-Time Rescheduling*, Vol. 147. ETH Zurich.
- Mazzarello, M., Ottaviani, E., 2007. A traffic management system for real-time traffic optimisation in railways. *Transp. Res. B* 41 (2), 246–274.

- Meng, L., Zhou, X., 2014. Simultaneous train rerouting and rescheduling on an N-track network: A model reformulation with network-based cumulative flow variables. *Transp. Res. B* 67, 208–234.
- Meng, L., Zhou, X., 2019. An integrated train service plan optimization model with variable demand: A team-based scheduling approach with dual cost information in a layered network. *Transp. Res. B* 125, 1–28.
- Mo, P., D'Ariano, A., Yang, L., Veelenturf, L.P., Gao, Z., 2021. An exact method for the integrated optimization of subway lines operation strategies with asymmetric passenger demand and operating costs. *Transp. Res. B* 149, 283–321.
- Mo, P., Yang, L., Gao, Z., 2019. Energy-efficient train operation strategy with speed profiles selection for an urban metro line. *Transp. Res. Rec. J. Transp. Res. Board* 2673 (4), 348–360.
- Ning, B., Dong, H., Zheng, W., Xun, J., Gao, S., Wang, H., Lingyun, M., Yidong, L., 2019. Integration of train control and online rescheduling for high-speed railways: challenges and future. *Acta Automat. Sinica* 45 (12), 2208–2217.
- Pachl, J., 2009. *Railway Operation and Control*, second ed. VTD Rail Publishing, Mountlake Terrace.
- Pellegrini, P., Marlière, G., Rodriguez, J., 2014. Optimal train routing and scheduling for managing traffic perturbations in complex junctions. *Transp. Res. B* 59, 58–80.
- Schöbel, A., 2017. An eigenmodel for iterative line planning, timetabling and vehicle scheduling in public transportation. *Transp. Res. C* 74, 348–365.
- Tan, Z., Lu, S., Xue, F., Bao, K., 2017. A speed trajectory optimization model for rail vehicles using mixed integer linear programming. In: *Proceeding of 2017 IEEE 20th International Conference on Intelligent Transportation Systems. ITSC, IEEE*, pp. 1–6.
- Wang, Y., 2014. *Optimal Trajectory Planning and Train Scheduling for Railway Systems* (Ph.D. thesis). Delft University of Technology.
- Wang, Y., De Schutter, B., van den Boom, T.J., Ning, B., 2013. Optimal trajectory planning for trains—A pseudospectral method and a mixed integer linear programming approach. *Transp. Res. C* 29, 97–114.
- Wang, P., Goverde, R.M., 2016. Multiple-phase train trajectory optimization with signalling and operational constraints. *Transp. Res. C* 69, 255–275.
- Wang, Y., Zhu, S., D'Ariano, A., Yin, J., Miao, J., Meng, L., 2021. Energy-efficient timetabling and rolling stock circulation planning based on automatic train operation levels for metro lines. *Transp. Res. C* 129, 103209.
- Williams, H.P., 2013. *Model Building in Mathematical Programming*. John Wiley & Sons.
- Wu, C., Zhang, W., Lu, S., Tan, Z., Xue, F., Yang, J., 2018. Train speed trajectory optimization with on-board energy storage device. *IEEE Trans. Intell. Transp. Syst.* 20 (11), 4092–4102.
- Xu, P., Corman, F., Peng, Q., 2016. Analyzing railway disruptions and their impact on delayed traffic in Chinese high-speed railway. *IFAC-PapersOnLine* 49 (3), 84–89.
- Xu, P., Corman, F., Peng, Q., Luan, X., 2017. A train rescheduling model integrating speed management during disruptions of high-speed traffic under a quasi-moving block system. *Transp. Res. B* 104, 638–666.
- Ye, H., Liu, R., 2017. Nonlinear programming methods based on closed-form expressions for optimal train control. *Transp. Res. C* 82, 102–123.
- Zhan, S., Wang, P., Wong, S., Lo, S., 2022. Energy-efficient high-speed train rescheduling during a major disruption. *Transp. Res. E* 157, 102492.
- Zhang, Y., D'Ariano, A., He, B., Peng, Q., 2019. Microscopic optimization model and algorithm for integrating train timetabling and track maintenance task scheduling. *Transp. Res. B* 127, 237–278.
- Zhou, M., Liu, X., Hou, Z., Shang, J., Yue, Y., Song, H., 2020. Integrated optimization of dispatching decision and speed trajectory for high-speed railway under disturbances. In: *2020 Chinese Automation Congress. CAC, IEEE*, pp. 4656–4661.
- Zhu, Y., Lv, X., Xu, Y., 2015. *Traction Calculation and Simulation System for EMUs* (in Chinese). China Railway Publishing House.