

Li WANG, Lixing YANG, Ziyou GAO, Yeran HUANG

# Robust train speed trajectory optimization: A stochastic constrained shortest path approach

© The Author(s) 2017. Published by Higher Education Press. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0>)

**Abstract** Train speed trajectory optimization is a significant issue in railway traffic systems, and it plays a key role in determining energy consumption and travel time of trains. Due to the complexity of real-world operational environments, a variety of factors can lead to the uncertainty in energy-consumption. To appropriately characterize the uncertainties and generate a robust speed trajectory, this study specifically proposes distance-speed networks over the inter-station and treats the uncertainty with respect to energy consumption as discrete sample-based random variables with correlation. The problem of interest is formulated as a stochastic constrained shortest path problem with travel time threshold constraints in which the expected total energy consumption is treated as the evaluation index. To generate an approximate optimal solution, a Lagrangian relaxation algorithm combined with dynamic programming algorithm is proposed to solve the optimal solutions. Numerical examples are implemented and analyzed to demonstrate the performance of proposed approaches.

**Keywords** train speed trajectory optimization, railway operation, stochastic programming

## 1 Introduction

A train speed trajectory, which is regarded as an operational guidance to drivers and an integral part of an

automatic train operation (ATO) system, specifies the movement of a single train over an inter-station or a railway line. Trains controlled on the railway line according to speed trajectory guarantee well-organized traffic. Focusing on the operational requirement of automation and efficiency, target speed trajectory is usually required to pre-generate real-time tracking of ATO controller as guidance during the train operation with the aim of preventing collision of the automatic train protection (ATP) system. Figure 1 illustrates the train operation process in real-world applications. To optimize speed trajectories, some indicators such as punctuality, riding comfort, energy efficiency, and switching frequency between various driving modes are considered; at the same time, considering track condition, stopping point, maximal allowable speed of ATP system, and vehicle/carriage data (which includes train mass and traction/braking performance). Guided by pre-generated target speed trajectories, the ATO controller or operator then releases the control instructions to drive the train on the transit line.

Owing to its inherent significance, the train speed trajectory optimization problem has attracted significant attention from researchers and engineers in recent decades. One of the popular topics is how to generate an energy-efficient speed trajectory to further reduce operation cost. With the assumption that trains can completely track the target speed trajectory, three methodologies have been proposed to optimize the speed curve to reduce operational energy consumption; these methodologies consist of analytical methods, numerical algorithms, and intelligence algorithms. Based on the optimal control theory, analytical methods, which include maximum principle (e.g., Asnis et al., 1985; Howlett, 2000; Liu and Golovitcher, 2003) and dynamic programming (e.g., Ko et al., 2004; Calderaro et al., 2014), require good properties of the considered objectives and aim to obtain optimal solutions even for complex problems. Some measures have suggested to improve properties of the objective function in speed curve optimization, such as simplification of certain conditions in

Received May 23, 2017; accepted September 2, 2017

Li WANG  
School of Modern Post, Beijing University of Posts and Telecommunications, Beijing 100876, China

Lixing YANG (✉), Ziyou GAO, Yeran HUANG  
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China  
E-mail: lxyang@bjtu.edu.cn

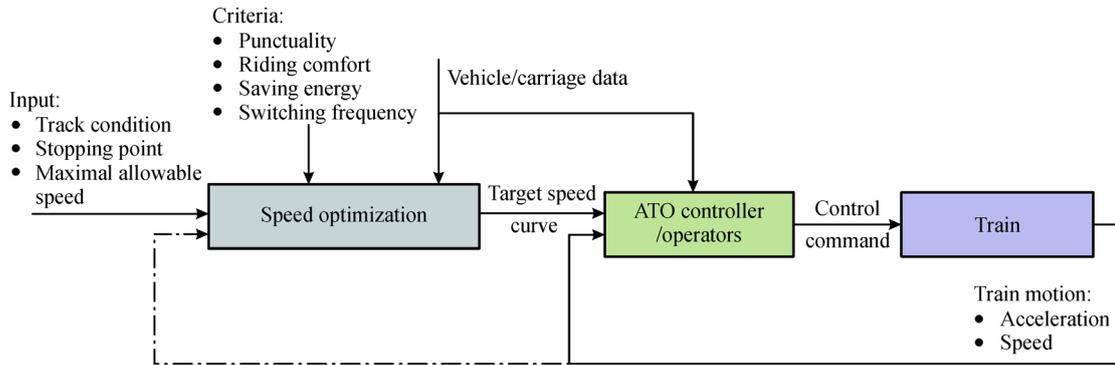


Fig. 1 Illustration of a driving system

the modeling process. Along this line, Howlett (2000) used Pontryagin principle to determine an optimal driving strategy of continuous and direct controls for fuel consumption minimization within a given running time. Liu and Golovitcher (2003) described an analytical process with maximum principle to compute the optimal operating successions of a rail vehicle to minimize energy consumption. Khmel'nitsky (2000) determined a detailed program for traction and brake applications to minimize energy consumption with a variable grade profile subject to arbitrary speed restrictions, where the optimal solution was obtained from the maximum principle analysis. Ko et al. (2004) investigated Bellman's dynamic programming to optimize train running speed with the objective of minimizing total consumed energy. Calderaro et al. (2014) used dynamic programming to find a set of pseudo-optimal speed cycles to minimize the electrical energy used for traction operations.

Compared with analytical methods, numerical algorithms can make a trade-off between accuracy and computational time. These algorithms, which have a less strict objective function, include gradient method and sequential quadratic programming. For instance, Miyatake and Matsuda (2009) pointed out that the charging/discharging command of on-board energy storage and vehicle speed profile should be optimized together based on optimality analysis; they developed a mathematical model based on sequential quadratic programming. Miyatake and Ko (2010) introduced dynamic programming, gradient method, and sequential quadratic programming to find energy-saving train speed profiles and compared the performance of the three methods. In addition, intelligent algorithms were proposed by simulating natural processes such as ant feeding, natural evolutionary law, and human intelligence process to find the near-optimal speed trajectories to achieve the minimum/maximum objective function during optimization. These algorithms include ant colony algorithm, genetic algorithm, tabu search, and simulated annealing algorithm. Ke et al. (2009) considered the speed curve optimization of

trains to reduce energy consumption on different signal block modes, that is, fixed block signaling and moving block signaling (Ke et al., 2012), which were solved by max-min ant algorithm. Kim and Chien (2011) developed simulated annealing algorithm to search for the optimal train operation for energy consumption minimization by considering track alignment, speed limit, and schedule adherence. Considering usage of regenerative energy between two trains, Tang et al. (2015) optimized the speed profile of one of the trains to reduce power consumption by enhancing genetic algorithm, which was calculated by the established electric model.

In practice, incomplete accurate tracking performance of ATO controller occurs. Thus, some studies have coordinated the optimized target speed trajectory with ATO controller performance to search for punctual and energy-saving strategies. For instance, Liu et al. (2015) proposed a more accurate model of energy calculation with ATO strategies to optimize the energy-saving speed curve by modifying the tabu search algorithm. Cao et al. (2016) proposed an optimization method of the recommended speed profile by considering the ATO tracking strategy, which was solved by the modified max-min ant algorithm. In addition, Wang et al. (2013) made a trade-off between energy consumption and riding comfort to search for the optimal trajectory through a pseudo-spectral method with varying line resistance, speed restrictions, and maximum traction force. Dominguez et al. (2012) designed optimal ATO speed profiles of metro trains to minimize the net energy at substations using the modular simulator considering the energy recovered from regenerative brake and stored in an on-board energy storage device.

Majority of the existing studies mainly focused on deterministic environments in which all the involved parameters are set as deterministic quantities. In the process of finding energy-efficient speed trajectories, various factors can lead to unreliable energy-consumption performance due to the complexity of the operation environment. For instance, in the Beijing metro system, the passenger demand during peak hours is typically much

larger than that during off-peak hours. Thus, even with the same speed trajectory, the energy consumption fluctuates between different service periods. In addition, the difference in the performance of individual trains exist, which leads to the uncertainty in energy consumption during actual operations. Recognizing these significant characteristics, the present study intends to investigate how a robust train speed trajectory over an inter-station is generated by considering the randomness/fluctuation of the energy consumption in the operation process. This is a novel idea because no related research has been found until today.

In detail, this study aims to contribute the following to the field of speed trajectory optimization: (1) represent the speed-distance space into a specified network by discretizing the distance into different intervals, and construct the velocity shift link between different velocity stamps by considering all possible acceleration rates; (2) use a sample-based link energy consumption to capture the uncertainty in the operation process; (3) formulate a problem as stochastic constrained shortest path to generate energy-efficient speed trajectories; and (4) design a combination of Lagrangian relaxation and dynamic programming algorithms to search for optimal solutions to the proposed model.

The rest of the paper is organized as follows. In Section 2, we provide a detailed statement of the problem of interest. In Section 3, we construct the speed-distance network for all potential speed trajectories. Section 4 formulates this problem as a stochastic constrained shortest path problem to generate the robust speed trajectory for the practical operations. A heuristic algorithm, which is a combination of Lagrangian relaxation and dynamic programming algorithms, is proposed to search for the near-optimal solution to the proposed model in Section 5. Numerical examples are implemented to demonstrate the performance of the proposed approaches in Section 6. Finally, we draw a conclusion.

## 2 Problem statements

In the process of determining optimal speed trajectory for trains, two evaluation indexes are considered when a train traverses along the railway line with a given speed trajectory; these are travel time and energy consumption. The travel time index corresponds to service quality in the transportation process because the railway system is widely regarded as one of the time-efficient transportation modes due to its punctuality and convenience. By contrast, the energy consumption index is essentially associated with the travel cost in the operation process. Generally, the variation tendency of inter-station travel time and energy consumption take opposite directions because high operation speed leads to increased energy usage and reduced inter-station travel time accordingly. To make a trade-off

between these two indexes, the existing literature always treats this problem as a multi-objective programming model and uses some traditional methods to handle two objectives, such as the linear weighted method and ideal point method. In these methods, the trajectory is optimized by considering the compromise between travel time and energy consumption.

In real-world applications, trains are required to operate in accordance with predetermined timetable in which the link travel time is set in the planning stage. Each train is scheduled to travel on the railway line according to given arrival and departure times at each station. For energy-saving operations, finding energy-efficient speed trajectories is also needed to further reduce travel costs. Thus, a speed trajectory optimization problem within the pre-specified travel time is at hand, which is also a significant issue for practical operations.

Energy consumption with respect to train operation mainly consists of two parts: one is traction and braking energy consumption, and the other is auxiliary energy consumed by auxiliary facilities such as lighting, air conditioning, and ventilation systems. However, since the auxiliary energy consumption is unrelated to the objective of this study, we only emphasize decreasing the traction and braking energy consumption in the speed trajectory optimization process. For practical operations, we have to specifically consider the different control modes in the train traversing process, including traction, cruising, coasting, and braking operations. In general, traction and cruising operations aim to keep or enhance the train speeds during operations. The coasting operation intends to decrease the speed through the resistance force (including air resistance, wheel resistance, grade resistance, and frictional resistance), and the braking operation is employed to finally stop the trains at stations. For these control modes, the coasting operation does not consume the traction energy, and the energy consumption occurs for the other three modes. With these analyses, the problem considered in this study is essentially to determine the sequence of operation statuses and their transfer sites along the railway lines, with the purpose of minimizing the energy consumption under the guaranteed travel time.

To clearly characterize train speed trajectories, we provide the illustration in Fig. 2 to show the detailed operations in the movement process of a train.

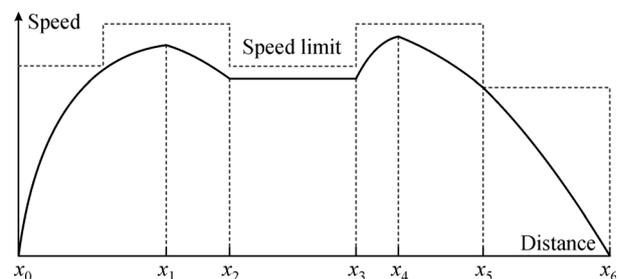


Fig. 2 Illustration of speed trajectory over an inter-station

Figure 2 illustrates the speed-distance trajectory over a given railway section, where the sites  $x_0$  and  $x_6$  on the  $x$ -coordinate correspond to the origin and destination stations, respectively, and any site along this coordinate specifically maps to a location on this inter-station. The  $y$ -coordinate represents the dimension of velocity. Typically, this speed trajectory consists of four types of operations. When a train leaves the site  $x_0$ , it needs to accelerate to site  $x_1$ . This interval corresponds to traction operation. At site  $x_1$ , the train is required to change from acceleration operation to coasting operation up to site  $x_2$ . There after, the cruising operation is employed to keep the constant speed over interval  $[x_2, x_3]$ . Then, the acceleration operation is reused to enhance the speed of the train in interval  $[x_3, x_4]$ . In interval  $[x_4, x_5]$ , the coasting operation is used again until site  $x_5$ . Finally, the braking operation is implemented in the last section to stop the train at the end of the section. In this process, the generated speed trajectory cannot exceed the pre-specified speed limit.

This illustration suggests that finding an optimal speed-distance trajectory is essentially specifying the sequence of traverse modes and the mode shift points along the inter-station. Generally, the pre-specified speed trajectory corresponds to two features, that is, travel time and energy consumption. To reduce energy consumption, choosing the optimal speed-distance trajectory with the guaranteed travel time requirement is needed. Note that this problem is essentially a complex nonlinear decision-making process that is difficult to solve effectively through analytical methods. Therefore, a variety of heuristics with various searching strategies, such as genetic algorithm and local search, have been used in the literature to find the approximate optimal solutions.

### 3 Speed-distance network

In the literature, trajectory optimization is always investigated in a certain environment, that is, all the involved parameters in the problem are assumed to be deterministic. However, various factors can lead to uncertainties in energy consumption, such as passenger loading amount and train performance. Therefore, in the process of optimizing train trajectories, considering uncertain factors is important in generating robust train trajectories. This study aims to investigate the method for producing robust train trajectories through a stochastic constrained shortest path model. To ensure comprehensiveness of this paper, we introduce the approaches to deal with the problem.

To employ the shortest path model, we first introduce the discretization method for this trajectory optimization problem, in which the entire speed-distance space is discretized into different stages, over which all the possible speed variations can be represented as a variety of links between different speed stamps. By using this method, we

finally construct a speed-distance variation network between origin and destination stations. As a result, the dynamic programming algorithm can be used to solve the optimal train trajectory. Figure 3 illustrates this discretization process.

As shown in Fig. 3, a speed-distance network is constructed over a section between two adjacent stations in which the speed and distance are all discretized in the 2D coordinate of distance and speed. In practice, this network can be regarded as an approximation of the original train trajectory space (that is, the area below the speed limit line over the inter-station). Specifically, the distance can be decomposed into a series of discrete sites by using a small distance interval with length  $\delta$  (e.g.,  $\delta = 1$  m). Over each discrete site, the potential velocity limit is also discretized into different speed stamps. Through this method, the entire feasible region of the speed trajectory can be discretized into various representative speed-distance nodes, over which the speed-distance network has to be constructed.

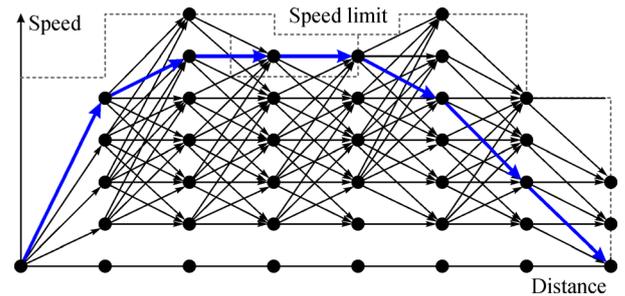


Fig. 3 Illustration of speed-distance network over an inter-station

To prepare the network, we have to consider the tradeoff between representation accuracy and complexity. Theoretically, if a much larger interval is used to discretize the speed-distance region, the accuracy of the generated optimal speed trajectory cannot be fully guaranteed, while the computational complexity can be decreased to significantly. On the contrary, if a small discretized interval is adopted, accuracy can be improved and complexity can be increased. In the speed-distance network, each speed trajectory can be represented by a path from the origin to the destination. Each link in this network is typically associated with two performance indicators, namely, link travel time and link energy consumption. The link travel time can be guaranteed if we take a suitable acceleration rate in practical operations. However, the link energy consumption may vary due to different train performances, driver habits, and loading number of passengers. In practice, we can use data to capture the randomness of the energy consumption over various links. The detailed process of calculating the link travel time and energy consumption is given in the following.

In the data preparation, acceleration and deceleration

rates can also be discretized into  $K$  cases denoted by  $a_1, a_2, \dots, a_K$ . Determining the links in the speed-distance network is easier with these acceleration rates. For instance, for each discretized speed-distance node  $A$  with speed  $v_1$ , we can adopt various acceleration rates to deduce all the potential speed in the next stage according to Newton's laws of motion. Typically, if acceleration rate  $a_k$  is adopted, then we can obtain the speed (denoted by  $v_2$ ) at the next stage. According to the equation

$$\delta = v_1 t + \frac{1}{2} a_k t^2, \quad (1)$$

we can deduce the travel time with each acceleration rate  $a_k$  as

$$t = \frac{-v_1 + \sqrt{v_1^2 + 2a_k \delta}}{a_k} \quad (a_k \geq 0),$$

$$t = \frac{-v_1 - \sqrt{v_1^2 + 2a_k \delta}}{a_k} \quad (a_k < 0). \quad (2)$$

Then, according to Newton's laws of motion, we obtain the following speed at the next stage:

$$v_2 = v_1 + a_k t. \quad (3)$$

In the process of data preparation, if we enumerate all the possible acceleration rates, the potential links rooted at node  $A$  can be deduced. Through this method, we can finally construct the speed-distance network to find energy-efficient speed trajectories in which the link travel time can be predetermined by the method.

As stated, even in each link, a variety of complex factors can lead to the randomness of energy consumption. To capture the randomness, this study uses the following method to produce random data. First, we provide a constant quality of the involved trains, denoted by  $m$ . Then, we calculate the total energy consumption  $\bar{E}_{ij}$  on a specific link  $(i, j)$  with travel time  $T_{ij}$  and acceleration  $a_k$  by simulation. That is,

$$\bar{E}_{ij} = \theta \int_{t \in T_{ij}} F v(t) dt. \quad (4)$$

In the preceding equation, parameter  $\theta$  is a transfer coefficient between the mechanical energy and electricity energy.  $F = ma_k + F_b$ , if  $a_k \geq 0$ ;  $F = |ma_k| - F_b$ , if  $a_k < 0$ , where  $F_b$  is the resistance referring to Ghoseiriet et al. (2004), and  $F_b = mg(A_1 v(t)^2 + A_2 v(t) + A_3)$ . With this information, we then treat the random link energy consumption (denoted by  $E_{ij}$ ) as

$$E_{ij} = \bar{E}_{ij} + \xi, \quad (5)$$

in which  $\xi$  is a random variable to show the energy-consumption fluctuation incurred by the uncertain factors.

In real-world operations, time efficiency is an important

index to evaluate the service quality of train operations. For the speed trajectory optimization, since the travel time in an inter-station is restricted within a time threshold to guarantee service quality, this problem can be finally formulated as a constrained shortest path problem in the proposed distance-speed trajectory network, which is discussed in detail in the following section.

## 4 Mathematical formulation

In this section, we formulate the problem as a stochastic programming model in which the randomness of energy consumption is specifically considered. The core issue in the formulation process is to determine the representation method for the random link travel time. Theoretically, two discretization methods can be employed to characterize randomness. One method is by using the independent discrete random variables associated with different links to capture the randomness of energy consumption, and the other is by using the sample-based discrete random variable with specific correlation over the entire speed-distance network. In the present study, we have adopted the second method to characterize the randomness of energy consumption. In practice, this type of representation is reasonable because each service train can provide sample data for the link energy consumption in real-world applications. Thus, the focus of this study is to generate a robust speed trajectory in the speed-distance network with the given sample-based random data, such that the expected energy consumption is minimized. To formulate the problem of interest, we first introduce the following notations and parameters:

- $N$ : set of nodes in speed-distance network,
- $A$ : set of arcs in speed-distance network,
- $(N, A)$ : input speed-distance network,
- $i, j$ : indexes of nodes in the network,
- $(i, j)$ : directed link from node  $i$  to  $j$ ,
- $t_{ij}$ : link travel time on link  $(i, j)$ ,
- $e_{ij}^\omega$ : link energy consumption over sample  $\omega$ ,
- $\omega$ : index of sample  $\omega$ ,
- $p_\omega$ : probability of sample  $\omega$ ,
- $\Omega$ : set of sample.

To find an optimal speed trajectory with the least expected energy consumption, this problem turns out to be a path-finding problem in the speed-distance network. Thus, we aim to formulate this problem as a stochastic constrained shortest path problem by using an approach similar to that proposed by Wang et al. (2016). To formulate this optimization problem, we introduce the decision variables involved in the problem of interest.

$y_{ij}^\omega$  is the binary indicator of selecting link  $(i, j)$  over sample  $\omega$ . If link  $(i, j)$  over sample  $\omega$  is selected, then  $y_{ij}^\omega = 1$ ; otherwise,  $y_{ij}^\omega = 0$ .

Using these decision variables, we have to introduce the

system constraints in the speed trajectory searching process. As shown in the network, this problem essentially needs to find a path from the origin to the destination over different samples. Therefore, the flow balance constraint is formulated as follows for each sample:

$$\sum_{(i,j) \in A} y_{ij}^\omega - \sum_{(j,i) \in A} y_{ji}^\omega = \begin{cases} 1, & \text{if } i = O, \\ -1, & \text{if } i = D, \omega \in \Omega, \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

In this formulation, the objective is to provide a unique speed trajectory as a guide for all the trains in real-world applications. As the flow balance constraint can only provide a sample-based speed trajectory for each sample, all the generated trajectories must not overlap each other because such a condition would cause difficulty. To generate an operational trajectory for all the trains under various operation conditions, this study intends to provide a robust solution for practical operations. For this purpose, we introduce the following unique speed trajectory constraint:

$$y_{ij}^{\omega'} = y_{ij}^{\omega''}, \omega', \omega'' \in \Omega. \quad (7)$$

Note that constraint (7) can be further reformulated as a total of  $|\Omega|$  equality constraints:

$$y_{ij}^1 = y_{ij}^2, y_{ij}^2 = y_{ij}^3, \dots, y_{ij}^{|\Omega|-1} = y_{ij}^{|\Omega|}, y_{ij}^{|\Omega|} = y_{ij}^1. \quad (8)$$

In addition, to transfer the passengers within the shortest possible time, the travel time should be guaranteed over each inter-station. For instance, on the level of schedule, we pre-specify the inter-station travel time in advance, denoted by  $T$ . Then, on the level of control, we have to ensure that regardless of the final speed profile produced, the link travel time should not exceed the time threshold for orderly operations according to the train schedule. Thus, to find the optimal solution, we have to add a time threshold side constraint to keep the travel time, as given in the following equation:

$$\sum_{(i,j) \in A} y_{ij}^\omega \cdot t_{ij} \leq T. \quad (9)$$

In this problem, different evaluation indexes can be used to evaluate the speed trajectory selection strategy. If the energy consumption is used as the evaluation index, the sample-based total energy consumption can be calculated as  $E[Y_\omega] = \sum_{(i,j) \in A} e_{ij}^\omega$ ,  $\omega \in \Omega$ . The expected constrained

shortest path problem can then be formulated as follows by considering the probability of each sample:

$$E[Y] = \sum_{\omega \in \Omega} \sum_{(i,j) \in A} p_\omega \cdot e_{ij}^\omega \cdot y_{ij}^\omega. \quad (10)$$

The train-speed trajectory problem can be formulated as the following stochastic constrained shortest path problem:

$$\begin{cases} \min E[Y] \\ \text{s.t.} \\ \text{flow balance constraints} \\ \text{unique trajectory constraints} \\ \text{side constraints} \\ \text{binary variable constraints} \end{cases} \quad (11)$$

In this model, the purpose is to find a robust speed trajectory with the least expected energy consumption. The first constraint aims to generate a series of sample-based speed trajectories, the second constraint imposes the unique trajectory constraint for the final results, and the third constraint requires that the inter-station travel time cannot exceed the pre-specified time threshold.

## 5 Solution algorithm

In Wang et al. (2016), since the Lagrangian relaxation-based algorithm demonstrates its effectiveness in solving the stochastic constrained shortest path problem, this study also proposes a similar algorithm to solve the model, which is detailed in the following discussion.

First, we reformulate and relax the hard constraints, namely, unique trajectory constraints and side constraint, by using Lagrangian relaxation approach. Specifically, the Lagrangian multipliers associated with unique speed trajectory constraints and siding constraints, denoted by  $\lambda_{ij}^\omega \geq 0, \forall (i,j) \in A, \omega \in \Omega$ , and  $\mu^\omega \geq 0, \omega \in \Omega$ , respectively, are introduced to construct the following relaxed model:

$$\begin{cases} R(\lambda, \mu) = \min \sum_{\omega \in \Omega} \sum_{(i,j) \in A} p_\omega \cdot e_{ij}^\omega \cdot y_{ij}^\omega \\ \quad + \sum_{(i,j) \in A} \left( \sum_{\omega \in \Omega} \lambda_{ij}^\omega (y_{ij}^\omega - y_{ij}^{\omega+1}) + \lambda_{ij}^{|\Omega|} (y_{ij}^{|\Omega|} - y_{ij}^1) \right) \\ \quad + \sum_{\omega \in \Omega} \mu^\omega \left( \sum_{(i,j) \in A} y_{ij}^\omega \cdot t_{ij} - T \right) \\ \text{s.t. flow balance constraints} \\ \text{binary variable constraints.} \end{cases} \quad (12)$$

By regrouping the variables of the objective function in the relaxed model (12), we can rewrite this model as follows:

$$\begin{cases} R(\lambda, \mu) = \min \sum_{\omega \in \Omega} \sum_{(i,j) \in A} (p_\omega \cdot e_{ij}^\omega + \mu^\omega \cdot t_{ij} + \gamma_{ij}^\omega) \cdot y_{ij}^\omega \\ \quad - \sum_{\omega \in \Omega} \mu^\omega \cdot T \\ \text{s.t. flow balance constraints} \\ \text{binary variable constraints} \end{cases} \quad (13)$$

where  $\gamma_{ij}^1 = \lambda_{ij}^1 - \lambda_{ij}^{|\Omega|}$ ,  $\gamma_{ij}^\omega = \lambda_{ij}^\omega - \lambda_{ij}^{\omega-1}$ ,  $\omega = 2, 3, \dots, |\Omega|$ .

With given multipliers  $\lambda_{ij}^\omega \geq 0$ ,  $\forall (i, j) \in A$ ,  $\omega \in \Omega$ , and  $\mu^\omega \geq 0$ ,  $\omega \in \Omega$ , the relaxed model (13) can be further decomposed into a total of  $|\Omega|$  standard shortest path problems with generalized link cost  $c_{ij}^s = p_\omega \cdot t_{ij}^s + \mu^\omega \cdot t_{ij} + \gamma_{ij}^\omega$  and a Lagrangian multiplier related constant. In detail, the decomposed standard shortest path problem over sample  $\omega$  can be formulated as follows.

$SP(\omega)$ :

$$\begin{cases} R_\omega(\lambda, \mu) = \min \sum_{(i,j) \in A} (p_\omega \cdot e_{ij}^\omega + \mu^\omega \cdot t_{ij} + \gamma_{ij}^\omega) \cdot y_{ij}^\omega \\ \text{s.t. flow balance constraints} \\ \text{binary variable constraints} \end{cases} \quad (14)$$

For the optimal objective functions between the relaxed model and subproblems, we can deduce the following relationship. In detail, for any given Lagrangian multiplier vector  $(\lambda, \mu)$ , we have

$$R(\lambda, \mu) = \sum_{\omega \in \Omega} R_\omega(\lambda, \mu) - \sum_{\omega \in \Omega} \mu^\omega \cdot T. \quad (15)$$

In addition, the Lagrangian dual problem can be formulated as follows:

$$L = \max_{\lambda, \mu \geq 0} R(\lambda, \mu).$$

In the following discussion, the subgradient and dynamic programming algorithms can be integrated to search for an approximate optimal solution to the original problem. The subgradient algorithm is used to generate the smallest gap between the lower and upper bounds, in which the best upper bound is treated as the approximate optimal objective function. Theoretically, a small duality gap corresponds to a high-quality solution. Once the lower and upper bounds overlap, the exact optimal solution is finally determined.

During the solution process, the subgradient algorithm can be treated as the out-loop algorithm framework to solve the Lagrangian dual problem, which aims to improve the lower bound of the original problem iteratively. At each iteration with the fixed Lagrangian multipliers, we have to solve the relaxed model to obtain its optimal objective. According to the decomposed models, solving the deterministic shortest path problems over different samples is necessary. Note that the speed-distance network is a single-direction network with directed links; thus, dynamic programming can be employed to find the shortest path problem in the given speed-distance network. In addition, we can generate a total of  $|\Omega|$  speed trajectories (that is, satisfying the siding constraint) when solving the sample-based shortest path problem. If the generated trajectories are feasible solutions to the original model, they will be used to update the upper bound of the searching process.

In the subgradient algorithm, the solution process is a

point-to-point search in which the subgradient direction is treated as the searching direction of the current point. Once the number of iterations reaches a predetermined threshold, or the duality gap is below a given level  $\varepsilon$ , the searching process is terminated, and the best upper bound is outputted as a close-to-optimal solution to the original problem. The detailed procedure of the designed algorithm is discussed in the following.

**Algorithm.** Lagrangian relaxation algorithm to solve model (11).

**Step 1: Initialization**

Set iteration number  $\eta = 1$ .

Initialize the Lagrangian multipliers  $\lambda_{ij}^\omega \geq 0$ ,  $\forall (i, j) \in A, \omega \in \Omega$  and  $\mu^\omega \geq 0$ ,  $\omega \in \Omega$ .

**Step 2: Solution to relaxed model**

Solve subproblems by the dynamic programming algorithm, and find a speed trajectory for each sample. We denote the objective value of the relaxed model at the current iteration as the lower bound.

**Step 3: Update of upper bound**

If the speed trajectory derived from a subproblem is not feasible to the original model (11), we adjust it as a feasible one through the  $K$ -shortest path algorithm. We set the objective value of each feasible solution to the original model as the current upper bound and update the upper bound at iteration  $\eta$  by the minimum. We calculate the relative gap between the upper and lower bounds at iteration  $\eta$ .

**Step 4: Update of Lagrangian multipliers**

We update the Lagrangian multipliers  $\lambda_{ij}^\omega$  and  $\mu^\omega$  by subgradients  $\lambda_{ij}^\omega + \theta_\eta (y_{ij}^\omega - y_{ij}^{\omega+1})$  and  $\mu^\omega + \theta_\eta \left( \sum_{(i,j) \in A} y_{ij}^\omega \cdot t_{ij} - T \right)$ , where  $\theta_\eta$  is the step size at iteration  $\eta$ .

**Step 5: Termination condition test**

If  $\eta$  is less than a predetermined maximum iteration number or the relative gap is less than a pre-specified value, we terminate the algorithm; otherwise,  $\eta = \eta + 1$ , and we return to Step 2.

---

## 6 Numerical examples

This section describes the series of numerical experiments implemented to test the effectiveness and performance of the proposed approaches. To solve the model in the examples, the designed heuristic algorithm is coded by C++ in Microsoft Visual Studio 2008 software and performed on a PC with 4.00 GB memory and 3.40 GHz processor.

In the experimental design, we considered a railway segment between two stations with a total length of 1000 m. To construct the speed-distance network, we

further subdivided this inter-station into 100 sites with the same length, that is, 10 m for each site and a total of 100 stages in the railway segment. Theoretically, the length of each site can be set as smaller if we want to describe the speed-distance more accurately, which will lead to an increase in computational time. According to the real-world operation conditions, the speed limit in distance intervals [0,200] and [800,1000] is  $55 \text{ km}\cdot\text{h}^{-1}$  (about  $15 \text{ m}\cdot\text{s}^{-1}$ ), and that of the [200,800] distance intervals is  $80 \text{ km}\cdot\text{h}^{-1}$  (about  $22 \text{ m}\cdot\text{s}^{-1}$ ). During implementation, this algorithm is terminated if the total number of iterations exceeds 15 or the relative gap is less than 0.001%. The best solution generated among the iterations is the approximate optimal solution to the model.

In each distance site, the potential velocity also has to be discretized into different speed stamps. First, the acceleration rates are assumed to be discretized into five cases, that is,  $1 \text{ m}\cdot\text{s}^{-2}$ ,  $0.5 \text{ m}\cdot\text{s}^{-2}$ ,  $0 \text{ m}\cdot\text{s}^{-2}$ ,  $-0.5 \text{ m}\cdot\text{s}^{-2}$ , and  $-1 \text{ m}\cdot\text{s}^{-2}$ . Second, in the acceleration stages, that is, the distance interval [0, 200], we assume that the acceleration rates are  $1 \text{ m}\cdot\text{s}^{-2}$ ,  $0.5 \text{ m}\cdot\text{s}^{-2}$ , and  $0 \text{ m}\cdot\text{s}^{-2}$ . In the deceleration stages, that is, the distance interval [800, 1000], three acceleration rates were obtained, namely,  $0 \text{ m}\cdot\text{s}^{-2}$ ,  $-0.5 \text{ m}\cdot\text{s}^{-2}$ , and  $-1 \text{ m}\cdot\text{s}^{-2}$ . In the distance interval [200, 800], all five cases were taken in the operation process. In this experiment, we determined the speed with the acceleration  $1 \text{ m}\cdot\text{s}^{-2}$  or  $-1 \text{ m}\cdot\text{s}^{-2}$  as the maximum speed in each stage under the acceleration or deceleration process. For example, as the initial speed is 0, the speed increases to approximately  $4.5 \text{ m}\cdot\text{s}^{-1}$  with the acceleration  $1 \text{ m}\cdot\text{s}^{-2}$ , and we denote this speed as the maximum in the first stage. In addition, we assume that a total of five different speed stamps exist in each stage; thus, we decrease the maximum speed by  $0.5 \text{ m}\cdot\text{s}^{-1}$  gradually. The five different speed stamps should be  $4.5 \text{ m}\cdot\text{s}^{-1}$ ,  $4.0 \text{ m}\cdot\text{s}^{-1}$ ,  $3.5 \text{ m}\cdot\text{s}^{-1}$ ,  $3.0 \text{ m}\cdot\text{s}^{-1}$ , and  $2.5 \text{ m}\cdot\text{s}^{-1}$  when the train operates in the first 10 m. In practice, more than five speeds are generated with the given acceleration rates at each stage; therefore, we correct all the speeds in the current stage as five, given the speed stamps for simplicity. With the speeds in two adjacent stages, we construct their relationships by links, and two indicators (travel time and random energy consumption) are associated with each link.

As stated, the proposed model has two critical parameters to influence the final optimal solutions; these are travel time budget and number of considered samples. The results are discussed in the following section.

**Impact of upper limit of total travel time.** In this experiment, we attempt to discuss the sensitivity of approximate optimal solution under different upper limits of total travel times, where we assume that the random energy consumption in each link has a total of 10 samples. Table 1 shows the computational results by varying the

upper limit of the total travel time between 80s and 89s. Evidently, all the relative gaps are less than 15.00%, where eight cases have relative gaps within 10.00%. This condition implies that the algorithm can obtain the high-quality solution with acceptable relative gaps between the upper and lower bounds. Also, note that the objective value, which is the expected energy consumption, is highly sensitive to the upper limit of the total travel time, which indicates that we have to carefully determine the upper limit in the solution process. The best upper limit should be 84s in this experiment, with the smallest relative gap of 5.10%.

**Table 1** Computational results with different upper limits

Upper limit/s	Upper bound/kWh	Lower bound/kWh	Relative gap/%
80	30.9875	28.1129	9.27
81	30.9710	29.1205	5.97
82	30.9913	27.5204	11.20
83	30.9980	26.5900	14.19
84	30.9575	29.3766	5.10
85	30.9577	28.8425	6.83
86	30.9659	28.4730	8.05
87	30.9766	28.7174	7.29
88	30.9487	28.6147	7.54
89	30.9729	28.3621	8.42

To further show the sensitivity of approximate optimal solution to the upper limit of travel time, Fig. 4 depicts the corresponding speed trajectories with different upper limits by different color curves. The figure suggests that the speed trajectories differ from each other when we vary the upper limit of the total travel time. Specifically, the speed trajectories significantly change when the speed increases from approximately  $12 \text{ m}\cdot\text{s}^{-1}$  to  $15 \text{ m}\cdot\text{s}^{-1}$  in the acceleration operation and decreases from  $15 \text{ m}\cdot\text{s}^{-1}$  to  $10 \text{ m}\cdot\text{s}^{-1}$  in the coasting operation.

**Impact of different sample size.** To verify the computational efficiency of the proposed algorithm, Fig. 5 presents six sets of experiments with different link energy consumption sample sizes, namely, 5, 10, 15, 20, 25, and 30. Moreover, we assume that the total travel time is less than 84s. Figure 5 implies that the computational times almost increased linearly with the increase in sample size, which indicates that the sample size has a considerable effect on computational efficiency. Particularly, in each iteration of the algorithm, the most time-consuming step is the generation of the optimal speed trajectory in each sample. Correspondingly, the approximate solutions, which are speed trajectories with different sample sizes, are depicted in Fig. 6. The maximum speed does not reach the speed limit of  $22 \text{ m}\cdot\text{s}^{-1}$  when the sample size is 30,

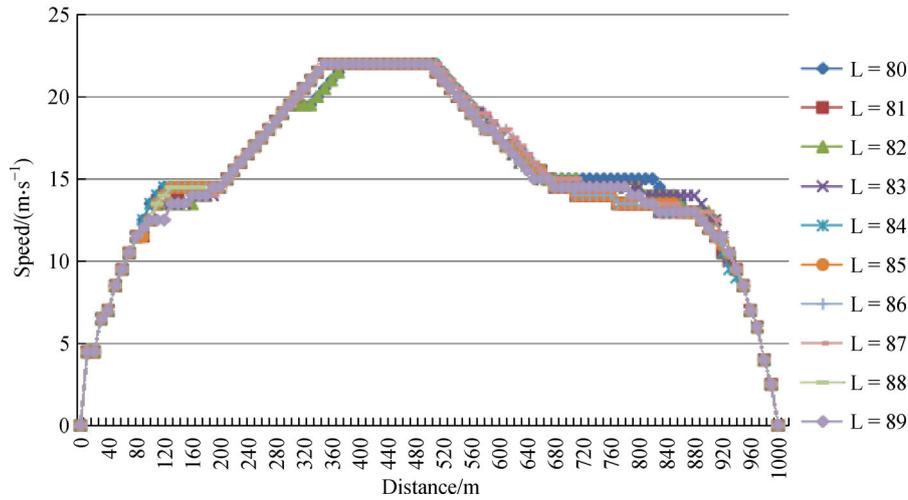


Fig. 4 Speed trajectories with different upper limits

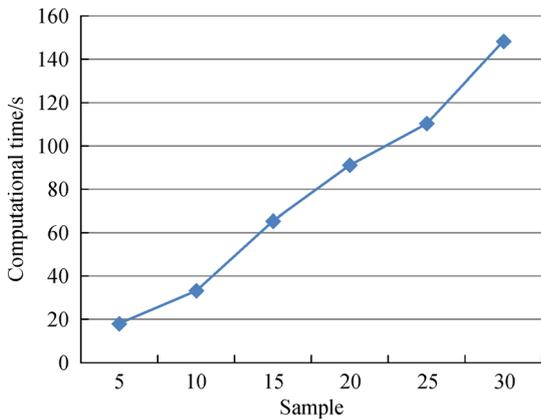


Fig. 5 Computational times with different samples

and the maximum speeds of other cases reach this speed limit.

## 7 Conclusions

Optimal speed trajectory plays a key role in urban rail transit systems. In this study, a novel method was presented to effectively formulate the problem of interest. The distance-speed network between two adjacent stations was constructed by dividing the distance into multiple sites of the same length, and the velocity shift link was constructed between different velocity stamps. The problem of interest is formulated as the shortest path problem. Moreover, due to the uncertainty of energy consumption, the sample-based method was adopted to

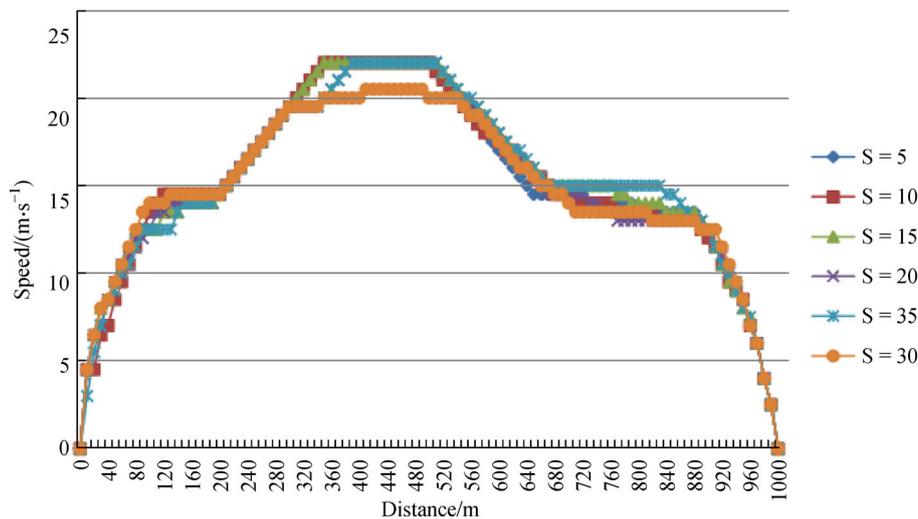


Fig. 6 Speed trajectories with different samples

depict randomness in the operation process. Considering the energy consumption and travel time of the involved train, a stochastic constrained shortest path model was formulated to illustrate the near-to-optimal speed trajectory. Finally, a Lagrangian relaxation-based algorithm was designed to solve the proposed model effectively, and its performance was demonstrated by a series of experiments.

In future research, we intend to focus on the optimal speed trajectories of multiple trains because in practice, more than one train operate at a single station. We also recognize that recovery energy is generated in the process of braking operations, and integrating the use of regenerative energy for multiple trains would be a meaningful research direction.

---

## References

- Asnis I A, Dmitruk A V, Osmolovskii N P (1985). Solution of the problem of the energetically optimal control of the motion of a train by the maximum principle. *U.S.S.R. Computational Mathematics and Mathematical Physics*, 25(6): 37–44
- Calderaro V, Galdi V, Graber G, Piccolo A, Cogliano D (2014). An algorithm to optimize speed profiles of the metro vehicles for minimizing energy consumption. In: *International Symposium on Power Electronics, Electrical Drives, Automation and Motion*
- Cao F, Fan L Q, Ke B R, Tang T (2016). Optimisation of recommended speed profile for train operation based on ant colony algorithm. *International Journal of Simulation and Process Modelling*, 11(3/4): 229
- Dominguez M, Cardador A F, Cucala A P, Pecharroman R R (2012). Energy savings in metropolitan railway substations through regenerative energy recovery and optimal design of ATO speed profiles. *IEEE Transactions on Automation Science and Engineering*, 9(3): 496–504
- Ghoseiri K, Szidarovszky F, Asgharpour M J (2004). A multi-objective train scheduling model and solution. *Transportation Research Part B: Methodological*, 38(10): 927–952
- Howlett P (2000). The optimal control of a train. *Annals of Operations Research*, 98(1–4): 65–87
- Ke B R, Chen M C, Lin C L (2009). Block-layout design using MAXCMIN ant system for saving energy on mass rapid transit systems. *IEEE Transactions on Intelligent Transportation Systems*, 10(2): 226–235
- Ke B R, Lin C L, Yang C C (2012). Optimisation of train energy-efficient operation for mass rapid transit systems. *IET Intelligent Transport Systems*, 6(1): 58–66
- Khmelnitsky E (2000). On an optimal control problem of train operation. *IEEE Transactions on Automatic Control*, 45(7): 1257–1266
- Kim K, Chien S I (2011). Optimal train operation for minimum energy consumption considering track alignment, speed limit, and schedule adherence. *Journal of Transportation Engineering*, 137(9): 665–674
- Ko H, Koseki T, Miyatake M (2004). Application of dynamic programming to optimization of running profile of a train. *Computers in railways IX*, 74: 10
- Liu S Q, Cao F, Xun J, Wang Y (2015). Energy-efficient operation of single train based on the control strategy of ATO. *IEEE 18th International Conference on Intelligent Transportation Systems*, 2580–2586
- Liu R F, Golovitcher I M (2003). Energy-efficient operation of rail vehicles. *Transportation Research Part A: Policy and Practice*, 37(10): 917–932
- Miyatake M, Ko H (2010). Optimization of train speed profile for minimum energy consumption. *IEEJ Transactions on Electrical and Electronic Engineering*, 5(3): 263–269
- Miyatake M, Matsuda K (2009). Energy saving speed and charge discharge control of a railway vehicle with on-board energy storage by means of an optimization model. *IEEJ Transactions on Electrical and Electronic Engineering*, 4(6): 771–778
- Tang H C, Dick C T, Feng X Y (2015). A coordinated train control algorithm to improve regenerative energy receptivity in metro transit systems. In *Transportation Research Record Board 94th Annual Meeting*, No. 15–1318
- Wang L, Yang L, Gao Z (2016). The constrained shortest path problem with stochastic correlated link travel times. *European Journal of Operational Research*, 255(1): 43–57
- Wang Y H, Schutter B D, Boom T J J V D, Ning B (2013). Optimal trajectory planning for trains-A pseudospectral method and a mixed integer linear programming approach. *IEEE Transportation Research Part C*, 29: 97–114