vehicles

MDPI

Article

Two-Step Optimization Method of Freight Train Speed Curve Based on Rolling Optimization Algorithm and MPC

Xubin Sun [1,*], Jingjing Li [1], Wei Zhang [2], Guiyang Sun [1], Xiyao Zhang [1] and Hongze Xu [1]

[1] School of Automation and Intelligence, Beijing Jiaotong University, Beijing 100044, China; 22120242@bjtu.edu.cn (J.L.); 23120249@bjtu.edu.cn (G.S.); 23125144@bjtu.edu.cn (X.Z.); hzxu@bjtu.edu.cn (H.X.)
[2] Casco Signal Ltd., Shanghai 200071, China; gan777lu@163.com
* Correspondence: xbsun@bjtu.edu.cn

Abstract: Given the considerable length and weight of freight trains, their operation can be quite challenging. Improper operation may lead to train decoupling and derailment. Driver Advisory Systems (DASs) are used in some countries to assist train drivers by providing the speed curves, which are desired to be easy to track. Multi-mass train model is a good choice to depict the in-train forces in train speed curve generating, but its application is often hindered by the computation time. A single mass train model is considered as another choice to simplify the computation. To exploit the advantages of the multi-mass and single-mass models, this paper proposes a Two-step Optimization Method to generate the optimal speed curves for the freight trains. In the first step, the Rolling Optimization Algorithm (ROA) is proposed to optimize the speed curve on the basis of the single-mass model, taking the train energy consumption and punctuality as the optimization objectives. In order to assist the driver in operating the train smoothly, the speed curve generated by the ROA was tested on DAS, but it could not be followed accurately in the actual operation. To solve this problem, a Model Predictive Control (MPC) algorithm based on a multi-mass model is adopted as the second optimization step, which takes the output speed curve of the ROA as the reference speed curve. The MPC algorithm will generate a new speed curve, taking in-train forces, energy consumption and punctuality as the optimization indices. Simulations are carried out using the data from the Dalailong railway in China to evaluate the proposed method. The simulation results show that the speed curves generated by the Two-step Optimization Method are smoother than that of the ROA, and the throttle sequences are more conducive for the driver to follow in practical operation. The simulation results show that the energy consumption is reduced by 17.1% compared to that of the ROA simulation. The speed curve also can be integrated into the onboard DAS or the Automatic Train Operation (ATO) system, aiming to obtain a smooth and energy-efficient train operation.

Keywords: MPC; rolling optimization; train operation optimization; energy saving

check for updates

Academic Editor: Mohammed Chadli

Received: 27 December 2024
Revised: 4 February 2025
Accepted: 7 February 2025
Published: 14 February 2025

Citation: Sun, X.; Li, J.; Zhang, W.; Sun, G.; Zhang, X.; Xu, H. Two-Step Optimization Method of Freight Train Speed Curve Based on Rolling Optimization Algorithm and MPC. *Vehicles* **2025**, *7*, 17. https://doi.org/10.3390/vehicles7010017

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In recent years, freight transportation has quickly increased [1]. As freight trains are heavy and long, it is difficult to operate freight trains smoothly. If the longitudinal impulse is too big, it will shorten the life of train couplers and even cause fatal accidents such as coupler breakdown and derailment. The operation strategies also influence the train's energy consumption. Due to the differences in driving skills and the experience of the drivers, the train energy consumption is also different. Even for the same driver, the train

operation is also affected by their state, physical and mental, operating conditions, line conditions, and the train condition. Therefore, it is necessary to improve the train operation by providing a smooth and energy-efficient train speed curve, which can be integrated into the Driver Advisory System (DAS) or Automatic Train Operation (ATO) system.

When optimizing the train speed curve, establishing a train dynamic model is an essential step. Single-mass and multi-mass models are both widely studied by scholars. The single-mass model is characterized by its comparatively fast computation speed. Howlett used the single-mass model to investigate the energy-efficient train control method [2]. Based on the single-mass model, Howlett used a new local energy minimization principle to calculate the optimal switching points with steep gradients [3]. Pudovikov simulated the train operation dynamic processes to improve the control effect of the automatic control system for freight train speeds [4]. The single-mass model cannot depict the in-train force, while the multi-mass model can consider the in-train force. Consequently, the optimization outputs based on the multi-mass model are more applicable to the real-world situation. Gruber introduced the concept of the multi-mass model by studying the problem of optimizing the heavy trains towed by multiple locomotives [5]. Chou conducted an optimal control study for energy saving and speed tracking problems using the LQR (Linear Quadratic Regulator) methodology based on the validated multi-mass model [6]. Zhuan proposed an output regulation method with measurement feedback control of heavy-duty trains based on the multi-mass model, controlling the train speed within a specified speed range [7].

The train speed curves are initially optimized by Pontryagin maximum principle rules and other algorithms. Howlett provided the key equations for optimal control strategies with minimum energy consumption and punctual operation under continuously changing gradients [8]. Khmelnitsky investigated the problem of train energy-saving operation under arbitrary changes in speed and gradient and realized the corresponding energy-saving control strategy in this context [9]. Martinis proposed a method for generating energy-saving speed curves for freight trains by analyzing data from a real-time vehicle monitoring system [10]. Ko used a dynamic programming method to generate energy-efficient train operating speed curves and reduced the computation time and speed errors by improving the dynamic programming method [11]. Albrecht proposed a two-layer optimization algorithm, which is able to compute and generate a sequence of energy-saving optimization operations for trains with a minimum number of condition transitions [12]. Guo used the particle swarm algorithm and the NSGA-II (Non-dominated Sorting Genetic Algorithm-II) algorithm to propose a multi-objective optimization algorithm based on the energy consumption and the running time [13]. Yi used the greedy algorithm and improved the multi-objective bald eagle search algorithm to solve a multi-objective optimization problem for freight trains [14]. Yang designed a fuzzy adaptive genetic algorithm (FAGA) for a multi-objective curve optimization strategy and used the improved high-order model-free adaptive iterative learning control algorithm to track the optimized speed curve [15].

There are many methods that can be used in freight train control to track the reference speed curve, in which the Model Predictive Control Algorithm considering the train multi-mass model is widely used in solving the train control problem. The MPC algorithm can also be seen as a short-distance optimization method for the train speed curve. Isna Silvia applied the MPC method to predict and adjust the dwell time and the running time of the train [16]. Song designed a constrained MPC closed-loop control algorithm and customized the target curve acting on the MPC according to the operating conditions of the loaded train [17]. Liu solved the optimal control problem based on the MPC framework under the constraints of operating conditions and analyzed the impact of different weight factors. The simulation results of MPC are compared with fuzzy control and PID (Proportion Integration Differentiation) control and verified the feasibility of the model predictive

controller [18]. A nonlinear MPC algorithm is proposed to optimize train adhesion control by identifying the trajectory parameters online using a dynamic forgetting factor recursive least squares method, and then these parameters are applied to the MPC to achieve optimal adhesion control [19]. In order to solve the optimal operation of heavy-duty trains, the MPC method was used to optimize train performance and operation through the analysis of various penalty factors, weight coefficients, etc., and the MPC reference speed is the self-setting step-type curves [20–23]. The above studies on MPC algorithms usually use a step curve to test the control performance, and it is possible to optimize the train operation by taking advantage of the MPC algorithm. MPC can only be used in short-term optimization of the freight train speed curve.

To leverage the advantages of the single-mass model and multi-mass model, a Two-Step Optimization Method is proposed in this paper. Firstly, using ROA to generate a speed curve based on a single-mass model. Secondly, inspired by the MPC control method, the paper will use the MPC method to re-optimize the train speed curve based on the multi-mass model. The main contributions of this paper are as follows:

- Considering the multi-mass characteristics of the train, the slopes of the line are processed by the gradient equivalent method, which will combine some adjacent slopes to reduce the calculation.
- The Rolling Optimization Algorithm (ROA) based on a single mass model is proposed, which will generate a speed curve with the objectives of energy-saving and punctuality.
- Inspired by the principle of MPC, this paper proposes a Two-step Optimization Method based on the multi-mass model. The ROA is used in the first step to generate the speed curve, which is taken as the reference curve in the second optimization step of the MPC method.
- Benefiting from the advantages of the single mass model and the multi-mass model, the speed curves generated by the Two-step Optimization Method are more in line with actual operation than that of the ROA.

2. The Two-Step Optimization Method

In this paper, a Two-Step Optimization Method is adopted to optimize the freight train speed curve, and the framework of the proposed algorithm is shown in Figure 1. The brown box indicates the contents of ROA based on single mass model, the purple box indicates the contents of MPC based on multi-mass models. ROA is designed based on the previous work of our group [24], which is on the basis of the single-mass model. In this algorithm, the length and weight of the train are taken into account to equalize the values of the line gradient according to the gradient equivalence method. The optimization index includes the train running time and energy consumption.

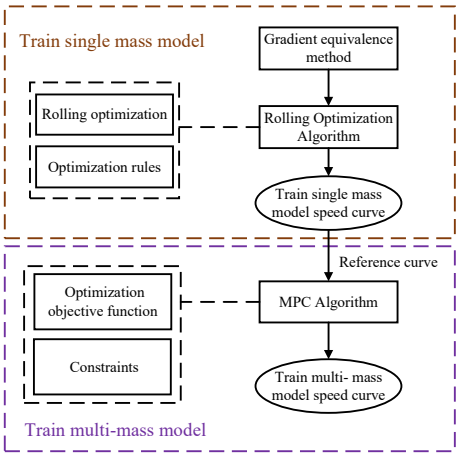


Figure 1. Two-step optimization method framework.

2.1. Gradient Equivalent Method

When the single-mass model is used in the freight train simulation, the gradient resistance force will change suddenly at the slope-changing positions. Therefore, a gradient equivalence method is proposed, taking the length and weight of the train into account. Assume that the weight of the train is uniformly distributed along the length of the train, and the train slope resistance is distributed along the length of the train.

Define the equivalent gradient $i_c(x)$ when the position of the train head is x, which can be calculated by

$$i_c(x) = \frac{\sum_{k=1}^{n} m_k l_k i_k(x)}{LM}, \tag{1}$$

where n is the number of all locomotive/wagon in the whole train, m_k is the weight of the kth locomotive/wagon, l_k is the length of the kth locomotive/wagon, $i_k(x)$ is the original gradient value where the kth locomotive/wagon is located, L is the total length of the train, and M is the total weight of the train. A gradient comparison is given in Figure 2, where the subfigure above gives the original gradient and the subfigure below gives the equivalent gradient of a demo line.

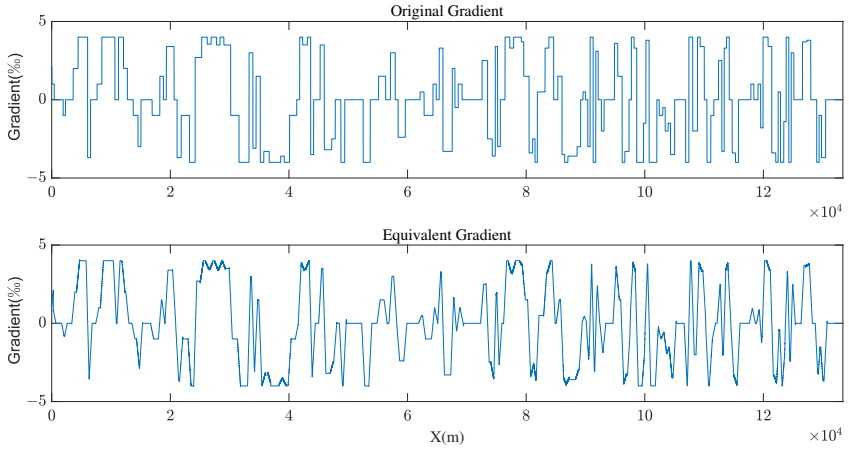


Figure 2. Comparison of the original and the equivalent gradient.

When the cruising speed is given, the line gradients can be classified into four types as shown in Table 1, where i_s is defined as the gradient on which the train can keep the cruise speed with maximum traction force. Similarly, i_x is defined as the gradient on which the train can keep the cruise speed with maximum braking force, the a, b, c, d are the symbols of different gradient types. Since there is no steep downhill section in the selected simulation line, it will not be discussed in the subsequent research.

Table 1. Classification of gradient types.

Gradient Types	Gradient Values	Type Symbols
General section	$i_s \geq i_c \geq 0$	a
Steep uphill section	$i_c > i_s$	b
Downhill section	$i_x < i_c < 0$	c
Steep downhill section	$i_c \leq i_x$	d

According to Equation (1), calculate the equivalent gradient at each position and combine the successive gradients with the same type as an equivalent slope. The demonstration of the equivalent slopes of the line is shown in Figure 3, the a, b, c are the same as the type symbols in the Table 1. In order to find the optimal operation solution, set three slopes as a group; six slope combinations can be obtained as [a b a], [a c a], [b a c], [b a b], [c a b],

[c a c]. As can be seen from Figure 3, the slope combinations include [a b a], [b a c], [a c a]. The gradient values are continuous after the calculation of the gradient equivalence algorithm, so there is no sudden change from a steep uphill section to a downhill section. In other words, there will be a general section between sections with varying gradients.

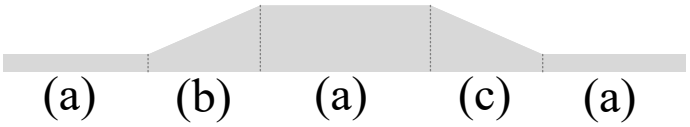


Figure 3. Slope type demonstration.

2.2. Rolling Optimization Algorithm(ROA) for Freight Train Speed Curve

A Rolling Optimization Algorithm is proposed to find the optimal freight train speed curve. The flowchart of the rolling optimization algorithm is shown in Figure 4. The cruise speed is calculated according to the train information, and the switching points of the optimal operation strategies are optimized by the bisection method. Based on each cruise speed, the speed curve is calculated in a rolling strategy. Firstly, equivalent gradients of the line are calculated by Equation (1). Secondly, we take three gradients as a group in each rolling step and calculate the optimal operation strategies according to the optimization rules. Roll forward to the next three gradients to continue the calculation, and stop the process until reaching the end of the line.

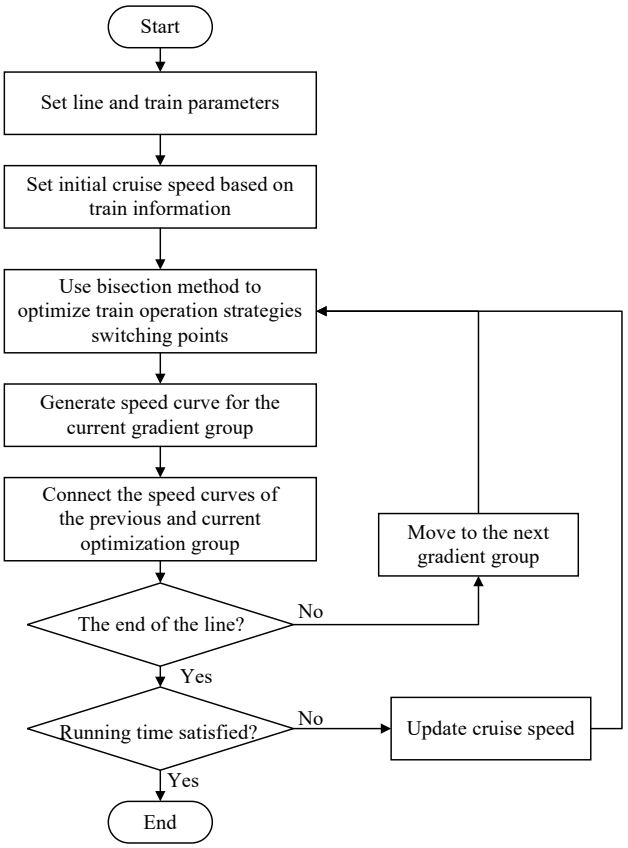


Figure 4. Flowchart of Rolling Optimization Algorithm for freight train speed curve.

2.2.1. Train Speed Curve Optimization for Each Gradient Group

The ROA takes punctuality and energy saving as the optimization objectives and establishes the optimization objective function as shown in

$$J = \sum_{i=1}^{N} \alpha \cdot |T_s - T_p| + \beta \cdot E_{S_s,S_e} \rightarrow min, \tag{2}$$

where T_s is the simulated train running time in the section, T_p is the planned running time in the section, E_{S_s,S_e} is the energy consumption of the train running in the section$[S_s, S_e]$, S_s is the starting point of the current train operating section, S_e is the endpoint of the current train operating section, α and β are the weights of the objective function, and N is the slope combination number.

Considering the time constraints and throttle constraints of the train during operation, the constraint analysis is given. Regarding the time constraints, the optimized train running time will have a certain error with the planned running time. Define the maximum deviation value as t_{max}, and then the section running time constraints shown in (3). In this way, the whole process running time constraints can also be defined as

$$\begin{gathered} |T_s^i - T_p^i| \leq t_{max} \\ |\sum_{i=1}^{N}(T_s^i - T_p^i)| \leq T_{max}, \end{gathered} \tag{3}$$

where T_{max} is the maximum error value of the whole running time.

Regarding the throttle constraints, they can be expressed by the following:

$$G_{min} \leq G \leq G_{max}, \tag{4}$$

where G is the current throttle of the train, G_{max} is the maximum throttle that the train is allowed, G_{min} is the minimum throttle that the train is allowed. The throttle constraints are set according to the Code for Calculation of Train Traction for the relevant parameters of the DF_{8B} diesel locomotive.

Therefore, the optimization objective function and constraints of the ROA are as follows:

$$\begin{gathered} J = \sum_{i=1}^{N} \alpha \cdot |T_s^i - T_p^i| + \beta \cdot E_{S_s,S_e} \rightarrow min \\ s.t. \begin{cases} |T_s^i - T_p^i| \leq t_{max} \\ |\sum_{i=1}^{N}(T_s^i - T_p^i)| \leq T_{max} \\ G_{min} \leq G \leq G_{max} \end{cases} . \end{gathered} \tag{5}$$

The main optimization rules of the ROA are given as follows. When a train enters an uphill section, it will search for an optimal switching point of the traction condition in advance. When the train pulls out of the uphill section, the running speed will not be much lower than the cruise speed. When a train enters a downhill section, it will search for an optimal switching point of the coast condition in advance. When the train pulls out of the downhill section, the running speed will not be much higher than the cruise speed. The comparison of different traction switching points for the uphill section is shown in the Figure 5. The comparison of different coast-switching points for the downhill section is shown in Figure 6. The range of the traction or coast switching point is the length of the general section. The v_{pro} is the limit speed, which is the speed that trains are not allowed to exceed.

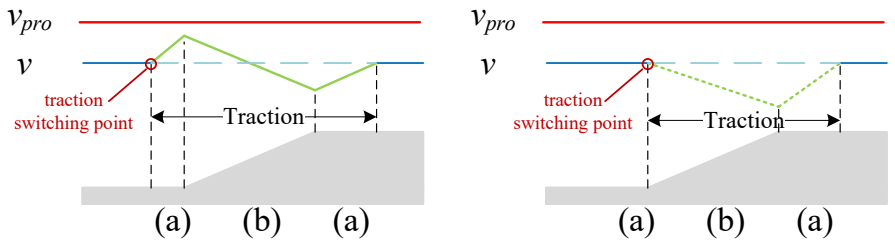


Figure 5. Comparison of the traction switching points for the uphill section.

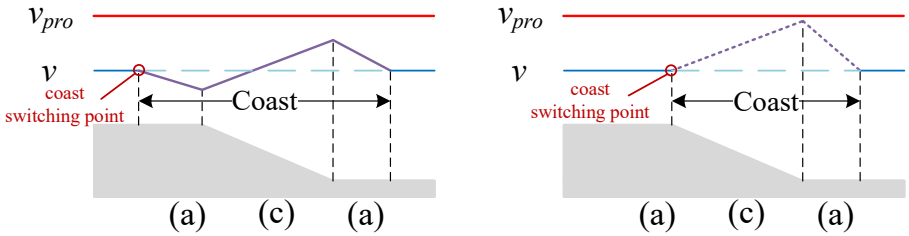


Figure 6. Comparison of the coast switching points for the downhill section.

Combined with the optimization objective function, there exists an optimal position of the switching point, which makes the objective function the minimum. Therefore, the trisection method is used to find the optimal position of the switching points in uphill and downhill sections, and the specific steps are as follows.

Step 1: Set the boundaries of the switching point $[SP_l, SP_r]$ as the start and end points of the gentle gradient section. Calculate the corresponding objective functions $[OF_l, OF_r]$ of the switching point boundaries.

Step 2: Divide $[SP_l, SP_r]$ into three equal parts and compute two trisections SP_{mid1} and SP_{mid2}, where SP_{mid1} is close to SP_l and SP_{mid2} is close to SP_r. Calculate the corresponding objective function values OF_{mid1} and OF_{mid2}.

Step 3: If $OF_{mid1} > OF_{mid2}$, update the left array boundary to SP_l according to the (6). If $OF_{mid1} < OF_{mid2}$, update the right array boundary to SP_r according to the (7).

$$SP_l = SP_{mid1} = SP_l + \frac{SP_r - SP_l}{3}. \tag{6}$$

$$SP_r = SP_{mid2} = SP_l + \frac{2 \times (SP_r - SP_l)}{3}. \tag{7}$$

Step 4: Repeat the above steps, and each step will narrow the interval down to two-thirds of its original length until the optimal switching point is found.

2.2.2. Speed Curves Connecting Method

According to the gradient equivalence method, three continuous slopes are combined as a group. The ROA is carried out on the basis of this combination to obtain the speed curve of each group. Then, the optimized speed curves are connected to obtain the train running speed curve of the whole line. Figure 7 shows the speed curves rolling connecting method of some of the groups. The blue line indicates the train speed of general section, the green line indicates the train speed of steep uphill section, and the purple line indicates the train speed of downhill section. The red vertical dashed line indicates the start point of speed curve where the connection needs to be made.

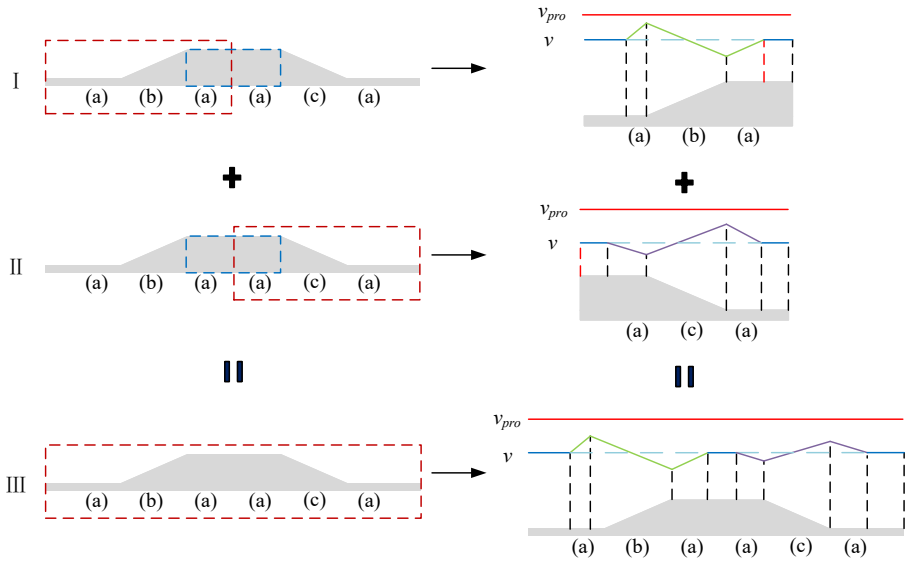


Figure 7. Speed curves connecting method.

2.2.3. Cruise Speed Updating Method

Firstly, the initial cruise speed $V_{initial}$ is calculated based on the start and end points of the train section and the planned running time, the calculation formula is shown in

$$V_{initial} = \frac{S_{ea} - S_{sa}}{T_{pa}}, \tag{8}$$

where T_{pa} is the planned running time in the whole train operating line, S_{sa} is the starting point of the whole train operating line, S_{ea} is the endpoint of the whole train operating line.

When the optimized train running time does not satisfy the running time constraints, the cruise speed needs to be adjusted with the following rule:

Step 1: Given the difference threshold T_{op} between the optimized train running time and the planned running time, compare the T_{sa} and T_{pa} in the section, where T_{sa} is the simulation running time of the whole train operating line, and let $T = |T_{sa} - T_{pa}|$. If $T < T_{op}$, the running time constraint is satisfied, and then the optimization is finished; otherwise, the cruise speed needs to be adjusted and re-optimized.

Step 2: Obtain the adjustment value ΔV of the cruise speed according to

$$\Delta V = \begin{cases} 1, T > 2T_{op} \\ \dfrac{T_{op}}{T}, T_{op} < T \leq 2T_{op} \end{cases}. \tag{9}$$

Step 3: When adjusting the speed of the cruise, if $T_{sa} > T_{pa}$, which indicates that the current cruise speed is comparatively low and the train is running slow, let $V_{new} = V + \Delta V$. If $T_{sa} < T_{pa}$, this indicates that the the current cruise speed is comparatively high and the train is running faster, then let $V_{new} = V - \Delta V$. Where V_{new} is the adjusted cruise speed.

Step 4: Use the V_{new} to re-optimize the speed curve until the running time constraint is satisfied.

2.3. MPC Algorithm for Freight Train Speed Curve Re-Optimization

The principle of the MPC algorithm can be specifically described as follows. At each sampling moment, an open-loop optimization problem is solved in real time for a finite

period of time based on the current system state information. Then, the first few groups of data in the generated control sequence are applied to the controlled model as the control output. At the next sample moment, the optimization problem is solved again using the updated system state information as the initial condition to predict the future state of the system repeatedly [25]. Thus, the main advantage of the MPC algorithm is that it balances the effects of ideal optimization and practical uncertainty on the future finite time horizon because the rolling optimization strategies always establish new optimization objectives on the latest information. This is more practical and effective than the traditional optimal control established under ideal conditions. The principle of the MPC algorithm is shown in Figure 8.

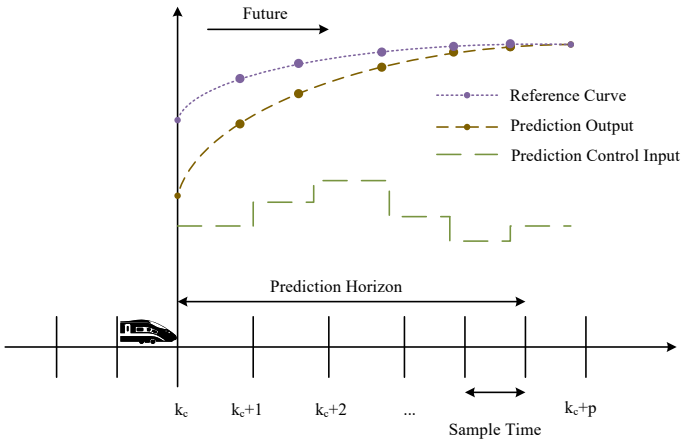


Figure 8. MPC algorithm demonstration.

2.3.1. The Multi-Mass Train Model

The multi-mass model takes into account not only the length of the train but also the coupler force, as shown in Figure 9, and it is built according to Longitudinal Train Dynamics. The coupler force is a direct index for evaluating the longitudinal impulse. For long-distance (heavy-duty) freight trains, the operation regulation recommends that the maximum coupler pulling force should be less than 2000 kN, and the maximum coupler pressing force should be less than 2250 kN [26]. The subsequent constraints on the coupler force will consider this recommendation.

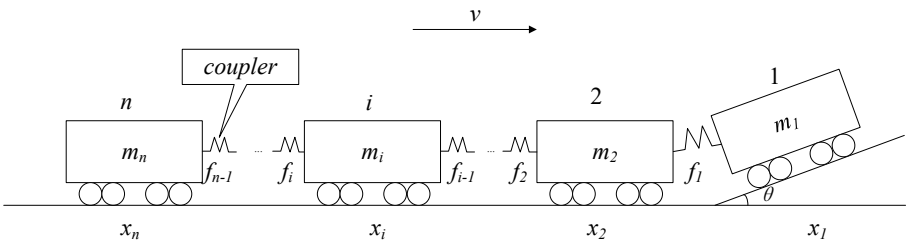


Figure 9. Train multi-mass model.

The coupler can be regarded as a spring with damping characteristics, and then the coupler force is calculated by

$$f_i = k_i(x_i - x_{i+1} - l_c) + d_i(v_i - v_{i+1}), \tag{10}$$

where k_i and d_i are the coefficients related to the elastic and damping force of the coupler, x_i and v_i are the displacement and speed of the ith locomotive/wagon, respectively, l_c is the distance of two neighboring couplers when they are in the nature state. In the subsequent

calculations, we only consider the elastic force to simplify the simulation. Based on the multi-mass model, the force analysis for a single locomotive/wagon is shown in Figure 10.

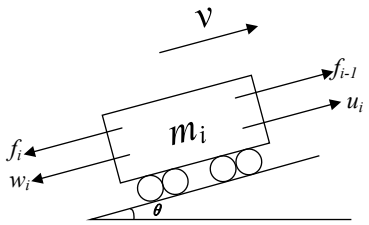


Figure 10. Force analysis for a single locomotive/wagon.

For the ith locomotive/wagon, u_i is the traction or braking force, f_{i-1} is the front coupler force, f_i is the latter coupler force, and m_i is the mass of the ith locomotive/wagon.

The first locomotive has no coupler force in front of it, and the last wagon has no coupler force behind it. Based on the force analysis of the locomotive and wagons, the whole train motion equation is defined by

$$\begin{cases} m_1\ddot{x}_1 = & u_1 - f_1 - w_1 \\ & \vdots \\ m_i\ddot{x}_i = & u_i - f_i + f_{i-1} - w_i, \\ & \vdots \\ m_n\ddot{x}_n = & u_n + f_{n-1} - w_n \end{cases} \tag{11}$$

where w_i is the total resistance force of the ith locomotive/wagon, which is

$$w_i = w_0 + w_g, \tag{12}$$

where w_g is the additional resistance on slopes. Other additional forces like curve resistance are not considered in this paper. The basic resistance force w_0 is defined by

$$w_0 = a + bv + cv^2. \tag{13}$$

In order to facilitate the subsequent design of the control algorithm, the train multi-mass model is simplified. Considering the train basic resistance force w_0 and unit slope additional resistance force w_g, define $\delta u_i = (a_i + c_i v_i^2)m_i + w_{g_i} (2 \leq i \leq n-1)$ and $\bar{u}_i = u_i - \delta u_i$. Define $x = [v_1, v_2, \ldots v_n, x_1, x_2, \ldots x_n]^T$ as the state vectors and $\bar{u} = [\bar{u}_1, \bar{u}_2, \ldots \bar{u}_n]^T$ as the decision vectors, and then the state equation of the linearized system is

$$\dot{x} = Ax + B\bar{u}, \tag{14}$$

where

$$A = \begin{bmatrix} A_{11} & A_{12} \\ I_{n\times n} & 0_{n\times n} \end{bmatrix},$$

$$A_{11} = \begin{bmatrix} -b_1 & 0 & \cdots & 0 \\ 0 & -b_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -b_n \end{bmatrix},$$

$$A_{12} = \begin{bmatrix} -\frac{k}{m_1} & \frac{k}{m_1} & 0 & \cdots & 0 & 0 & 0 \\ \frac{k}{m_2} & -\frac{2k}{m_2} & \frac{k}{m_2} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{k}{m_{n-1}} & -\frac{2k}{m_{n-1}} & \frac{k}{m_{n-1}} \\ 0 & 0 & 0 & \cdots & 0 & \frac{k}{m_n} & -\frac{k}{m_n} \end{bmatrix},$$

$$B = \begin{bmatrix} B_1 \\ 0_{n\times n} \end{bmatrix},$$

$$B_1 = \begin{bmatrix} \frac{1}{m_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{m_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{m_n} \end{bmatrix}.$$

The model after linearization is discretized in the subsequent simulation application, and the discretized model can be expressed as

$$x(k_c+1|k_c) = Ax(k_c) + B\bar{u}(k_c). \tag{15}$$

2.3.2. MPC Framework

According to the state Equation (15) of the multi-mass model of the freight train, assuming that the train operation state variable $x(k_c)$ is measurable at the moment k_c, the derivation of the state equation leads to the predicted state variable $\hat{x}(k_c+i|k_c)$ and is calculated as follows:

$$\begin{aligned} \hat{x}(k_c+1|k_c) &= Ax(k_c) + B\bar{u}(k_c) \\ \hat{x}(k_c+2|k_c) &= A\hat{x}(k_c+1|k_c) + B\bar{u}(k_c+1|k_c) \\ &= A^2x(k_c) + AB\bar{u}(k_c) + B\bar{u}(k_c+1) \\ &\cdots \\ \hat{x}(k_c+N_p|k_c) &= A^{N_p}x(k_c) + A^{N_p-1}B\bar{u}(k_c) + \\ &\quad A^{N_p-2}B\bar{u}(k_c+1) + \cdots + \\ &\quad A^{N_p-N_c}B\bar{u}(k_c+N_c-1) \end{aligned} \tag{16}$$

Generalizing the above equation yields

$$X = Gx(k_c) + \Gamma U, \tag{17}$$

where

$$X = \left[\hat{x}(k_c+1|k_c), \hat{x}(k_c+2|k_c), \cdots, \hat{x}(k_c+N_p|k_c)\right]^T,$$

$$U = \left[\bar{u}(k_c), \bar{u}(k_c+1), \cdots, \bar{u}(k_c+N_p)\right]^T,$$

$$G = \left[A, A^2, \cdots, A^{N_p}\right]^T,$$

$$\Gamma = \begin{bmatrix} B & 0 & \cdots & 0 \\ AB & B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N_p-1}B & A^{N_p-2}B & \cdots & A^{N_p-N_c}B \end{bmatrix},$$

where X is the train state at each predicted moment, U is the corresponding control input, and G, Γ are the coefficient matrices. The predicted state X can be predicted by (17), and the control sequence U within the prediction range can be obtained.

2.3.3. Optimization Objective Function

In freight train operation, energy saving, punctuality and coupler force are the key factors to ensure the safe operation of the train; the operation optimization objective function of the freight train is constructed by

$$J = \int_{t_0}^{t_1} \left[\sum_{i=1}^{n-1} k_a f_i^2 + \sum_{i=1}^{n} k_b u_i^2 + \sum_{i=1}^{n} k_d (v_i - v_r)^2 \right] dt, \tag{18}$$

where v_r is the reference speed, k_a, k_b, k_d are the weighting coefficients for the coupler force, energy consumption and speed tracking error, respectively, and t_0, t_1 represent the start and end time of the prediction period, respectively. The above objective function will be transformed into a discrete one for numerical calculation.

Based on (14), the above objective function is transformed into

$$J = \int_{t_0}^{t_1} \left(x^T L x + \bar{u}^T P \bar{u} + Q_1^T x + Q_2^T \bar{u} \right) dt \rightarrow min, \tag{19}$$

where

$$L = \begin{bmatrix} L_{11} & 0 \\ 0 & L_{22} \end{bmatrix},$$

$$L_{11} = diag\{k_d, \cdots, k_d\},$$

$$L_{22} = \begin{bmatrix} k_a k_1^2 & -k_a & \cdots & 0 & 0 \\ -k_a & 2k_a k_2^2 & -k_a & \cdots & 0 \\ 0 & -k_a & 2k_a k_3^2 & -k_a & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -k_a & k_a k_{n-1}^2 \end{bmatrix},$$

$$P = k_b I_{n \times n},$$

$$Q_1^T = -2k_d v_r [1, \cdots, 1, 0, \cdots, 0],$$

$$Q_2^T = 2k_b [\delta u_1, \cdots, \delta u_n].$$

Q_1 and Q_2 are related to the reference velocity v_r, which is determined based on the train position, while δu_i depends on the train position and speed; a computational method for handling these variables based on the train position and speed is given. Firstly, the speed in the predicted state $\hat{x}(k_c + i|k_c)$ is denoted as $v(k_c + i|k_c)$, and the position of

the train at the sample moment k_c is denoted as $s(k_c)$. Then, the position of the train at the sample moment $k_c + i$ can be expressed as

$$s(k_c + i|k_c) = s(k_c) + \sum_{j=1}^{i} T[\frac{1}{2}(v(k_c + j - 1|k_c) + v(k_c + j|k_c))], \tag{20}$$

where T is the sample time.

From the above equation, the predictions of Q_1 and Q_2 can be expressed, respectively, as a function of s by

$$\begin{aligned} \overline{Q_1}(s) &= \left[Q_1(k)^T, \cdots, Q_1(k + N_p)^T\right]^T \\ \overline{Q_2}(s) &= \left[Q_2(k)^T, \cdots, Q_2(k + N_p)^T\right]^T. \end{aligned} \tag{21}$$

Therefore, the train operation optimization objective function under the MPC framework can be expressed as

$$J = U^T H U + 2U^T f \rightarrow min, \tag{22}$$

where

$$H = \Gamma^T \overline{L} \Gamma + \overline{P},$$

$$f = \Gamma^T \overline{L} G x(k_c) + \frac{1}{2}(\Gamma^T \overline{Q_1} + \overline{Q_2}),$$

$$\overline{L} = diag\{L, \cdots, L\},$$

$$\overline{P} = diag\{P, \cdots, P\}.$$

2.3.4. Train Operational Constraints

The freight train operation should consider the constraints of the coupler force, and the traction and braking force of the locomotive. Therefore, the constraints of the freight train are given as follows:

$$\begin{aligned} f^{min} \leq f_{in_k} \leq f^{max}, \quad & k = 1, 2, \cdots, n - 1 \\ u_i^{min} \leq \overline{u_i} \leq u_i^{max}, \quad & i = 1, 2, \cdots, n \end{aligned}, \tag{23}$$

where f_{in_k} is the kth locomotive/wagon coupler force, f^{min} and f^{max} are the lower and upper bounds of coupler force, $\overline{u_i}$ is the ith traction and braking force of the locomotive/wagon, u_i^{min} and u_i^{max} are lower and upper bounds of the ith traction and braking force of locomotive/wagon.

The above constraints can be expressed in terms of state quantities:

$$X^{min} \leq X_{in} \leq X^{max}, \tag{24}$$

$$\begin{bmatrix} I_{(nN_c \times nN_c)} \\ -I_{(nN_c \times nN_c)} \end{bmatrix} U \leq \begin{bmatrix} U^{max}_{(nN_c \times 1)} \\ -U^{min}_{(nN_c \times 1)} \end{bmatrix}. \tag{25}$$

X^{max}, X^{min} and U^{max}, U^{min} are upper and lower bounds of X_{in} and U, respectively.

Firstly, according to the calculation method of train coupler force, a matrix Z is constructed to extract the relative displacement between neighboring compartments in the state variable, where $Z = \left[0_{(n-1)\times n} \quad I_{(n-1)\times(n-1)}\right]_{(n-1)\times(2n-1)}$. Then, $x_{in}(k) = Z x(k)$.

Secondly, according to the derivation process in (16), the same reasoning can be used to express the calculation of the coupler force as

$$X_{in} = G_x x(k_c) + \Gamma_x U, \tag{26}$$

where

$$G_x = \left[ZA, ZA^2, ZA^3, \cdots, ZA^{N_c}\right]^T,$$

$$\Gamma_x = \begin{bmatrix} ZB & 0 & \cdots & 0 \\ ZAB & ZB & \cdots & 0 \\ ZA^2B & ZAB & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ ZA^{N_c-1}B & ZA^{N_c-2}B & \cdots & ZB \end{bmatrix}.$$

Thus, the constraints on the coupler force can be expressed as follows:

$$\begin{bmatrix} \Gamma_x \\ -\Gamma_x \end{bmatrix} U \leq \begin{bmatrix} X^{max} - G_x x(k_c) \\ -X^{min} + G_x x(k_c) \end{bmatrix}. \tag{27}$$

By combining the train control force constraint (25) and the train coupler force constraint (27), the constraints for this optimization function can be obtained

$$RU \leq \sigma, \tag{28}$$

where

$$R = \begin{bmatrix} I_{(nN_c \times nN_c)} \\ -I_{(nN_c \times nN_c)} \\ \Gamma_x \\ -\Gamma_x \end{bmatrix},$$

$$\sigma = \begin{bmatrix} U^{max}_{(nN_c \times 1)} \\ -U^{min}_{(nN_c \times 1)} \\ X^{max} - \Gamma_x x(k_c) \\ -X^{min} + \Gamma_x x(k_c) \end{bmatrix}.$$

Finally, the overall optimization problem is formulated by the optimization objective function (22) and the constraints (28). This is a typical Quadratic Programming problem, which will be solved subsequently using the active set method.

3. Simulation Results

The simulation data in this paper, such as gradient and speed limit, are all the actual data of the Dalailong railway in China. The simulated train is composed of a DF_{8B} diesel locomotive and 42 C80 wagons, with a total length of 525.1 m and a total weight of 4338 t. The coupler elastic force coefficient k_i is 25,000. Part of the railway from Longkougang Station to Changyibei Station is considered in the simulation, with a length of 133.4 km. This simulation was performed in Matlab.

Table 2 gives the train operation results of the MPC algorithm with different objective function weighting coefficients, while the prediction step and optimization step of the MPC algorithm are fixed ($N_p = 6, N_c = 1$) when changing the weighting coefficients of the objective function. In Table 2, f_{max} and f_{min} denotes the maximum pulling coupler force and the maximum pressing coupler force for all locomotives/wagons during the simulation, respectively, and ΔT denotes the operation time difference between the MPC

algorithm and the ROA. Figures 11 and 12 give the maximum pulling coupler force and the maximum pressing coupler force for all locomotives/wagons during simulation with different weighting coefficients. When k_a = 0.1, k_b = 10, and k_d = 10,000, ΔT is 20 s, the maximum pulling and pressing coupler forces are within the recommended range, and the train speed is within the limit. Concurrently, the maximum pulling and pressing coupler forces, under the other objective function weighting coefficients, fall within the recommended limits.

Table 2. Train operation results with different objective function weights.

k_a	k_b	k_d	f_{max} (kN)	f_{min} (kN)	ΔT (s)
0.1	10	10,000	687.2	−804.2	20
0.1	10	30,000	838.9	−983.5	15
0.1	10	50,000	1343.6	−1292.4	25
0.01	10	10,000	1021.8	−1034.6	15
0.01	10	30,000	1701.7	−1974.5	20
0.01	10	50,000	1520.9	−1604.7	25

Figures 13 and 14 give the Two-step Optimization Method simulation speed curves for different values of k_a and k_d when k_b = 10, N_p = 6 and N_c = 1. The speed limit curves from high speed to low speed have been replaced by the maximum regular braking curves in Figures 13 and 14. As can be seen in Figure 13, the simulation speed curves with different objective function weighting coefficients have little change, and the speed curves are smoother compared to the ROA simulation curve. In Figure 14, it can be seen that when the train is operated through the speed limit section, none of them exceeds the speed limit value. The coupler force becomes larger as k_a decreases or k_d increases.

Table 3 gives the train operation results of the MPC algorithm with a different prediction step (N_p) and optimization step (N_c), when the weight coefficients of the objective function are fixed to k_a = 0.1, k_b = 10, k_d = 10,000. Figure 15 gives the maximum pulling coupler force and the maximum pressing coupler force for all locomotives/wagons during simulation with different N_p and N_c. The optimization result is N_p = 6, N_c = 1.

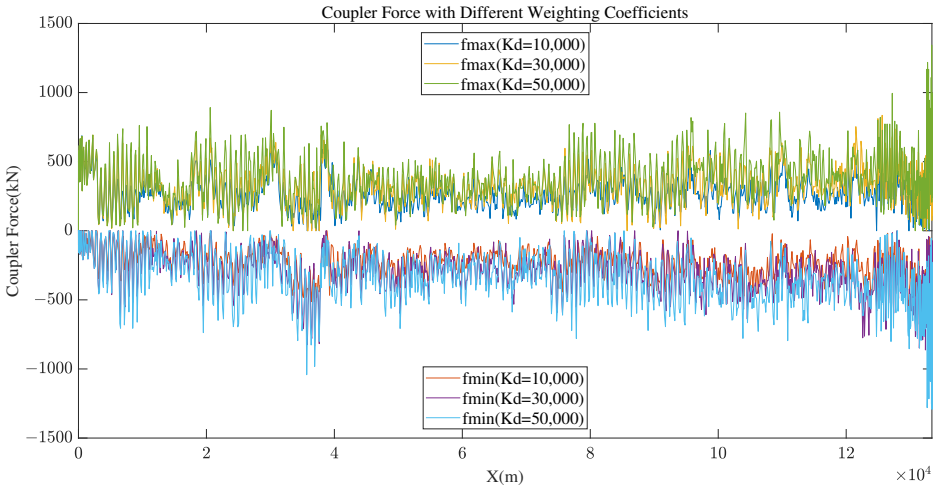


Figure 11. Coupler force with different k_d when k_a = 0.1, k_b = 10.

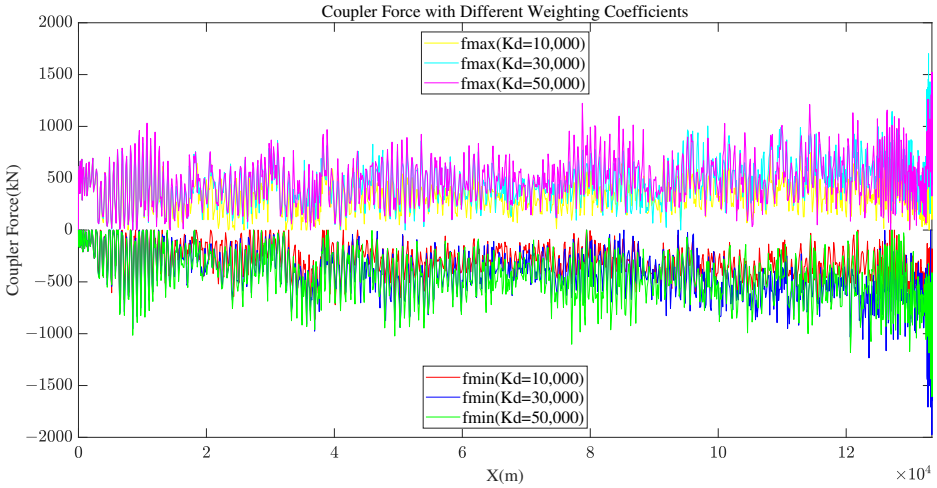


Figure 12. Coupler force with different k_d when $k_a = 0.01$, $k_b = 10$.

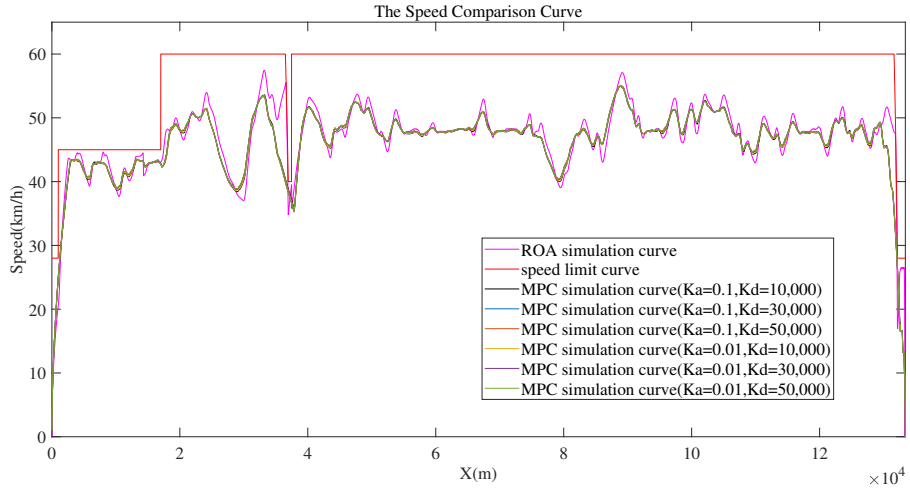


Figure 13. Simulation speed curves with different weighting coefficients.

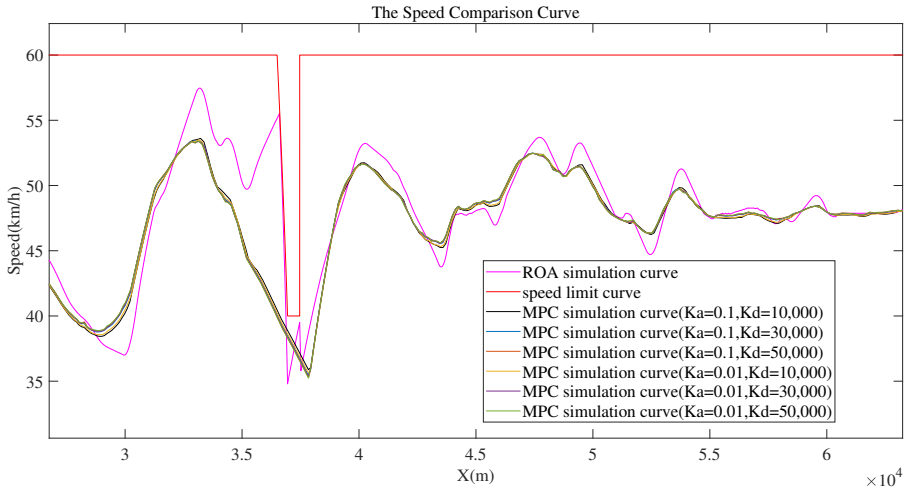


Figure 14. Partially enlarged simulation speed curves with different weighting coefficients.

Table 3. Train operation results with different prediction steps and optimization steps.

N_p	N_c	f_{max} (kN)	f_{min} (kN)	ΔT (s)
6	6	763.9	−875.8	35
6	3	831.9	−1005.4	35
6	1	687.2	−804.2	20
3	3	785.7	−706.5	160
3	1	868.4	−851.1	150

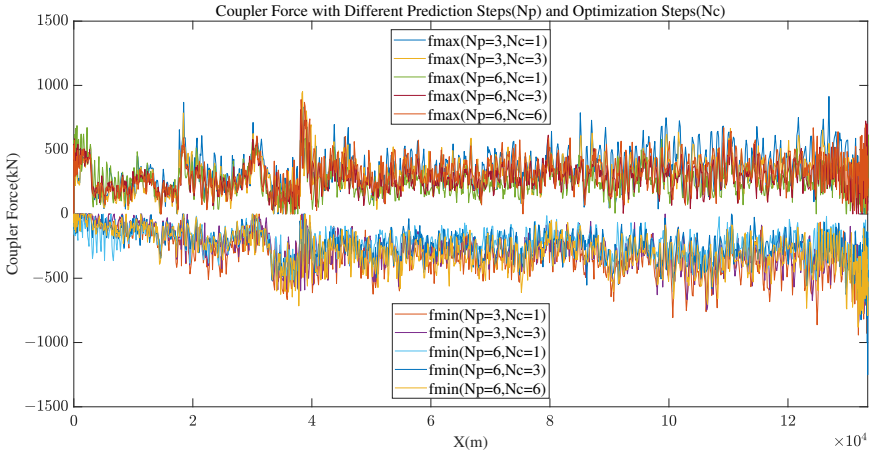


Figure 15. Coupler force with different N_p and N_c.

Figures 16 and 17 give the Two-step Optimization Method simulation speed curves for different values of N_p and N_c when k_a = 0.1, k_b = 10, k_d = 10,000. From Figure 16, it can be seen that when the values of N_p and N_c are changed, there are some differences in the generated speed curves, but they are smoother compared to the ROA simulation curve. It can be seen from Figure 17 that when the train is operated through the speed limit section, the speed limit value is not exceeded whatever the values of N_p and N_c. When changing the value of N_p and N_c, the train speed curves change obviously, while the change in coupler force is not significant. When N_p increases from 3 to 6 with a fixed N_c, the simulation speed curves closer to the ROA simulation curve and more in line with speed limits. Meanwhile, with the decrease in N_c, while the N_p is fixed, the simulation speed gradually changes towards the speed limit range, and when N_p = 6, N_c = 1 the simulation results achieve a reasonable performance. Therefore, the train will have better operation with longer prediction steps and shorter optimization steps.

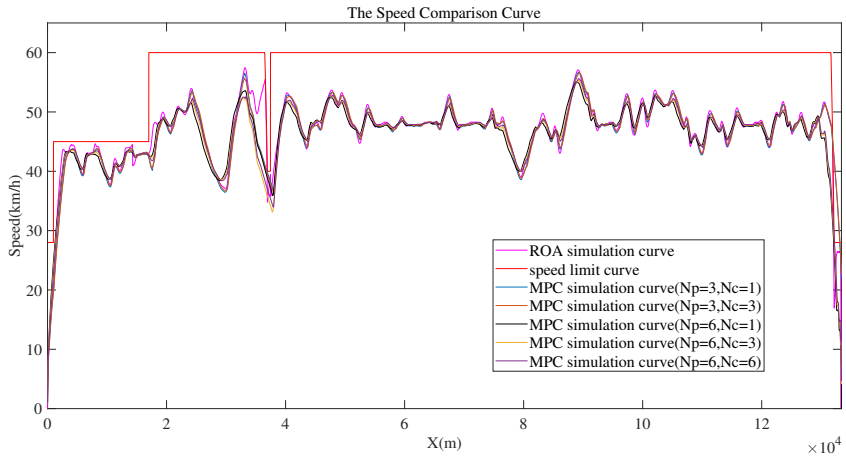


Figure 16. Simulation speed curves with different N_p and N_c.

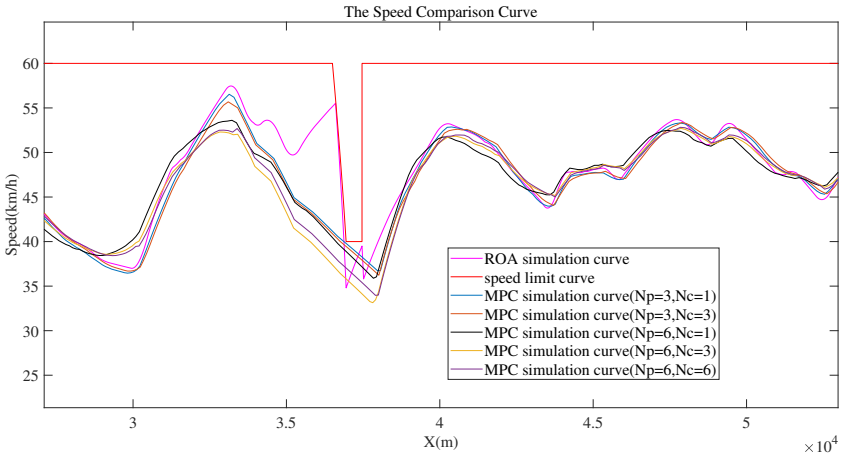


Figure 17. Partially enlarged simulation speed curves with different N_p and N_c.

Summarizing the above simulation results, when $k_a = 0.1$, $k_b = 10$, $k_d = 10{,}000$, $N_p = 6$ and $N_c = 1$, the maximum pulling coupler force is 34.4% of the pulling coupler force limit value, and the maximum pressing coupler force is 35.7% of the pressing coupler force limit value. It can be shown that the optimized train operation is smoother.

Figure 18 gives the simulation results when $k_a = 0.1$, $k_b = 10$, $k_d = 10{,}000$, $N_p = 6$ and $N_c = 1$. In the first subfigure, the red line is the speed limit curve, the purple line is the reference speed curve, the green line is the actual speed curve, and the black line is the simulation curve of the MPC algorithm. The actual speed curve is plotted from the data in the LKJ (Train Operation Monitoring Equipment). The speed limit curves from high speed to low speed also have been replaced by the maximum regular braking curves in Figures 16–18. In the second subfigure, the blue line represents the maximum pulling coupler force and the red line represents the maximum pressing coupler force in all locomotives/wagons at each sampling time. The third subfigure represents the line gradient curve. The horizontal coordinate indicates the distance. From the first subfigure, it can be seen that the speed curve obtained from the Two-step Optimization Method has a smaller range of speed fluctuations and is smoother. In the second subfigure, the coupler force variations of the MPC method are small and within safe limits.

Figure 19 gives the simulation results of the train control strategy. The first subfigure represents the train control strategy of the MPC algorithm. The second subfigure represents that of the Rolling Optimization Algorithm. The third subfigure represents the gradient curve. The vertical coordinate values in the first two subfigures represent the throttles, where $Gear = 0$ means that the train has neither traction nor braking force, $Gear > 0$ means that the train is in traction, and $Gear < 0$ means that the train is braking. The horizontal coordinate indicates the distance. It can be seen that the control strategies generated by the MPC algorithm and the Rolling Optimization Algorithm are similar, and when the train travels between 60 km and 72 km, the throttles generated by the MPC algorithm change less. Hence, the control strategy generated after the Two-step Optimization Method is easier to follow for the drivers during the actual operation.

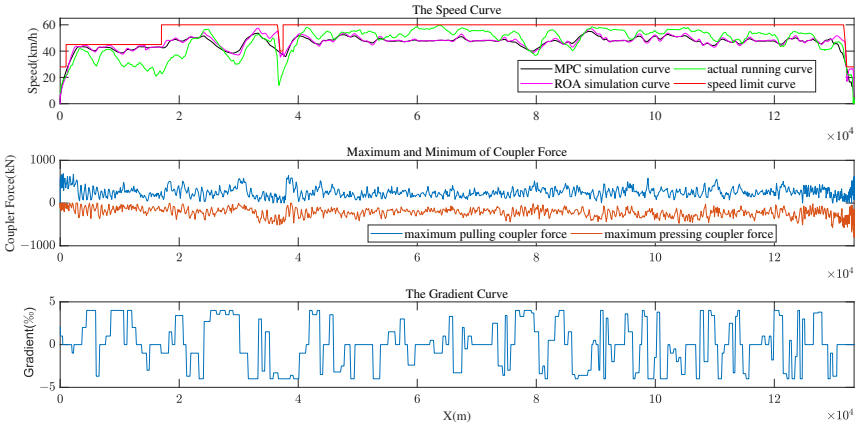


Figure 18. Train speed curves and coupler force comparisons.

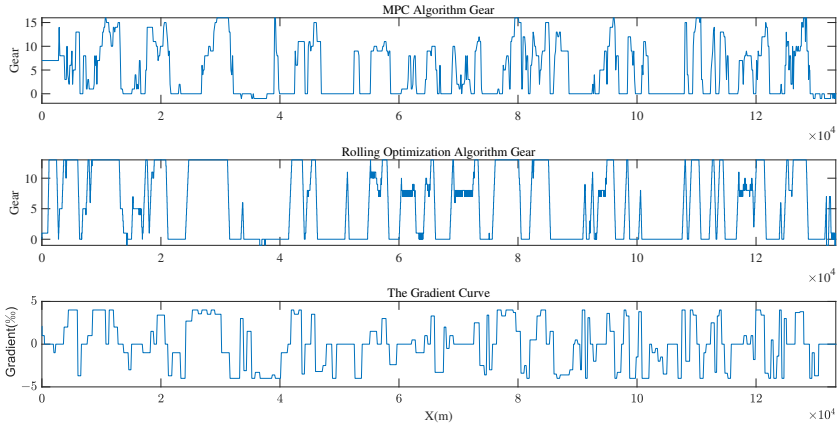


Figure 19. Control strategy comparison chart.

The train fuel consumption based on ROA simulation is 8.09 kg/10^4 t·km, the train fuel consumption based on MPC algorithm simulation is 6.71 kg/10^4 t·km, and the actual train fuel consumption calculated by LKJ data is 8.49 kg/10^4 t·km. Therefore, the Two-step Optimization Method can reduce energy consumption by 20.9% and 17.1% compared to actual operation and ROA simulation, respectively. The proposed algorithm can be integrated into the Driver Advisory System (DAS). The computational time of the proposed algorithm is less than one minute. The preparation time before the train departure is much longer than one minute, so the computational time is within the operating limit of the Driver Advisory System (DAS). The proposed algorithm can be integrated into the ATO system to generate the optimized reference speed curve, which can be tracked by the control algorithm of the ATO system.

4. Discussion

In this paper, a Two-step Optimization Method of freight train speed curve is proposed. The Rolling Optimization Algorithm is used to optimize the speed curve, and the optimized speed curve is used as the reference curve of the MPC algorithm for the second optimization step. The ROA takes into account the gradient equivalence method and the single-mass train model, while the MPC algorithm employs a multi-mass train model to consider in-train forces and establish an objective function. The optimal train operation results formed by different parameters are calculated by adjusting the weight coefficients of the objective function and the prediction step and optimization step of the MPC algorithm.

Our simulation results, based on data from China's Dalailong railway, demonstrate that the speed curves generated by the Two-step Optimization Method exhibit superior smoothness and are more aligned with actual operational constraints. Compared to actual operations, this method achieves a remarkable 20.9% reduction in energy consumption and a 17.1% reduction when compared to the ROA simulation. The optimized throttle sequences require less frequent adjustments, providing a more practical guide for drivers during actual operations.

Looking ahead, we plan to further test and refine the train-suggested speed curve optimization program based on the outcomes of the Two-step Optimization Method. The ultimate goal is to integrate this optimization algorithm into the Driver Advisory System (DAS) to offer drivers a more effective reference, thereby reducing train energy consumption and enhancing the overall efficiency of freight train operations.

Author Contributions: Conceptualization, X.S.; methodology, J.L. and W.Z.; formal analysis, J.L.; investigation and data curation, G.S. and X.Z.; writing—original draft, J.L.; writing—review and editing, W.Z.; supervision, X.S.; funding acquisition, H.X. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Key R&D Program of China (2023YFB4302104-1).

Data Availability Statement: The data are available upon request.

Conflicts of Interest: Author Wei Zhang was employed by the company Casco Signal Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

DAS	Driver Advisory System
ROA	Rolling Optimization Algorithm
MPC	Model Predictive Control
ATO	Automatic Train Operation
LQR	Linear Quadratic Regulator
NSGA-II	Non-dominated Sorting Genetic Algorithm-II
FAGA	Fuzzy Adaptive Genetic Algorithm
PID	Proportion Integration Differentiation
LKJ	Train Operation Monitoring Equipment

References

1. National Railway Administration of the People's Republic of China. *Railway Statistics Bulletin 2023*; Beijing, China, 2024.
2. Howlett, P.; Milroy, I.; Pudney, P. Energy-efficient train control. *IFAC Proc. Vol.* **1993**, *26*, 1081–1088. [CrossRef]
3. Howlett, P.; Pudney, P.; Vu, X. Local energy minimization in optimal train control. *Automatica* **2009**, *45*, 2692–2698. [CrossRef]
4. Pudovikov, O.E.; Bespał'Ko, S.V.; Kiselev, M.D.; Serdobintsev, E.V. Application of a reference train model in an automatic control system of freight-train speed. *Russ. Electr. Engin.* **2017**, *88*, 563–567. [CrossRef]
5. Gruber, P.; Bayoumi, M.M. Suboptimal control strategies for multilocomotive powered trains. In Proceedings of the 1980 19th IEEE Conference on Decision and Control including the Symposium on Adaptive Processes, Albuquerque, NM, USA, 10–12 December 1980; pp. 319–327.
6. Chou, M.; Xia, X.; Kayser, C. Modelling and model validation of heavy-haul trains equipped with electronically controlled pneumatic brake systems. *Control Eng. Pract.* **2007**, *15*, 501–509. [CrossRef]
7. Zhuan, X.; Xia, X. Speed regulation with measured output feedback in the control of heavy haul trains. *Automatica* **2008**, *44*, 242–247. [CrossRef]

8. Howlett, P.G.; Cheng, J. Optimal driving strategies for a train on a line with continuously varying gradient. *J. Aust. Math. Soc. Ser. B Appl. Math.* **1997**, *38*, 388–410. [CrossRef]
9. Khmelnitsky, E. On an optimal control problem of train operation. *IEEE Trans. Autom. Control* **2000**, *45*, 1257–1266. [CrossRef]
10. Martinis, V.D.; Weidmann, U.A. Improving energy efficiency for freight trains during operation: The use of simulation. In Proceedings of the 2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC), Florence, Italy, 7–10 June 2016; pp. 1–6.
11. Ko, H.; Koseki, T.; Miyatake, M. Application of dynamic programming to optimization of running profile of a train. *Comput. Railw. IX* **2004**, *74*, 103–112.
12. Albrecht, T.; WGassel, C.; Binder, A.; van Luipen, J. Dealing with operational constraints in energy efficient driving. In Proceedings of the IET Conference on Railway Traction Systems (RTS 2010), Birmingham, UK, 13–15 April 2010; pp. 1–7.
13. Guo, Y.; Qiu, L.; Ma, J.E. Multi-objective optimization of high-speed train running speed trajectory based on particle swarm and NSGA-II fusion algorithm. In Proceedings of the 2022 IEEE Transportation Electrification Conference and Expo, Asia-Pacific (ITEC Asia-Pacific), Haining, China, 28–31 October 2022; pp. 1–5.
14. Yi, L.Z.; Hang, D.K.; Li, W. Research on multi-objective optimization of freight train operation process based on improved bald eagle search algorithm. *J. Comput.* **2022**, *33*, 135–150.
15. Yang, H.; Xu, K.X.; Fu, Y.T. Research on multi-objective optimal control of heavy haul train based on iImproved genetic algorithm. In Proceedings of the 2022 4th International Conference on Industrial Artificial Intelligence (IAI), Shenyang, China, 24–27 August 2022; pp. 1–6.
16. Isna, S.S.; Ari, S. Application of model predictive control on metro train scheduling problems. *Int. J. Veh. Inf. Commun. Syst.* **2023**, *8*, 103–118.
17. Song, Y.X. Research on Train Automatic Operation Control Algorithm for Heavy-Hual Trains Based on Model Predictive Control. Master's Thesis, Southwest Jiaotong University, Chengdu, China, 2022.
18. Liu, Y.; Zhang, Z.F.; Jiang, F. Research on model predictive control method of heavy-haul trains based on multi-point model. In Proceedings of the 2023 IEEE Vehicle Power and Propulsion Conference (VPPC), Milan, Italy, 24–27 October 2023; pp. 1–7.
19. He, X.M.; Wang, S.; Alanamu, B.B. Research on optimal train adhesion control based on nonlinear model predictive control. In Proceedings of the International Conference on Smart Transportation and City Engineering (STCE 2023), Chongqing, China, 16–18 December 2023.
20. Zhang, L.J.; Li, Q.; Zhuan, X.T. Energy-efficient operation of heavy haul trains in an MPC framework. In Proceedings of the 2013 IEEE International Conference on Intelligent Rail Transportation Proceedings, Beijing, China, 30 August–1 September 2013; pp. 105–110.
21. Zhang, L.J.; Zhuan, X.T.; Xia, X.H. Optimal operation of heavy haul trains using model predictive control methodology. In Proceedings of the 2011 IEEE International Conference on Service Operations, Logistics and Informatics, Beijing, China, 10–12 July 2011; pp. 402–407.
22. Zhang, L.J.; Zhuan, X.T. Optimal operation of heavy-haul trains equipped with electronically controlled pneumatic brake systems using model predictive control methodology. *IEEE Trans. Control Syst. Technol.* **2014**, *22*, 13–22. [CrossRef]
23. Zhang, L.J.; Zhuan, X.T. Development of an Optimal Operation Approach in the MPC Framework for Heavy-Haul Trains. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 1391–1400. [CrossRef]
24. Zhang W. Research on Receding Optimization of Energy-Saving Speed Curve of Freight Trains. Master's Thesis, Beijng Jiaotong University, Beijing, China, 2023.
25. Bujarbaruah, M.; Zhang, X.J.; Borrelli, F. Adaptive MPC with Chance Constraints for FIR Systems. In Proceedings of the 2018 Annual American Control Conference (ACC), Milwaukee, WI, USA, 27–29 June 2018; pp. 2312–2317.
26. Zhai, W.M. Research on the dynamics performance evaluation standard for freight trains and the suggested schemes(last part continued)—Evaluation standard for vertical wheel rail force and coupler force. *Roll. Stock* **2002**, *3*, 10–13+1.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.