



UNIVERSITY OF LEEDS

This is a repository copy of *Robust Optimization of Energy-Saving Train Trajectories under Passenger Load Uncertainty Based on p-NSGA-II*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/189547/>

Version: Accepted Version

---

**Article:**

Xing, C, Li, K [orcid.org/0000-0001-6657-0522](https://orcid.org/0000-0001-6657-0522), Zhang, L et al. (1 more author) (2023) Robust Optimization of Energy-Saving Train Trajectories under Passenger Load Uncertainty Based on p-NSGA-II. IEEE Transactions on Transportation Electrification, 9 (1). 1826 -1844. ISSN 2332-7782

<https://doi.org/10.1109/tte.2022.3194698>

---

© 2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# Robust Optimization of Energy-Saving Train Trajectories under Passenger Load Uncertainty Based on p-NSGA-II

Chen Xing, Kang Li, *Senior member, IEEE*, Li Zhang, *Senior member, IEEE*, and Zhongbei Tian, *Member, IEEE*

**Abstract**—Railway electrification has attracted substantial interests in recent years as a key part of the global effort to achieve transport decarbonisation. To improve the energy efficiency of train operations, of particular interest is the optimization of train speed trajectories. However, most studies formulate the problem as a single-objective optimization model and do not take into account train mass uncertainty associated with the passenger load variations. This paper formulates a bi-objective robust optimization model to minimize both the energy consumption and journey time, in which the robustness against the uncertain train mass is considered and viewed as a decision-maker preference. A novel multi-objective optimization algorithm namely p-NSGA-II is proposed, incorporating the original NSGA-II and a proposed preference dominance criterion to handle the DM preference. With the proposed p-NSGA-II, not only all solutions will converge to the optimal Pareto front but also solutions with better robustness in the Pareto front will be automatically selected and retained, meanwhile the spread of the selected solutions is maintained. The effectiveness of the p-NSGA-II to generate a set of performance-robust driving schemes is verified by numerical case studies. The results show that the p-NSGA-II can achieve up to 40.59% robustness improvement compared to the original NSGA-II.

**Index Terms**—Train energy-saving speed trajectories, robust multi-objective optimization, train load uncertainty, p-NSGA-II.

## NOMENCLATURE

### A. Abbreviations(Alphabetically)

|                  |   |
|------------------|---|
| <i>ADP</i>       | Approximate dynamic programming             |
| <i>ATO</i>       | Automatic train operation                   |
| <i>DM</i>        | Decision maker                              |
| <i>GA</i>        | Genetic algorithm                           |
| <i>HS</i>        | High speed                                  |
| <i>LS</i>        | Low speed                                   |
| <i>MILP</i>      | Mixed-integer linear programming            |
| <i>MOEA</i>      | Multi-objective evolutionary algorithm      |
| <i>MOPSO</i>     | Multi-objective particle swarm optimization |
| <i>MS</i>        | Medium speed                                |
| <i>NSGA – II</i> | Nondominated sorting genetic algorithm-II   |
| <i>PMP</i>       | Pontryagin’s maximum principle              |
| <i>PR</i>        | Performance robustness                      |

|            |                         |
|------------|-------------------------|
| <i>RO</i>  | Robust optimization     |
| <i>ROI</i> | Region of interest      |
| <i>SO</i>  | Stochastic optimization |
| <i>SR</i>  | Scheme robustness       |

### B. Parameters and Variables (Alphabetically)

|                         |  |
|-------------------------|--|
| $\alpha, \beta, \gamma$ | Empirical coefficients of basic resistance |
| $\bar{M}$               | Average train mass                         |
| $\bar{M}_p$             | Average passenger mass                     |
| $\theta$                | Ramp angle                                 |
| $a$                     | Acceleration rate                          |
| $E$                     | Energy consumption                         |
| $F$                     | Traction/braking force                     |
| $F_{max}$               | Maximum traction/braking force             |
| $M$                     | Train mass                                 |
| $M_p$                   | Passenger mass                             |
| $M_v$                   | Empty vehicle mass                         |
| $P_{max}$               | Rated motor power                          |
| $R_1$                   | Basic resistance                           |
| $r_1$                   | Per unit basic resistance                  |
| $R_2$                   | Additional resistance                      |
| $T$                     | Journey time                               |
| $V$                     | Speed control sequence of whole journey    |
| $v_{i,b}$               | Braking speed of subsection $i$            |
| $v_{i,c}$               | Cruising speed of subsection $i$           |
| $v_{i,e}$               | End speed of subsection $i$                |
| $v_{i,s}$               | Start speed of subsection $i$              |
| $V_i$                   | Speed control sequence of subsection $i$   |

## I. INTRODUCTION

**R**AILWAY electrification has attracted substantial interests in recent years as a key part of the global effort to achieve transport decarbonisation. Given the changes in the energy mix landscape and the challenging decarbonisation targets to which a number of governments have committed, it is vital to improve the energy efficiency in all sectors, including train operations [1], [2]. High-frequency, high-density train schedules often lead to dramatically increased energy consumption which is becoming a major concern for railway companies. Many measures have been taken to reduce energy consumption, such as rescheduling of operating timetables [3]–[7], installation of energy storage systems [8]–[14], and reducing vehicle resistance. Among these, optimization of the train speed trajectory is seen as an effective and economical approach that requires no no additional investment in infrastructure and equipment.

This work was supported in part by the Strategic Innovation Fund (SIF). (Corresponding author: Kang Li.)  
 Chen Xing: elcx@leeds.ac.uk  
 Kang Li: K.Li1@leeds.ac.uk  
 Li Zhang: L.Zhang@leeds.ac.uk  
 Zhongbei Tian: Zhongbei.Tian@liverpool.ac.uk

It is also the most cost-effective approach to address the uncertainties at the train planning and operation stage for given timetable and power supply capacity.

Much effort so far has been put into developing energy-saving operating schemes for trains. In earlier work, the optimization of the train speed trajectory is usually formulated as a mixed-integer linear programming (MILP) problem for minimum energy consumption [14]–[17], while some approximations using linearization are required. With later advances in computer performance, more studies were focused on developing meta-heuristic algorithm based optimization models [18]–[25]. In [26], [27], multiple-phase trajectory controls are proposed and additional practical factors have been incorporated in the optimization models, such as passenger comfort and punctuality. In summary, studies on single-objective optimization of a train speed trajectory are mature.

However, single-objective optimization may suffer from a few limitations and fail to meet the requirements where multiple conflicting objectives exist simultaneously. In a modern railway system, an automatic train operation (ATO) system is widely deployed to achieve energy-saving driving [28], [29], and several popular driving patterns have been proposed, such as the cruising pattern and coasting-remotoring pattern [30]–[32]. Generally speaking, the driving trajectories need to be pre-programmed into the ATO system based on the required arrival time. Thus, to meet different requirements on the journey time (timetable), it is preferred to provide in advance a set of driving profiles with different arrival time, rather than only one option. However, single-objective optimization can only produce a single solution rather than a comprehensive set of solutions. To tackle this drawback, the train speed trajectory optimization has been formulated as a bi-objective problem, where the objectives include energy consumption and running time. In this way, a set of Pareto-optimal driving schemes characterized by various arrival times can be produced. For example, a bi-objective speed trajectory optimization model is proposed in [33]. A multi-objective particle swarm optimization (MOPSO) is developed and applied to obtain a set of Pareto-optimal speed profiles. This type of bi-objective optimization model is also adopted in [34], [35] and [32]. The bi-objective optimization models minimizing both the energy consumption and journey time can effectively handle the speed trajectory planning problem, fit well the existing onboard ATO system and meet various operational requirements.

In the aforementioned studies, the optimization problems are formulated as deterministic models and related parameters are assumed to be constant. In reality, operation of a railway system is usually affected by uncertain factors which may greatly disturb the operation or even lead to worse consequences. In the context of engineering applications, various methods have been proposed to handle the uncertainties to mitigate their impacts [36]–[38]. However, few studies on railway operation optimization have considered the uncertainties. [39] considers the uncertainty in traction force and train resistance. The optimization for the speed trajectory was formulated as a Markov decision process and solved by an approximate dynamic programming (ADP) based method. However, the train mass, as a significant source of uncertainties, was not considered.

So far, the train mass uncertainty associated with complex passenger distributions is still seen as the greatest challenge for a speed trajectory optimization. As pointed out in [32], [40], the maximum passenger mass is around 30% of the total train mass which can result in up to 35% energy increment and 8% running time increment. Therefore it is vital to investigate the robustness enhancement to mitigate the negative effects of the random passenger load. In [41], the authors adopted a method combining an expert system and reinforcement learning to handle the railway system uncertainties such as delays. While it needs a vast amount of operational data to develop the expert rules. In [40], a stochastic programming model for metro timetable rescheduling to reduce the time delay caused by the passenger number uncertainty is developed. However the speed trajectory itself is not optimized. In [42], a two-phase stochastic programming model integrating timetable and speed profile optimization is formulated, considering the train mass uncertainty. However, the stochastic optimization model requires prior information which may result in a high computational cost for the information sampling. Meanwhile, in these studies, the optimization problems are formulated as single-objective programming models which do not meet the requirements of ATO system.

As discussed earlier, several factors, such as energy consumption, journey time and robustness which are affected by the uncertainty of the passenger load, need to be considered simultaneously in the train speed trajectory optimization. Meanwhile, most real-life multi-objective problems suffer from having a large solution space, which may result in a slow and time-consuming optimization process, even non-convergence may occur. One of the measures is to incorporate decision maker's (DM) preferences into the optimization process to reduce the solution space. For example, in [43], a new DM preference-based multi-objective evolutionary optimization algorithm (MOEA) utilizing configured weight intervals is proposed for the ship weather routing. In [44], a novel MOEA incorporating the weight aggregation strategy as the priori DM's preference, is developed to optimize the load scheduling of electric vehicles. In [45], a scoring and dynamic hierarchy-based NSGA-II is designed for workflow scheduling, where a scoring criterion is considered as the DM preference. Besides, [46]–[51] also introduce some DM preference-based MOEAs. However, these methods are developed for specific problems and not suitable for train speed trajectory optimization problem elaborated earlier. In [32], the speed profile optimization is formulated as a bi-objective optimization problem and optimized by MOPSO. The trainload variations are considered in formulating the optimization constraints which can be viewed as using the DM preference to guide the search. However, the formulated optimization model involves a pre-defined sensitivity parameter that needs to be set in advance which will significantly affect the final results. Further, the adopted method in [32] may distort the distribution of the optimal Pareto front. Finally, the cruising operation is not considered in the driving patterns, which may lead to suboptimal and less energy-efficient solutions.

This paper aims to address the aforementioned challenges and to develop robust energy-saving speed profiles considering

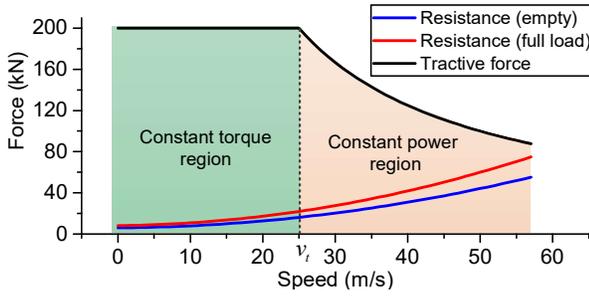


Fig. 1. Operation characteristic of British rail Class 390.

the uncertainty of train mass. The main contributions are summarized as follows:

1) A bi-objective train speed trajectory optimization problem is formulated where the robustness affected by the uncertain train mass is considered. The robustness of each driving scheme is formulated as the maximum variations of running time and energy consumption due to uncertain passenger loads.

2) To derive a set of optimal driving schemes suitable for a realistic train journey, a train movement control model is adopted which divides the whole journey into multiple subsections based on different speed limits and track gradients. The speed trajectory of each subsection is optimized separately, which helps to explore the flexibility and superiority of different combinations of train operations.

3) A new DM preference-based multi-objective algorithm, namely p-NSGA-II, is proposed. The p-NSGA-II can automatically guide the search to the optimal Pareto front and further to the region of interest (ROI) based on DM preference. Meanwhile, the distribution of the obtained optimal Pareto front can be maintained.

4) The robustness of the train speed trajectories is modeled as the DM preference. By adopting p-NSGA-II, a set of driving schemes, which optimize both energy consumption and journey time, are produced with better robustness. The effectiveness of the formulated optimization model and p-NSGA-II is verified based on extensive case studies.

The remainder of the paper is organized as follows. Section II develops the bi-objective optimization model considering the uncertainty of train mass. In Section III, the new DM preference-based algorithm p-NSGA-II is proposed and its implementation procedures are detailed. In Section IV, the effectiveness of the p-NSGA-II based robust optimization of the speed trajectory is verified through two case studies. The performance of the proposed p-NSGA-II is compared with other algorithms including the original NSGA-II and a robust multi-objective optimization method. Finally, Section V concludes the paper.

## II. MODEL FORMULATION

### A. Train Kinetic Model

Based on the fundamental Newtonian equations of train motion, the train kinetic model can be described as (1), known

as Lomonosoff's equation [35], which has been verified in earlier research.

$$\begin{cases} Ma = F - R_1 - R_2 \\ a = \frac{dv}{dt} \end{cases} \quad (1)$$

where the train mass  $M$  consists of the empty vehicle mass  $M_v$  and uncertain mass of passengers  $M_p$ , i.e.  $M = M_v + M_p$ . The vehicle mass is seen as a deterministic parameter. While the passengers mass is random at different stations and during different operational periods.  $a$  is the train's acceleration rate.  $F$  represents the tractive force or braking force.  $R_1$  and  $R_2$  are the basic resistance and additional resistance of the train respectively, as will be detailed shortly.  $v$  is the train velocity.

In Fig. 1, the black line shows the motor characteristic curve of the British rail Class 390, a typical electric high-speed passenger train [52]. Whether the train is in the state of traction or braking, it is always governed by this characteristic curve. The whole curve can be split into two regions, with constant torque region or constant power. When the train speed is lower than  $v_t$ , the train is operated in the constant torque region, and at higher speed runs in the constant power region. A piecewise function (2) can represent the motor characteristic curve.

$$F = \begin{cases} uF_{\max} & 0 < v \leq v_t \\ u \frac{P_{\max}}{v} & v_t < v \leq v_{\max} \end{cases} \quad (2)$$

where  $F_{\max}$  represents the maximum traction/braking force.  $P_{\max}$  is the rated motor power.  $u \in [-1, 1]$  controls the train's operation command.  $F$  is the tractive force when  $u$  is positive, and braking force when  $u$  is negative.

1) *Basic Resistance*: The running train is resisted by the track friction and air force. Based on the Davis equation [35], this kind of resistance can be expressed as (3), containing the empirical coefficients  $\alpha$ ,  $\beta$  and  $\gamma$ , the train mass  $M$  and speed  $v$ .  $r_1$  represents per unit basic resistance. The basic resistance will not exceed the traction force. The curves of basic resistance versus train velocity are also shown in Fig. 1, for both empty-vehicle and full-load conditions.

$$R_1(v) = r_1(v) Mg = (\alpha + \beta v + \gamma v^2) Mg \quad (3)$$

2) *Additional Resistance*: The additional resistance depends on the track geography and is mainly composed of ramp resistance, curve resistance and tunnel resistance. In most research, only ramp resistance is considered in (4) since it dominates the additional resistance.

$$R_2 = Mg \sin(\theta) \quad (4)$$

where  $\theta$  represents the ramp angle. In addition, the train movement can be expressed as:

$$s(t) = \int_0^t v(t) dt \quad (5)$$

where  $s$  is the running distance of the train.

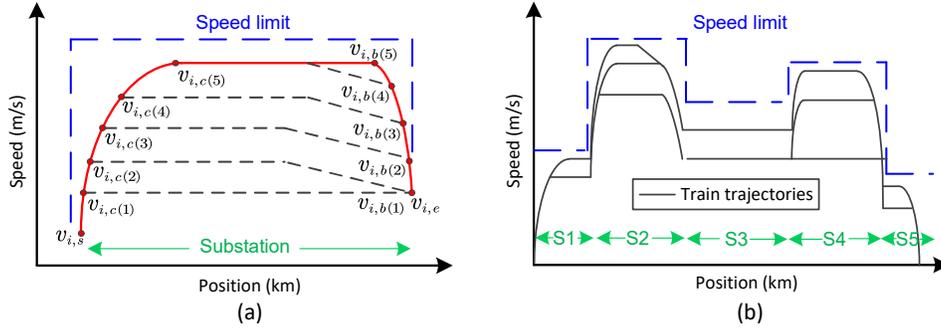


Fig. 2. Train Speed trajectory control. (a) Speed trajectory control of subsection; (b) Speed trajectories control of whole journey.

### B. Train Driving Pattern and Speed Trajectory Control

In order to construct the optimal and practically feasible train speed trajectories, a method introduced in [34] is adopted. The whole journey between two consecutive stations is divided into  $n$  subsections based on different speed limitations as shown in Fig. 2. For each subsection, several popular driving patterns have been proposed in the literature, such as the cruising pattern and coasting-remotoring pattern [30]–[32]. In particular, one driving strategy, which considers all possible operations including acceleration, cruising, coasting and braking, has been comprehensively adopted in the ATO system [4], [21], [34], [35], and it helps to explore the flexibility and superiority of different combinations of train operations. As stated in [27], [35], based on Pontryagin's maximum principle (PMP), if the train is aimed to reach the destination as quickly as possible, the train is supposed to reach the maximum allowable speed first with full acceleration and keep the maximum speed (cruising) until the final braking point. Then the full braking is performed to stop the train. More energy-saving driving schemes can be achieved if considering coasting operation and reducing cruising speed. In the following, the combinations of four operations including full acceleration, cruising, coasting and full braking are considered and the driving pattern is introduced. As shown in Fig. 2(a), in each subsection, four control variables ( $v_s, v_c, v_b, v_e$ ) are defined to determine a specific sectional speed trajectory.  $v_s, v_c, v_b$  and  $v_e$  represent the start speed, cruising speed, braking speed and end speed of each subsection respectively. The speed trajectory of each subsection can be depicted as: if  $v_c > v_s$ , the train is operated in full acceleration mode until the train reaches to the cruising speed  $v_c$  and then the cruising operation follows. If  $v_c = v_s$ , the train is operated in the cruising mode directly. The cruising time depends on the running distance of each subsection. After cruising, the coasting operation follows immediately until the train reaches the set braking speed if  $v_b > v_e$ . Finally, the train brakes with maximum braking force ( $u = -1$  in (2)) until the train reaches the end speed at the end point of the subsection. If  $v_b = v_e$ , the train is operated in the coasting mode until it reaches the end speed of the subsection. Theoretically, based on the above description, once the values for  $v_s, v_c, v_b$  and  $v_e$  are chosen in advance, the positions of cruising, coasting and braking points can be calculated, and the speed trajectory of each subsection can be plotted, the

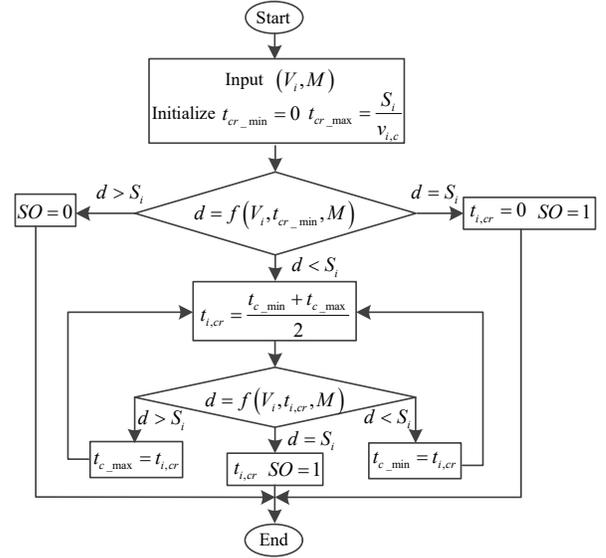


Fig. 3. Flowchart of SO evaluation and  $t_{i,cr}$  calculation.

credibility of which has been verified in [32], [34], [35]. Due to the operational constraints, some rules need to be clarified as follows:

- 1) All the values of  $v_s, v_c, v_b$  and  $v_e$  should be lower than the speed limit in each subsection.
- 2)  $v_c < v_s$  and  $v_b < v_e$  are disabled.
- 3) For two successive subsections, the end speed of the former subsection should be equal to the start speed of the later subsection, i.e.  $v_{i,e} = v_{i+1,s}$ .

For the whole journey including  $n$  subsections, the control variables can be expressed as:  $V = (v_{1,s}, v_{1,c}, v_{1,b}, v_{1,e}, \dots, v_{i,s}, \dots, v_{i,e}, \dots, v_{n,s}, v_{n,c}, v_{n,b}, v_{n,e})$  where  $v_{1,s} = v_{n,e} = 0$  means the initial and final speeds of the journey are equal to zero. Thus there are  $3n - 1$  control variables in total which need to be optimized to form the whole train speed trajectory as shown in Fig. 2(b).

### C. Robust Optimization Model of Train speed Trajectory

1) *Original Train Speed Trajectory Optimization Model without Considering Robustness*: As discussed above, once the set of speed control variables  $V$  is confirmed, the whole speed trajectory can be formulated for a specific train mass.

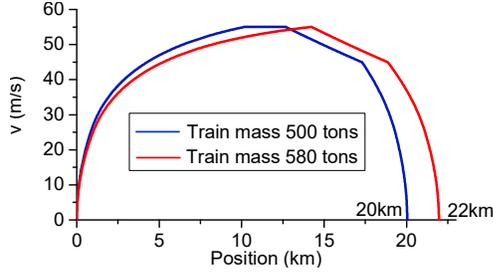


Fig. 4. Scheme invalidity caused by passengers' randomness.

The associated total energy consumption and running time can be expressed as (6) and (7) respectively.

$$E = f_1(V, M) = \sum_{i=1}^n \left( \int_0^{T_i} \max\{0, F(t)\} v(t) dt - \eta \int_0^{T_i} \min\{F(t), 0\} v(t) dt \right) \quad (6)$$

$$T = f_2(V, M) = \sum_{i=1}^n (t_{i,a} + t_{i,cr} + t_{i,co} + t_{i,b}) \quad (7)$$

where  $t_{i,a}$ ,  $t_{i,cr}$ ,  $t_{i,co}$  and  $t_{i,b}$  represent the durations of the motoring, cruising, coasting and braking of each subsection.  $\eta$  represents the energy efficiency of the regenerative braking. Generally speaking, the energy consumption and running time also vary with different driving schemes, even for the same journey distance. The optimal speed trajectory should be a trade-off between the energy consumption and running time. Thus the original train trajectory optimization can be formulated as a bi-objective optimization problem in (8) where the train mass is set to the mean value  $\bar{M}$ , which is the sum of the empty vehicle mass and mean mass of the passengers, i.e.  $\bar{M} = M_v + \bar{M}_p$ . In the formulated optimization model, both objectives of minimizing the energy consumption and running time are equally important.

$$\min \begin{cases} E = f_1(V, \bar{M}) \\ T = f_2(V, \bar{M}) \end{cases} \quad (8)$$

In real operation, the train is required to arrive at the destination at time  $T$  and should be subject to speed and acceleration limits of each subsection and the driving rules stated in Section II.B. The corresponding constraints are shown in (9).

$$s.t. \begin{cases} v_{1,s} = v_{n,e} = 0 \\ v_{i,e} = v_{i+1,s} \quad (i = 1, \dots, n-1) \\ s(0) = 0; s(T) = S \\ v_{i,s}, v_{i,c}, v_{i,b} \text{ and } v_{i,e} \leq v_{i,\max} \quad (i = 1, \dots, n) \\ v_{i,s} \leq v_{i,c}; v_{i,e} \leq v_{i,b} \quad (i = 1, \dots, n) \\ |a_i| \leq a_{i,\max} \quad (i = 1, \dots, n) \end{cases} \quad (9)$$

2) *Scheme Robustness (SR)*: Previous research results have suggested that the main uncertain factor affecting the train operation is the passenger load dependent train mass. One driving scheme with a certain set of speed control variable  $V$  may be valid for the specific train mass values but not for others. The scheme robustness (SR) in this paper proposed is used to guarantee the validity of the chosen scheme  $V$  even

under the worst case. For a specific subsection with a particular set of driving control variables  $V_i = (v_{i,s}, v_{i,c}, v_{i,b}, v_{i,e})$ , the corresponding acceleration, coasting and braking distances can be confirmed based on the train model (1)-(5). While before producing a sectional train trajectory, an intermediate variable, the cruising time  $t_{i,cr}$  needs to be calculated to determine the cruising distance  $d_{i,c} = t_{i,cr} \cdot v_{i,c}$  accordingly. An iterative method is adopted to calculate the cruising time as shown in the flowchart in Fig. 3. When a set of driving control variables of a subsection with a particular train mass value  $M$  is input, initially the cruising time  $t_{i,cr}$  is assumed as 0, and the corresponding running distance  $d$  can be calculated. If  $d = S_i$ , where  $S_i$  is the distance of the subsection, it implies that the train can reach the end point of the subsection under  $(V_i, M)$  with  $t_{i,cr} = 0$ , hence  $SR=1$ . While  $d < S_i$  implies that a valid cruising time value  $t_{i,cr} > 0$  must exist. Then the binary search (or half-interval search) is used to find the cruising time value. Conversely,  $d > S_i$  implies that a valid cruising time value can not be found under  $(V_i, M)$ . Consequently, the scheme  $(V_i)$  is invalid under train mass  $M$ , hence  $SR=0$ . The bi-objective optimization problem in (8)-(9) is solved under the mean mass value  $\bar{M}$ . The worst case – the case most likely resulting in the scheme invalidity must occur at either the minimum  $M_{\min}$  (empty vehicle mass) or the maximum  $M_{\max}$  (full-load train mass). Evaluation of SR, for each selected driving scheme  $(V_i)$  should be performed by applying either  $M_{\min}$  or  $M_{\max}$  instead of average  $\bar{M}$ . If the results satisfy Equation (10), which is a logical constraint added in the original optimization model (8), the chosen driving scheme passes the scheme robust requirement. Taking an example as shown in Fig. 4 with initial speed and end speed equal to 0, i.e.  $v_{i,s} = v_{i,e} = 0$ , the chosen scheme is valid with train mass of 500 tons. However, when the train mass increases to 580 tons, even if the cruising time is 0, the running distance is far longer than the sectional distance 20 km. In this case, the scheme  $(0, 55 \text{ m/s}, 45 \text{ m/s}, 0)$  is regarded as a non-scheme-robust scheme. i.e.  $SR(0, 55, 45, 0, 580) = 0$ . It should be noted that the driving scheme of each subsection needs to be checked to guarantee the validity under the extreme train mass values, i.e.  $SR(V_i, M_{\min} | M_{\max}) = 1 \quad (i = 1, \dots, n)$ , so that the whole train speed trajectory  $V$  is valid.

$$SR(V_i, M_{\min}) = 1 \text{ and } SR(V_i, M_{\max}) = 1 \quad (10)$$

$$SR(\cdot) = \begin{cases} 0 & \text{invalid scheme} \\ 1 & \text{valid scheme} \end{cases} \quad (11)$$

3) *Performance Robustness (PR)*: Scheme robustness, which implies the feasibility of the chosen scheme even under the worst train mass case, can be distinguished by the performance robustness. The energy consumption and running time of the chosen driving scheme should be affected as little as possible due to the train mass variation. This is referred to as performance robustness (PR) of the driving scheme and it is evaluated by the scheme sensitivity, which can be formulated as (12). Obviously, a driving scheme with low sensitivity to the variation of the passenger load is preferred and also means

the energy consumption and running time in the whole journey will not deviate far from their expected values.

$$PR = \max \left( \lambda_1 \cdot \frac{|t - t_{\min}|}{t} + \lambda_2 \cdot \frac{|E - E_{\min}|}{E}, \lambda_1 \cdot \frac{|t - t_{\max}|}{t} + \lambda_2 \cdot \frac{|E - E_{\max}|}{E} \right) \quad (12)$$

where  $\lambda_1$  and  $\lambda_2$  are the weighting factors.  $\lambda_1, \lambda_2 \in [0, 1]$  and  $\lambda_1 + \lambda_2 = 1$ . And

$$\begin{cases} t_{\min} = T(V, M_{\min}); E_{\min} = E(V, M_{\min}) \\ t_{\max} = T(V, M_{\max}); E_{\max} = E(V, M_{\max}) \\ t = T(V, \bar{M}); E = E(V, \bar{M}) \end{cases} \quad (13)$$

Compared with the energy consumption and the running time, the performance robustness is often considered to be less important and hence it can be regarded as a subordinate objective. The main goal is still to find the scheme-robust Pareto solutions from the two-dimensional (2-D) plane of energy consumption and running time. Then the Pareto solutions are further filtered considering the performance robustness, meanwhile maintaining the Pareto spread and evenness.

In summary, the robust optimization problem for the train speed trajectories can be described as: Firstly without considering the uncertainty of the train mass, the original optimization problem is formulated as (8) with constraints (9). Then the scheme robustness (SR) is taken as the constraint (10) to filter out the driving schemes that cannot fit all train mass values. Finally, the performance robustness (PR) in (12) is considered in the proposed p-NSGA-II introduced in Section III to guide the search to select the driving schemes that are affected as little as possible by the train mass uncertainty. Notably, for a specific driving strategy, both the SR and PR are two different robustness indicators which are used to evaluate the performance of solutions at different stages of the optimization.

### III. OPTIMIZATION ALGORITHM – p-NSGA-II

In the robust optimization model of the speed trajectories, the performance robustness has a lower priority than energy consumption and running time. In this case, the performance robustness criterion can be seen as a type of DM preference. Thus the whole problem becomes a DM preference-based bi-objective optimization. The DM preference is specified as aspirations, leading the search to gradually converge towards the DM region of interest (ROI). While existing preference-based approaches mostly rely on a reference point or a reference direction [46], [47], [53], these are not adopted in the optimization problem proposed here. In this section, a new DM preference-based optimization approach is proposed, which is extended by original NSGA-II and named as p-NSGA-II. The proposed p-NSGA-II is a type of interactive DM approaches. In an interactive decision-making approach, the DM preferences are incorporated progressively during the optimization process. The p-dominance criterion incorporating the DM preference is involved in each iteration. This can reduce the searching space and enable a DM effectively guiding the search towards regions of interest and away from

exploring non-interesting solutions. The DM preference in the proposed optimization procedure is based on performance robustness and is manifested through a modification of the dominance relation.

#### A. Definition of p-Dominance

Most MOEAs are based on the nondomination relation. In this section, the nondomination relation is briefly introduced and a new dominance relation - preference-dominance (p-dominance) for dealing with DM preference is developed from this. This combines the original nondomination principle and DM preference. The main characteristic of p-dominance is that solutions located at the Pareto front can be further selected based on the DM preference. Given a general form of the multi-objective optimization problem:

$$\min (f_1(x), f_2(x), \dots, f_K(x)) \quad \text{s.t. } x \in X \quad (14)$$

where  $x$  is the vector of the decision variables,  $X$  represents the set of feasible solutions, and  $f_k(x)$  ( $k = 1, 2, \dots, K$ ) are the various objective functions. It is assumed that every feasible solution  $x_p$  has three attributes:

- 1) *Nondomination rank* ( $x_p^{n-rank}$ )
- 2) *Preference rank* ( $x_p^{p-rank}$ )
- 3) *Crowding distance* ( $x_p^{cd}$ )

The nondomination relation can be described as: for two feasible solutions  $x_p$  and  $x_q$ , if

$$f_k(x_p) \leq f_k(x_q) \quad \forall k = 1, 2, \dots, K \quad (15)$$

with existing at least one integer in  $k$  satisfying  $f_k(x_p) < f_k(x_q)$ , then  $x_p$  dominates  $x_q$  ( $x_p$  has a lower nondomination rank than  $x_q$ ). In other words, if a solution is not dominated by any other solutions, it is located at the first nondomination rank. The set of solutions in the first nondomination rank constitutes the Pareto front. Based on the above principle, the nondomination ranks  $x^{n-rank}$  of all solutions can be sorted.

The attribute of crowding distance, which has been introduced in [54], reflects the density of solutions surrounding the particular proposed solution  $x_p$ . To obtain an estimate of the density of solutions around a particular solution in the population, the average distance of two points on either side of this point along each of the objectives is calculated by (16).

$$x_p^{cd_n} = \frac{f_k(x_{p+1}) - f_k(x_{p-1})}{f_k^{\max} - f_k^{\min}}, \quad x_p^{cd} = \sum_{k=1}^K x_p^{cd_n} \quad (16)$$

where  $x_{p+1}$  and  $x_{p-1}$  are adjacent solutions to  $x_p$ .  $f_k^{\max}$  and  $f_k^{\min}$  represent the maximum and minimum values of  $k$ -th objective respectively. The solutions with a larger crowding distance are preferred, and it also implies that these solutions are in a less crowded region. The approach for evaluating the preference ranks will be illustrated in Section III.B.

Based on the above three attributes, the Preference-Dominance Relation is defined as:

**Definition (Preference – Dominance Relation):** Given two solutions  $x_p$  and  $x_q$ , if one of the following statements is satisfied,  $x_q$  is  $p$ -dominated by  $x_p$ .

- 1) if ( $x_p^{n-rank} < x_q^{n-rank}$ )

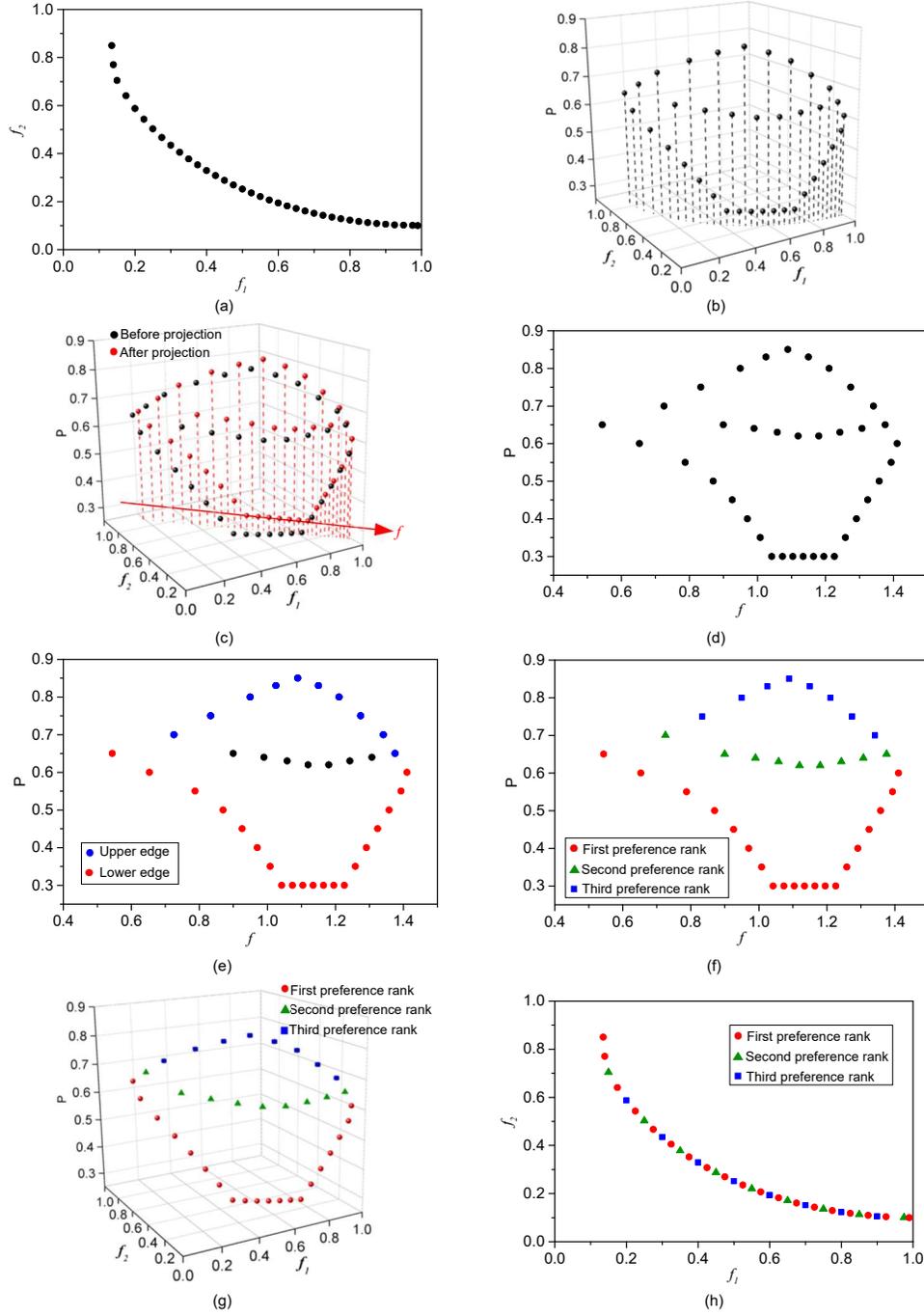


Fig. 5. Procedures of preference ranking.

- 2) or  $(x_p^{n-rank} = x_q^{n-rank}$  and  $x_p^{p-rank} < x_q^{p-rank}$ )
- 3) or  $(x_p^{n-rank} = x_q^{n-rank}$  and  $x_p^{p-rank} = x_q^{p-rank}$  and  $x_p^{cd} > x_q^{cd}$ )

That means, for two solutions with different nondomination ranks, the one with the lower rank is preferred. Then if two solutions are located with the same nondomination rank, the one with a lower preference rank is preferred. Finally the crowding distance attribute plays a minimum role in the p-dominance. If and only if both solutions are located in the same nondomination and preference ranks, the crowding distance will determine the ultimate p-dominance relationship

between both solutions.

### B. Preference Ranking

Different from the optimization objectives, the DM preference usually functions as a filter to guide the solutions towards the region of interest (ROI). Generally speaking, the expected ROI is a part of the optimal Pareto front of the original optimization problem. In some cases, the DM preference is defined by the distance between the solutions and reference points. At times, the preference is described as a convergence reference direction. They evolve accordingly to reference-point

and reference-direction based DM algorithms. However, more generally, the DM preference is evaluated by defining the preference function based on a specific optimization problem. For example, the preference function for the problem of train speed trajectory optimization is the performance robustness expressed by (12).

It is assumed, in this section, that the preference function is predefined and the corresponding values for all solutions are evaluated in advance, so the approach to sort out their corresponding preference ranks is described below.

Taking  $p(x)$  as the preference function for a bi-objective optimization problem defined by (14) with  $i=1,2$ , the DM preference is specified to minimize this function. For evaluating the preference rank of each solution, one can assume that there is a set of nondominated solutions  $x_1, x_2, x_3, \dots, x_P$  located in the same front, their corresponding objective function values constitute the set of points in the two-dimensional (2-D) solution plane, as shown in Fig. 5(a). Furthermore, with their preference function values evaluated by  $p(x)$ , the original set of 2-D points is extended to the points cloud in three-dimensional (3-D) space as illustrated in Fig. 5(b). Note that in this graph x-axis and y-axis represent both objectives  $f_1(x)$  and  $f_2(x)$ , and z-axis represents the preference function  $p(x)$ .

The proposed method for deriving the preference rank for each solution can be divided into three steps – projection of solutions, outline points identification and preference ranking.

1) *Projection*: The aim of this preliminary step is to project all the solution points in the 3-D space plot in Fig. 5(b) onto a new 2-D plane, hence facilitating rank sorting. Since the corresponding coordinate of a solution  $x_p$  in Fig. 5(b) is  $(f_1(x_p), f_2(x_p), p(x_p))$ , one can define a new axis,  $f(x)$ -axis, which collapses axes  $f_1(x)$  and  $f_2(x)$ , and then project each solution point onto the new 2-D plane defined by  $(f(x), p(x))$ . The value of each solution  $x_p$  in  $f(x)$ -axis can be calculated by:

$$f(x_p) = \sqrt{(f_1(x_p))^2 + (f_2^{\max}(x) + f_2^{\min}(x) - f_2(x_p))^2} \quad (17)$$

where  $f_2^{\max}(x)$  and  $f_2^{\min}(x)$  represent, respectively, the maximum and minimum values of  $f_2(x)$ , as  $x$  ranges over the set of all nondominated solutions. In this way, all solutions are mapped to the  $f(x)$ -axis, which is shown in Fig. 5(c). Also according to the characteristics of nondominated solutions, all solutions do not overlap on the  $f(x)$ -axis and the density of the solutions can be maintained to the utmost extent. As each solution's  $p(x)$  value is retained, all solutions now have their respective  $f(x)$ -axis and  $p(x)$ -axis values and hence can be projected to the 2-D plane with  $f(x)$  and  $p(x)$  axes, as shown in Fig. 5(d).

2) *Outline Points Identification*: This step aims to identify the outline points of all solutions in the new 2-D plane defined by  $f(x)$  and  $p(x)$  axes.  $\alpha$ -shape method is adopted which is widely applied for finding the outline points of a point set [55]. Algorithm 1 ( $\alpha$  shape method) illustrates how to determine whether the solution points are the outline points of the shape.

It is noted that, in Algorithm 1 ( $\alpha$  shape method), if the lengths  $L_{c_1n}$  or  $L_{c_2n}$  ( $n = 1, \dots, N - 2$ ) are larger than  $\alpha$ , that means there are no points contained inside the circle with

---

### Algorithm 1 $\alpha$ shape method

---

- 1: Input  $N$  solution points with their coordinates
  - 2: Set the circle radius  $\alpha$
  - 3: Select a pair of points  $i$  and  $j$  from the solution points
  - 4: Calculate the length  $L_{ij}$  between points  $i$  and  $j$
  - 5: IF  $L_{ij} \leq 2\alpha$
  - 6: Construct two circles with same radius  $\alpha$  that pass through points  $i$  and  $j$ , see Fig. 6(b)
  - 7: Calculate the circle center coordinates  $C_1$  and  $C_2$
  - 8: Calculate the lengths  $L_{c_1n}$  and  $L_{c_2n}$  ( $n = 1, \dots, N - 2$ ) between the circle center ( $C_1$  or  $C_2$ ) and other points except for points  $i$  and  $j$
  - 9: IF  $L_{c_1n} > \alpha$  or  $L_{c_2n} > \alpha$  ( $n = 1, \dots, N - 2$ )
  - 10: Points  $i$  and  $j$  are identified as outline points
  - 11: ENDIF
  - 12: ENDIF
  - 13: Return to Line 3 until all point pairs of size  $\frac{N \cdot (N - 1)}{2}$  are checked
  - 14: Output all outline points
  - 15: END
- 

radius  $\alpha$ , whose circle center is  $C_1$  or  $C_2$ . At this moment, the selected pair of points  $i$  and  $j$  are located at the edge of the shape constructed by the discrete points, i.e. outline points.

A simple example is used to explain this method. As shown in Fig. 6(a), there are six points in a 2-D plane. A pair of points  $i$  and  $j$  are selected and labeled in Fig. 6(b). Two circles (blue circle and red circle) with radius  $\alpha$  can be constructed which pass through both points  $i$  and  $j$  as can be seen in Fig. 6(b). Although the red circle contains one other point (red point), there is no other point locating within the blue circle. Thus both points  $i$  and  $j$  are therefore identified as outline points. In the same way, all outline points (blue points in Fig. 6(c)) can be identified, and the shape constituted by the points is shown in Fig. 6(d), while the red point is not the outline point. For this method, the circle radius  $\alpha$  will affect the result of the outline points identification. If  $\alpha$  is too small, the set of points will be split into multiple shapes. If  $\alpha$  is too large, the shape will be rather irregular. In this paper, radius  $\alpha$  is set to the minimum value to ensure that all points are grouped in one region. Then the classification of lower and upper outline points is achieved as follows: when the outline points are identified, it is easy to find the leftmost point based on the value of the  $f$ -axis as shown in Fig. 6(e). The leftmost point is determined as the first lower edge outline point. During the process of outline point identification, the leftmost point with two identified outline points  $p$  and  $q$  can constitute two point pairs  $(L, p)$  and  $(L, q)$  satisfying the conditions of outline point identification. These two points  $p$  and  $q$  are considered as lower and upper edge outline points respectively. Similarly, considering one of the two identified outline points like point  $p$ , besides the point pair  $(L, p)$ , there exists another point  $r$  that can constitute another point pair  $(p, r)$  with point  $p$ . This point  $r$  can be considered located in the same outline with point  $p$ , lower outline or upper outline. The process is repeated until the rightmost point is searched. In this way, the whole set of the outline points is divided into two sets by the leftmost point

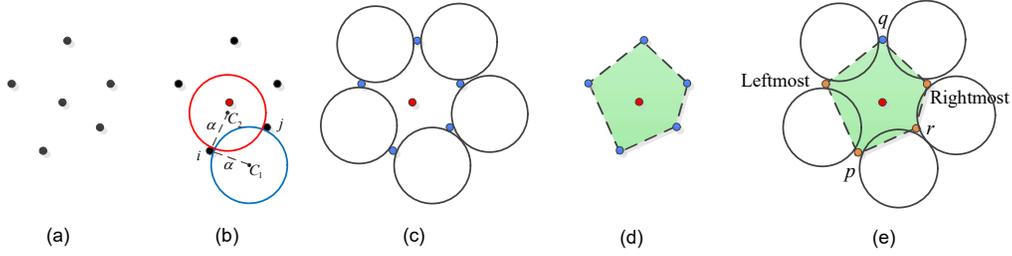


Fig. 6. Outline points identification.

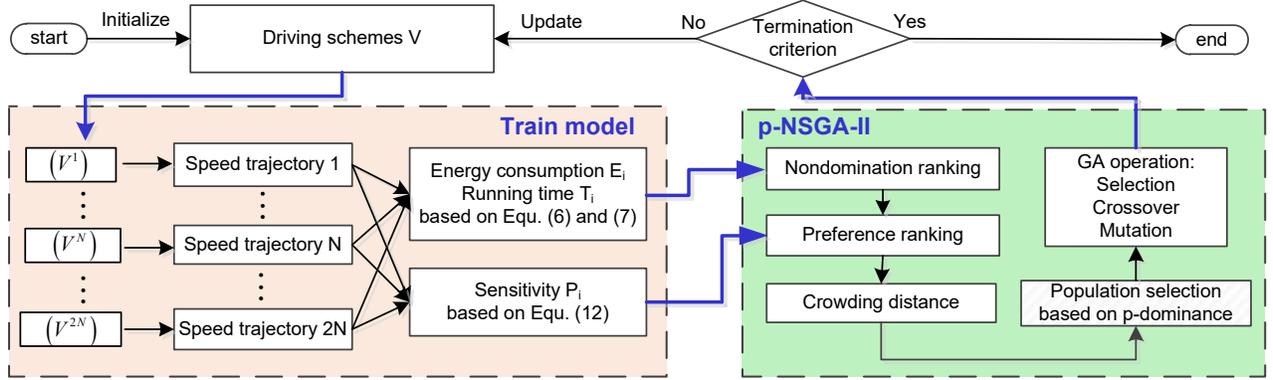


Fig. 7. Framework for searching the robust driving schemes.

and rightmost point. Then, the average  $p$ -axis values of both sets of points are calculated. The point set with smaller  $p$ -axis values can be seen as the lower outline point set labeled in orange in Fig. 6(e).

Applying the above described scheme to all the solution points in Fig. 5(d), the outline points can be identified, and these can then be divided into lower edge outline points (red points) and upper edge outline points (blue points) as illustrated in Fig. 5(e). While the black points are not the outline points.

3) *Ranking*: Since the DM preference aims at minimizing  $p(x)$ , those solution points in Fig. 5(e), possessing lower preference values and constituting the lower edge of the shape, are preferred. These solutions are assigned the first preference rank. For the rest of the solutions, the outline points identification and ranking procedure are repeated until all solutions are sorted into different preference ranks. In Fig. 5(f), the solutions are sorted into three ranks and are remapped into the original 3-D space as shown in Fig. 5(g). Finally, all the original solutions located in the same nondomination front as those illustrated in Fig. 5(a) are sorted into different preference ranks as presented in Fig. 5(h).

### C. Procedure of $p$ -NSGA-II

Incorporating the proposed  $p$ -dominance relation and preference ranking method into the typical multi-objective evolutionary algorithms (MOEAs), a variant of the original NSGA-II is proposed which is called  $p$ -NSGA-II. Its implementation procedure is now described:

- Step 1 GA operation: Initially, a random parent population  $P_0$  of size  $N$  is created. To ensure generality, the

parent population is denoted as  $P_i$  in the subsequent iterations. After applying GA operators on parent population  $P_i$ , including selection, crossover and mutation, the offspring population  $Q_i$  also of size  $N$  is created. Then the parent and offspring populations  $P_i$  and  $Q_i$  are combined to form a new population  $C_i$  of size  $2N$ .

- Step 2 Fitness and preference value evaluation: For each individual in the population  $C_i$ , their corresponding objective functions  $f_1, f_2, \dots, f_k$ , are evaluated together with the preference function  $p(x)$ .
- Step 3 Nondomination rank sorting: Based on the nondomination relation and objective functions evaluated, the nondomination rank of the individuals in a population  $C_i$  can be sorted. The sorting is according to the objective function values of the ones not being dominated by any other individuals hence are the most nondominated and are assigned Rank 1, and the individuals only dominated by individuals in Rank 1 are assigned Rank 2, and so on up to the  $t$ -th rank in the set. Once the sum of solutions from the first rank to  $t$ -th rank exceeds  $N$ , only solutions from first to  $(t-1)$ -th ranks are chosen, while the solutions in  $t$ -th rank should be further sorted. The  $t$ -th rank is called the last nondomination rank.
- Step 4 Preference ranking: For the solutions in the  $t$ -th nondomination rank, they continue to be sorted into different preference ranks using the proposed performance ranking method in Section III.B. Similarly, the solutions from the first to the  $(h-1)$ -th preference ranks are chosen in turn. While the solutions in the last preference rank, i.e. the  $h$ -th preference rank, will be operated in the next step.

- Step 5 Crowding distance sorting: The solutions in the  $t$ -th nondomination rank and  $h$ -th preference rank are sorted using the crowding-comparison operator. The solutions with larger crowding distance are chosen until the chosen solutions from Steps 3-5 fill all slots in a population, forming a new parent population  $P_{i+1}$ .
- Step 6 Stopping criterion: this is defined by the maximum number of iterations.

Steps 3-5 perform the p-dominance procedure. In the whole process, the elitism is ensured. All solutions will be automatically guided towards the Pareto front based on the nondomination relation. Then the solutions located at the Pareto front will continue to converge to the ROI according to the DM preference.

To assess the complexity of the algorithm, we consider one iteration with a population of size  $N$  and the basic operations and their worst-case complexities are as follows:

- 1) nondominated sorting is  $O(KN^2)$
- 2) preference ranking is  $O(N^3)$
- 3) crowding-distance assignment is  $O(KN \log(N))$
- 4) p-dominance sorting is  $O(N \log(N))$

where  $K$  represents the number of objectives. The computational complexities of the nondominated sorting and crowding-distance assignment operations have been discussed in [54]. Please refer to [54] for details. For preference ranking operation, it includes three steps: projection, outline points identification and ranking. The computational complexities of projection and ranking operations are  $O(1)$ . The outline points identification is achieved by the  $\alpha$ -shape method. During this process, as illustrated in Algorithm 1, any two points will constitute a point pair.  $N$  points can constitute  $N(N-1)$  pairs. For any other points except for the selected pair of points, it is necessary to check if they are enclosed by the circles constituted by the selected pair of points. It is obvious that the computational complexity of outline points identification is  $O(N^3)$ . Thus, the overall computational complexity of preference ranking operation is  $O(N^3)$  which is governed by the outline points identification ( $\alpha$ -shape method). While the computational complexity of p-dominance sorting is  $O(N \log(N))$ . By summarizing the above analysis for the different operations in the proposed p-NSGA-II, the overall complexity can be considered as  $O(N^3)$ , which is governed by the preference ranking of the p-NSGA-II. Furthermore, it should be noted that, since the parent population and offspring population are mixed together during the algorithm, the actual size of the population is  $2N$ . Thus the actual computational complexities need to be adjusted correspondingly by replacing  $N$  with  $2N$ .

Combining the optimization model for train speed trajectories as illustrated in Section II, the whole framework for searching the robust driving schemes by the proposed p-NSGA-II is summarized in Fig. 7, and details are given as follows. Step 1: Initialization. Randomly initialize the  $2N$  driving strategies ( $V^1$ ) to ( $V^{2N}$ ) at the beginning of the algorithm and each driving strategy ( $V^i$ ) corresponds to one specific train speed trajectory. Step 2: Apply the  $2N$  driving strategies one by one to the train model (1)-(5) to calculate their energy consumption and running time ( $E_i, t_i$ ) under the mean

value of the train mass using (6) and (7). Besides, under the worst scenario, evaluate the performance sensitivity  $P_i$  of each driving strategy based on the maximum energy and running time deviations by (12). Step 3: Verify the scheme robustness (SR) of each driving strategy and remove non-robust solutions. Step 4: For the remaining driving strategies, determine the nondomination rank and calculate the crowding distance with the energy consumption and running time ( $E_i, t_i$ ) based on Section III.A. Taking performance sensitivity  $P_i$  as the DM preference in the proposed p-NSGA-II, conduct the preference ranking considering the energy consumption, running time and also sensitivity ( $E_i, t_i, P_i$ ) based on Section III.B. Step 5: Given the nondomination rank, preference rank and crowding distance of each driving strategy, select  $N$  preferred driving strategies from the initial  $2N$  driving strategies based on the principle of p-dominance defined in Section III.A. Step 6: Produce  $N$  new offspring driving strategies using the genetic algorithm (GA) operations such as crossover and mutation. Step 7: Mix the parent set and newly generated offspring set by Step 6 to generate the new  $2N$  driving strategies. Step 8: Return to Step 1 to repeat the procedures until the termination condition is satisfied. Notably, the railway decision maker (dispatcher) can set the preference according to (12) and the proposed p-NSGA-II is able to find the DM preference-based driving strategies accordingly.

#### IV. CASE STUDY AND DISCUSSIONS

To verify the effectiveness of the proposed optimization method, two case studies are performed using MATLAB R2021a on a PC with 2.80GHz Intel i7-7700HQ CPU and 16GB RAM. Both use British Rail Class 390 as the example, which has the characteristic curve as shown in Fig. 1. The related parameters are listed in Table I [52]. Case 1, as a baseline case, only considers the basic resistance to verify the proposed p-NSGA-II. Case 2 considers basic resistance, additional resistance caused by the track gradients as well as various speed limits. The case studies confirm the effectiveness of the proposed p-NSGA-II in searching for the robust and optimal train speed trajectory that are suitable for real-life operations.

##### A. Case 1

In this case, the whole track is on a flat plain and only the constant speed limit is considered. Thus the control variable

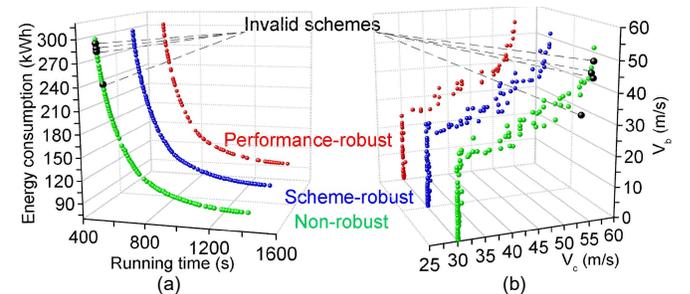


Fig. 8. Comparisons of the Pareto fronts and solutions distribution under different situations in case 1. (a) Pareto fronts; (b) Solutions distribution.

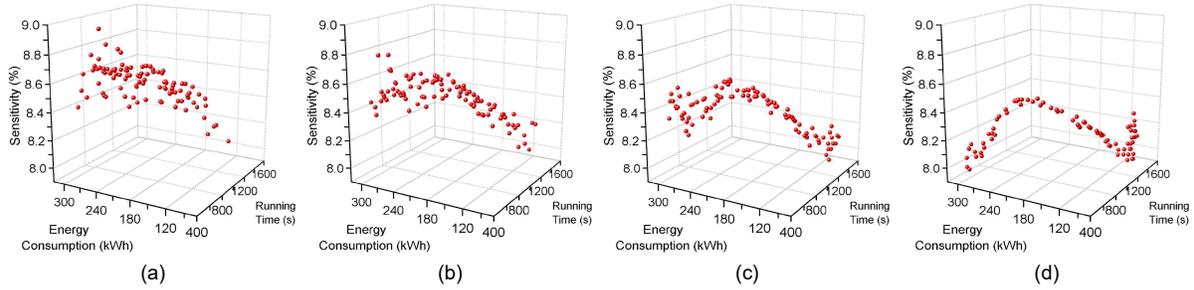


Fig. 9. Iterations of p-NSGA-II in case 1. (a) 20 iterations; (b) 50 iterations; (c) 75 iterations; (d) 100 iterations.

TABLE I  
CASE STUDY PARAMETERS

| Parameter                                 | Value                         |
|---|-------------------------------|
| Journey distance $S$                      | 20 km                         |
| Empty vehicle mass $M_v$                  | 420 ton                       |
| Maximum passenger mass $M_{p-max}$        | 160 ton                       |
| Basic resistance $r_1$                    | $1.456 + 0.0183v + 0.0034v^2$ |
| Speed limit $v_{max}$                     | 55 m/s                        |
| Maximum traction/braking effort $F_{max}$ | 200 kN                        |
| Maximum power $P_{max}$                   | 5000 kW                       |
| Regeneration efficiency $\eta$            | 0.8                           |
| Range of cruising speed $v_c$             | [30, 55] m/s                  |
| Range of braking speed $v_b$              | [0, 55] m/s                   |

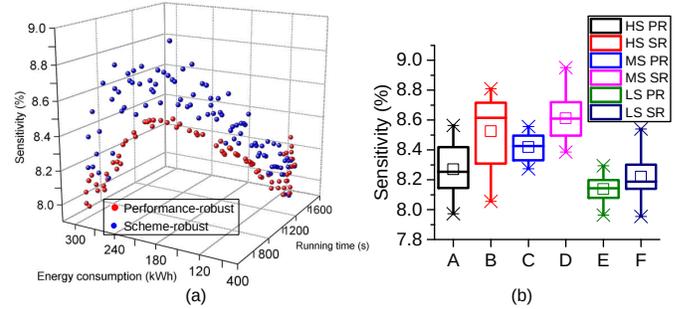


Fig. 11. Comparisons of sensitivity between scheme-robust solutions and performance-robust solutions ( $\lambda_1 = 0.5$ ,  $\lambda_2 = 0.5$ ) in case 1.

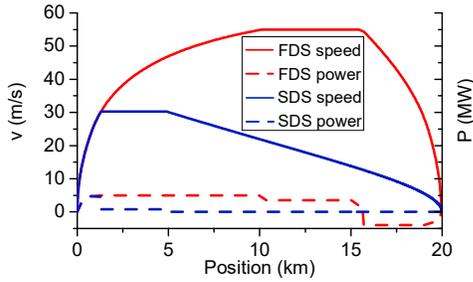


Fig. 10. Fastest driving scheme and slowest driving scheme.

is  $V = (0, v_c, v_b, 0)$  where  $v_c$  and  $v_b$  need to be determined to get the set of driving schemes. The NSGA-II is applied first to solve the bi-objective train speed trajectory optimization problem defined by (8)-(9) without considering any robustness. For the NSGA-II, the solution number in a population is 100 and the maximum iteration number is set to 100. The crossover probability is 0.9 and mutation probability is 0.05. The train mass of 500 tons equaling the sum of empty vehicle mass (420 tons) and average passenger load (80 tons), is applied to the optimization model. The Pareto front obtained in this initial run is shown in Fig. 8(a) in terms of their two objective function values.

All Pareto solutions (green set) in Fig. 8(a) are checked by scheme robustness constraint in (10). Among the whole set, four solutions, labeled in black in Fig. 8(a), are non-scheme-robust, hence can not fit all operation conditions. The optimization procedure is implemented again with the scheme robustness applied as a constraint during NSGA-II iterations, hence all Pareto solutions are checked in each

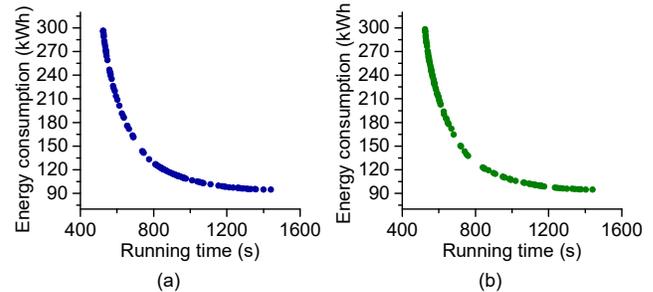


Fig. 12. Pareto fronts of two sets of weighting factors. (a)  $\lambda_1 = 1$ ,  $\lambda_2 = 0$ ; (b)  $\lambda_1 = 0$ ,  $\lambda_2 = 1$ .

iteration and the non-scheme-robust solutions are filtered out. Consequently, the Pareto solutions (blue set) are all scheme-robust and the optimization result is also shown in Fig. 8(a). Fig. 8(b) shows the distributions of the non-scheme-robust solutions and scheme-robust solutions in terms of their  $v_c$  and  $v_b$  respectively. It can be seen that applying or not applying scheme robustness constraint does not significantly affect the distributions of the Pareto solutions. In general, the non-scheme-robust driving strategies are caused by insufficient braking distance and braking force. When the chosen driving strategy possesses a high cruising speed and a relatively low braking speed, the acceleration distance is prolonged to catch up with the target cruising speed as the train mass increases. Meanwhile, to decelerate the train to the set braking speed by the inertia from a high cruising speed, the coasting distance may be too long and the braking distance is likely insufficient. Thus this is also evident that as shown in Fig. 8, the driving strategies with a big gap between the cruising and coasting

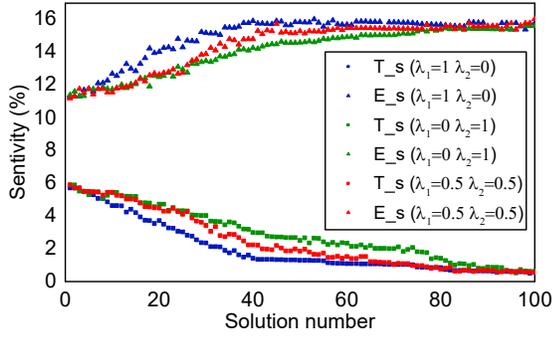


Fig. 13. Comparison of solution sensitivity under three different sets of weighting factors in case 1.

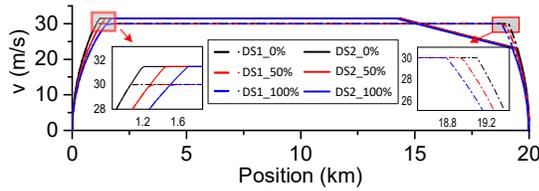


Fig. 14. Comparison of time-robust and non-robust driving strategies.

speeds are not included in the robust Pareto sets.

Considering the performance robustness of each solution, the proposed p-NSGA-II is adopted to produce a performance-robust Pareto front. Equation (12), as the DM preference function, is applied in the p-NSGA-II framework to evaluate the sensitivity of each driving scheme. The goal of the p-NSGA-II is to find solutions with lower sensitivity (more performance-robust). To achieve equilibrium on the robustness of energy consumption and running time, in this case,  $\lambda_1$  and  $\lambda_2$  in (12) are both set to 0.5. The parameters of GA operations are the same as above.

As the iteration goes on, the population gradually converges to the lower nondomination ranks, which is consistent with the description in Section III. After 20 iterations, all solutions converge to the first nondomination rank. In the 3-D space, the x-axis, y-axis and z-axis represent energy consumption, running time and driving scheme sensitivity respectively, the solution distribution is shown in Fig. 9(a). As the iteration continues, from Fig. 9(b)-(d), it is apparent that the population is further guided towards the more performance-robust region.

The performance-robust Pareto front (red set) solved by p-NSGA-II and the solution distribution are also shown in Figs. 8(a) and (b) respectively. Compared with the scheme-robust Pareto front, the spread of the performance-robust Pareto front is less even than the former but is still acceptable. And the solution distribution is also not significantly changed. Table II lists the results for three different robust situations. It is shown that, when considering robustness, the spread of the Pareto fronts will not become worse compared to the Pareto front without considering robustness.

Fig. 10 illustrates the fastest driving scheme (FDS) and the slowest driving scheme (SDS) among the performance-robust Pareto solutions. The FDS accelerates the train as much as possible to the maximum speed without any coasting

operation. While for the slowest driving scheme (most energy-saving driving scheme), the cruising speed nearly equals the minimum speed limit. The coasting operation is maximally utilized to reduce energy consumption. The running time of FDS is only 36.5% of SDS, but consumes 3.19 times as much energy as the SDS.

In Fig. 11(a), the comparison of the performance-robust solutions (red balls) with the scheme-robust solutions (blue balls) under  $\lambda_1 = 0.5$  and  $\lambda_2 = 0.5$ , demonstrates that the p-NSGA-II does produce a set of solutions with overall lower sensitivity to passenger load variation. To further explain the differences of robustness between both sets of solutions, each set of solutions is divided into three groups based on the running time. They are high-speed (HS) group (33 solutions), medium-speed (MS) group (33 solutions) and low-speed (LS) group (34 solutions), respectively. Fig. 11(b) presents, amongst the six groups, all three groups from the performance-robust solutions have better indices of robustness than the other three groups from scheme-robust solutions. To compare the overall robustness between the scheme-robust solutions and performance-robust solutions, the normalized average sensitivity values of both sets of solutions are calculated by (18), where  $S_{SR}$  and  $S_{PR}$  represent the sets of sensitivity values of the scheme-robust solutions and performance-robust solutions respectively. Compared to the scheme-robust solutions, the average sensitivity reduction (robustness improvement) of the performance-robust solutions is 40.59% calculated by (18). The superiority of the proposed p-NSGA-II is thus validated.

$$S_{SR/PR}^{ave} = \frac{1}{N} \sum_{n=1}^N \frac{S_{SR/PR}^n - \min(S_{SR} \vee S_{PR})}{\max(S_{SR} \vee S_{PR}) - \min(S_{SR} \vee S_{PR})} \quad (18)$$

$$R = \frac{S_{SR}^{ave} - S_{PR}^{ave}}{S_{SR}^{ave}} \cdot 100\% \quad (19)$$

To further analyze the effects of weighting factors,  $\lambda_1$  and  $\lambda_2$ , in performance robustness equation (12) on optimization results, another two sets of values are adopted. The first set is  $\lambda_1 = 1$  and  $\lambda_2 = 0$ , this means only the robustness of the running time is considered, while the energy consumption robustness is ignored. Another set is  $\lambda_1 = 0$  and  $\lambda_2 = 1$ , so only energy consumption robustness is considered, neglecting the running time. When the weighting factors are set to be  $\lambda_1 = 0$  and  $\lambda_2 = 1$ , the objective is to find a set of optimal solutions with the best robustness for energy consumption by p-NSGA-II. While if the weighting factors are set  $\lambda_1 = 1$  and  $\lambda_2 = 0$ , p-NSGA-II will produce a set of optimal solutions with the best robustness for the journey time. When the weighting factors are set  $\lambda_1 = 0.5$  and  $\lambda_2 = 0.5$ , a set of balanced-robust solutions are obtained. The Pareto fronts of both cases are shown in Figs. 12 (a) and (b). To compare each solution, the sensitivities of the running time and energy consumption are evaluated by (20) and (21) below respectively.

$$T_{-s} = \max \left( \frac{|t - t_{\min}|}{t}, \frac{|t - t_{\max}|}{t} \right) \quad (20)$$

$$E_{-s} = \max \left( \frac{|E - E_{\min}|}{E}, \frac{|E - E_{\max}|}{E} \right) \quad (21)$$

The comparison of the sensitivity of the running time and energy consumption under three sets of weighting factors is shown in Fig. 13 and Table II. In Fig. 13, 100 solutions under each set of weighting factors are sorted based on the running time. The number 1 corresponds to the solution with the shortest running time, while the number 100 corresponds to that with the longest running time. From Fig. 13 and Table II, it is evident that:

- 1) For all three sets of weighting factors, as the journey time increases, the running time sensitivity decreases but energy consumption sensitivity increases. The sensitivities of running time and energy consumption are negatively correlated. Under the driving schemes with a high average speed, the variation of energy consumption due to the train mass uncertainty is relatively small. While those driving schemes with a low average speed perform well in running time robustness.
- 2) The weighting factors determine the priority of DM preference. Choosing different sets of weighting factors is to trade off the robustness of running time and the energy consumption. If only the robustness of running time is considered, the solutions' robustness of energy consumption will perform poorly. On the other hand, when only consider the robustness of energy consumption, the robustness of running time will be poorer. When  $\lambda_1 = 1$  and  $\lambda_2 = 0$ , the solutions are most robust for running time. When  $\lambda_1 = 0$  and  $\lambda_2 = 1$ , the set of solutions are most robust for energy consumption. The solutions under any other weighting factors, for example  $\lambda_1 = 0.5$  and  $\lambda_2 = 0.5$ , will be somewhere in between.
- 3) The average energy sensitivity is 13.7% for the most energy robust set ( $\lambda_1 = 0$ ,  $\lambda_2 = 1$ ) compared to 14.4% of the most time robust set ( $\lambda_1 = 1$ ,  $\lambda_2 = 0$ ). Take 15% energy deviation as the threshold since the maximum energy deviations are larger than 15% without optimization as shown in [31]. The energy sensitivity of 62 solutions out of 100 (62%) in the most energy-robust set is lower than the threshold, while only 28 solutions (28%) in the most time-robust set are lower than the threshold. The most time robust set has a lower average time sensitivity of 1.6% compared to 2.5% of the most energy-robust set. The average delay time of the most time-robust set is 13.7 s compared to 18.2 s of the most energy robust set. Based on the experience from Beijing Yizhuang Line Links [31], [42], assume the time sensitivity threshold is set to 4%, the most time robust set has 85% satisfactory rate compared to 70% of the most energy-robust set. The effectiveness of the proposed p-NSGA-II is therefore evident.

A comparison of a time robust speed profile and another speed profile only catching up with the braking curve is performed. For the first driving strategy, a target cruising speed 30 m/s without optimization is set. The train is accelerated to the target speed and then kept the constant speed until the train reaches the latest braking point. The second driving strategy (31.5 m/s, 23.1 m/s) is optimized by the proposed p-NSGA-II under  $\lambda_1 = 1$ ,  $\lambda_2 = 0$  (most time robust). The

speed trajectories under different passenger loads are shown in Fig. 14. The journey time is 725 s under empty load, 738 s under 50% load and 753 s under 100% load for DS1, and 726 s, 738 s and 749 s for DS2. It is evident that, under the mean value of the train mass, the running time of both driving strategies is the same. While the running time deviations of both driving strategies under the mean value of the train mass are 13 s and 12 s to the empty load, and 15 s and 11 s to the full load. The maximum deviations of running time for both driving strategies are 15 s and 12 s respectively, corresponding to the maximum running time sensitivities of 2.0% and 1.6%. Thus the performance of the proposed p-NSGA-II is validated.

The performance robustness reflects the maximum sensitivities of energy consumption and running time to the variation of the passenger load. This performance value is primarily determined by both the resultant driving strategies that are generated by our algorithm and by the DM's preference which are reflected in the two preference weights  $\lambda_1$  and  $\lambda_2$ . In other words, infinite solutions may exist in the Pareto front while their sensitivities to energy consumption and running are different. The final selected Pareto solutions are determined by the DM (dispatcher) preference and are generated by the proposed p-NSGA-II.

To further illustrate the benefits of our proposed algorithm p-NSGA-II and the optimization framework, the comparisons between robust optimization (RO), stochastic optimization (SO) and our proposed method are performed. A widely adopted framework of the uncertain multi-objective optimization is shown in the following [32], [56]:

$$\min (f_1(\hat{x}, \bar{m}), f_2(\hat{x}, \bar{m}), \dots, f_K(\hat{x}, \bar{m})) \quad (22)$$

$$s.t. \frac{\|\hat{f}^P(\hat{x}, m) - \hat{f}(\hat{x}, \bar{m})\|}{\|\hat{f}(\hat{x}, \bar{m})\|} \leq \eta \quad \hat{x} \in S \quad (23)$$

where  $\hat{f}(\hat{x}) = (f_1(\hat{x}), f_2(\hat{x}), \dots, f_K(\hat{x}))$ .  $\hat{x}$  is the vector of the control variables.  $m$  represents the uncontrollable uncertain parameters.  $\bar{m}$  is the mean value of the uncontrollable uncertain parameters.  $\hat{f}^P$  represents the perturbed objective vector and  $\eta$  is the maximum value of sensitivity required so that the decision maker can adjust the maximum value of the sensitivity of the solutions. The perturbed objective vector can be chosen either to be the worst case to formulate a RO problem, or to be the mean effective value of the neighborhood to formulate a SO problem. In the following, both optimization models will be generated and the two objectives still remain to be the same, i.e. minimizing the energy consumption and running time for the given mean value of the train mass. The sensitivity  $\eta$  is set to different values. Notably, the perturbed objective vector in the SO problem needs to be calculated by the Monte Carlo method, and assume the mass of passengers is distributed as discrete normal distribution  $M \sim N(\mu, \sigma^2)$  for a specific substation and operational period. The RO optimization model in (22) and (23) was adopted in [32]. Both RO and SO are performed by traditional NSGA-II.

Figs. 15-16 show the Pareto fronts and the sensitivities by RO. In Fig. 15, the dark green and pink dots represent the energy sensitivity and time sensitivity of each solution

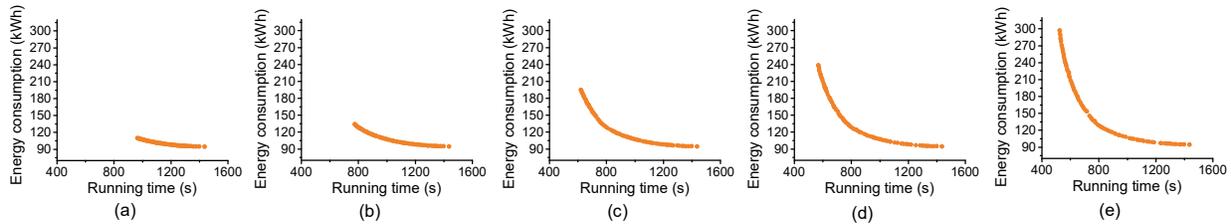


Fig. 15. RO Pareto fronts. (a)  $\eta = 0.02$ ; (b)  $\eta = 0.03$ ; (c)  $\eta = 0.05$ ; (d)  $\eta = 0.06$ ; (e)  $\eta = 0.07$ .

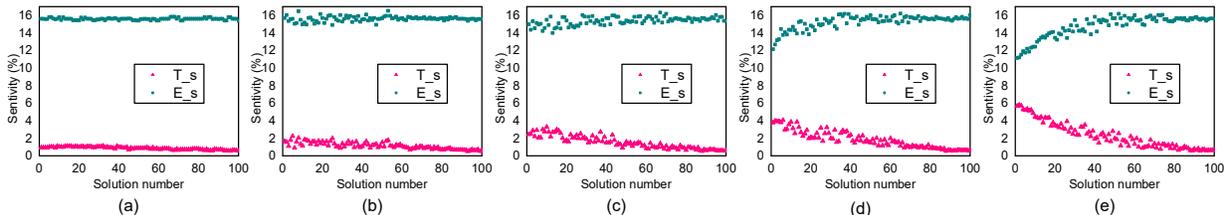


Fig. 16. RO sensitivities. (a)  $\eta = 0.02$ ; (b)  $\eta = 0.03$ ; (c)  $\eta = 0.05$ ; (d)  $\eta = 0.06$ ; (e)  $\eta = 0.07$ .

respectively. In Figs. 15 (a)-(d),  $\eta$  varies from 0.02 to 0.06, and it is obvious that the solution distributions are narrowed down compared to the solutions generated by our proposed p-NSGA-II. As shown in Table II, the Pareto distributions of energy decrease by 92.6%, 80.8%, 51.4% and 30.3% respectively, while the time distributions decrease by 48.4%, 27.5%, 10.7% and 4.8% respectively. These results are consistent with [32]. For  $\eta = 0.07$ , the Pareto distribution is similar to the results generated by our proposed p-NSGA-II. Compared to the p-NSGA-II solution sets with  $\lambda_1 = 1$ ,  $\lambda_2 = 0$  (most time robust), and  $\lambda_1 = 0$ ,  $\lambda_2 = 1$  (most energy robust) as shown in Fig. 13, the sensitivity values of the RO solutions with  $\eta = 0.07$  in Fig.16(e) are distributed disorderly within the boundary constituted by the solutions optimized by p-NSGA-II, which implies that the  $\eta$  value can no longer play a role in selecting the robust solutions at this stage. When  $\eta$  is set to 0.01, there has been even no solution satisfying the sensitivity limit that can be found in the feasible region. Therefore, the solutions as the results of different  $\eta$  values are ambiguous and unpredictable. Sometimes, an inappropriate  $\eta$  value can lead to a narrowed solution distribution or even a null solution set. And the RO algorithm proposed in [32] cannot well manage different robustness indicators. Our proposed p-NSGA-II, on the one hand, can set the preferences of different robust indicators to generate the most-time-robust, most-energy-robust, or balanced solution sets. Furthermore, the solution distribution will not be narrowed while taking into account the requirements from dispatchers.

The SO was also applied. The passenger mass values are distributed as  $M \sim N(80, 26^2)$  so that the range of the passenger mass is kept the same for different approaches. The Monte Carlo samples are set to 50 to calculate the perturbed objective vector. However, it was found that the perturbed objective vectors with  $\eta$  decreasing from 0.03 to 0.001 are very close to the objective vector calculated with the mean value of the train mass, which implies that the mean deviations of the energy consumption and running time are insensitive to the

preset values. As shown in Fig. 17(a), the Pareto front is not further filtered compared with the original optimization with the mean value of the train mass even if the preset  $\eta$  value is sufficiently small ( $\eta=0.001$ ). The solution sensitivities are also distributed disorderly as shown in Fig. 17(b). The perturbed objective vector by Monte Carlo samples shows the long-term patterns of energy consumption and running time for a specific driving strategy. It is evident that the solutions optimized by p-NSGA-II with the mean value of the train mass can sufficiently reflect the long-term operation situations and achieve long-term energy saving. While SO does not limit the maximum sensitivities (deviations) of energy consumption and running time. Furthermore, from Table II, it can be found that the computational time of p-NSGA-II is longer than that of other algorithms except for SO method. This is expected because p-NSGA-II has the process of preference ranking, which leads to extra time-spending. While for SO, Monte Carlo sampling is needed hence resulting in far more computational time than p-NSGA-II (nearly three times).

In summary, four algorithms, namely the conventional NSGA-II, RO algorithm, SO algorithm and the proposed p-NSGA-II are tested to verify the effectiveness and superiority of p-NSGA-II. The NSGA-II does not consider the robustness performance. Thus its solutions have the worse robustness in both energy consumption and journey time compared to the solutions obtained by p-NSGA-II under  $\lambda_1 = 0.5$ ,  $\lambda_2 = 0.5$ . And the robustness of energy consumption is worse than that obtained by p-NSGA-II under  $\lambda_1 = 0$ ,  $\lambda_2 = 1$ . The robustness of journey time is worse than that obtained by  $\lambda_1 = 1$ ,  $\lambda_2 = 0$ . When RO algorithm proposed in [32] is applied, a pre-defined parameter  $\eta$  needs to be determined, which controls the convergent direction. As shown in Table II, when  $\eta$  is set to have a small value such as  $\eta = 0.02$ , the distribution of the Pareto solutions is narrowed down significantly. If the dispatcher expects to choose the driving scheme with a short journey time below 965 s, the RO algorithm under  $\eta = 0.02$  can not provide any satisfactory solutions. The same issue also

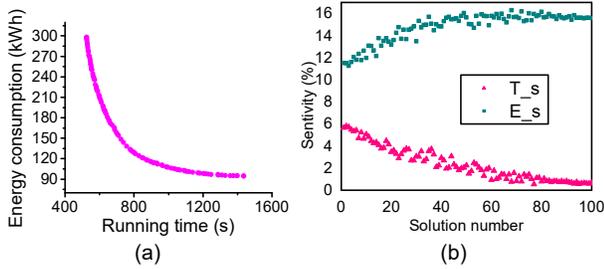


Fig. 17. SO results. (a) SO Pareto front; (b) SO sensitivities.

happens when  $\eta$  is set 0.03, 0.05 and 0.06, and no suitable solutions can be provided to the dispatcher. When  $\eta$  is set a relatively large value such as  $\eta = 0.07$ , the distribution of the Pareto solutions is close to that obtained by p-NSGA-II. However, the average sensitivities of energy consumption and journey time obtained by RO algorithm under  $\eta = 0.07$  are 14.3% and 1.9% which are worse than those of 13.9% and 1.8% obtained by p-NSGA-II under  $\lambda_1 = 0.5, \lambda_2 = 0.5$  as shown in Table II. And the average sensitivity of energy consumption is worse than that obtained by p-NSGA-II under  $\lambda_1 = 0, \lambda_2 = 1$ . The average sensitivity of journey time is worse than that obtained by  $\lambda_1 = 1, \lambda_2 = 0$ . Further, the RO algorithm can not control the energy consumption robustness and journey time robustness separately. It is evident from Table II that the RO algorithm shows these drawbacks when compared to the proposed p-NSGA-II. For the SO algorithm, first, it is more time-consuming than RO algorithm and p-NSGA-II due to the application of the Monte Carlo sampling method, and it also shows the worse robust performance for both energy consumption and journey time compared to the p-NSGA-II. For the proposed p-NSGA-II, not only can the energy consumption robustness and journey time robustness be controlled separately by adjusting weighting factors, but also the distribution of the obtained Pareto front is always maintained. As shown in Table II, under three different sets of weighting factors, the distribution of the solutions is comprehensive and close to that obtained by original NSGA-II without considering robustness. When the weighting factors are  $\lambda_1 = 0, \lambda_2 = 1$ , a set of solutions with lowest average sensitivity 13.7% of energy consumption is obtained compared to that obtained by other algorithms. When the weighting factors are  $\lambda_1 = 1, \lambda_2 = 0$ , a set of solutions with lowest average sensitivity 1.6% of journey time are obtained. It is consistent with the expectation for the weighting factors.

### B. Case 2

In this case, the track considered is not on a flat surface but has ramps with different gradients as illustrated in Fig. 18. It results in the whole journey being divided into 5 subsections with different speed limits based on the track gradients which are also revealed in Fig. 18. Thus the corresponding control variable set is  $V = (0, v_{1,c}, v_{1,b}, v_{1,e}, \dots, v_{5,s}, v_{5,c}, v_{5,b}, 0)$ . There are 14 control variables in total to be determined considering  $v_{i,e} = v_{i+1,s}$ . In order to avoid the low average speed of the train resulting in a long journey time, the minimum

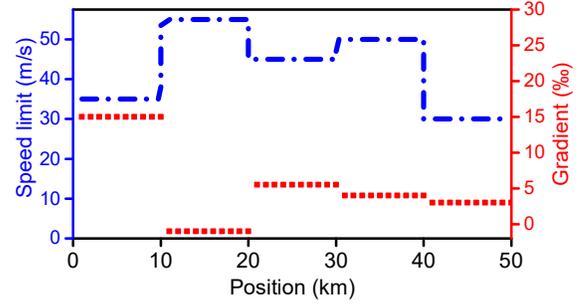


Fig. 18. Track ramps in case 2.

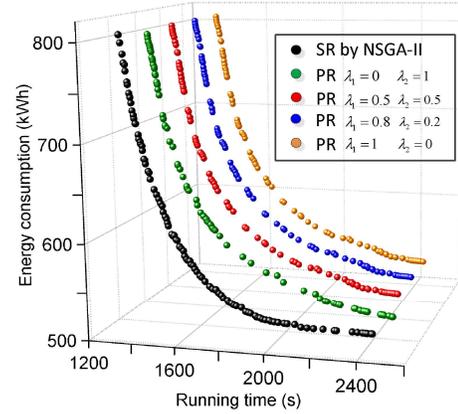


Fig. 19. Comparisons of the Pareto fronts in case 2.

cruising speed of the first four subsections is set to 30 m/s, and 15 m/s for the last subsection.

First, the optimization problem is solved by NSGA-II only considering scheme robustness but not performance robustness. The parameters for GA operations are still the same as those used in Case 1. Then the performance robustness is considered. Four cases with weighting factors ( $\lambda_1 = 0, \lambda_2 = 1$ ) - most energy-robust, ( $\lambda_1 = 0.5, \lambda_2 = 0.5$ ) - balanced, ( $\lambda_1 = 0.8, \lambda_2 = 0.2$ ) - time robust priority, ( $\lambda_1 = 1, \lambda_2 = 0$ ) - most time-robust, are performed. The Pareto fronts optimized by NSGA-II and proposed p-NSGA-II are shown in Fig. 19. The spread of the Pareto fronts and corresponding computational time are listed in Table II. The spread of the Pareto fronts optimized by p-NSGA-II is not getting significantly worse compared to that optimized by NSGA-II. Fig. 20 shows the fastest and most energy-saving driving strategies obtained by p-NSGA-II which are labeled in red line and blue line respectively. The fastest driving strategy is obviously energy-consuming taking 809.6 kWh with a shortest journey time 1381 s. The most energy-saving driving scheme takes 523.1 kWh with a 2482 s journey time which has a 35.4% decrease of energy consumption compared to that of the fastest driving scheme. For the fastest driving strategy, the train always reaches the maximum allowable speed in each subsection and the coasting operation is rarely or never applied. For the most energy-saving driving strategy obtained, the full braking operation is rarely or never applied, instead the coasting operation is applied as much as possible.

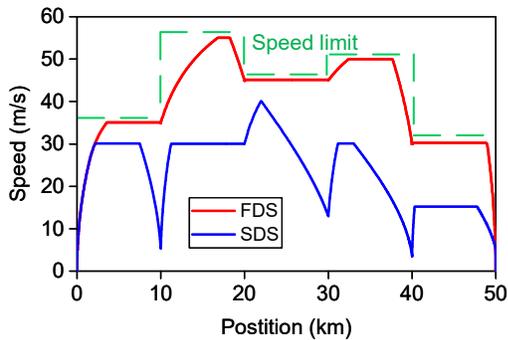


Fig. 20. Fastest and most energy-saving driving schemes (FDS/ SDS).

This is consistent with the conclusion drawn in [27].

Figs. 21-24 demonstrate the evolutionary convergence process of the p-NSGA-II under different weighting factors. As the iterations goes on, all solutions are guided to the Pareto front with lower sensitivities. This implies that a set of more robust driving schemes can be obtainable. The time sensitivities and energy sensitivities of the solutions optimized by the NSGA-II and p-NSGA-II are compared as shown in Fig. 25 (a) and (b) respectively. In Fig. 25 (a), the set of solutions under  $\lambda_1 = 1$  and  $\lambda_2 = 0$  have the overall lowest time sensitivities compared to the solution set under other weighting factors. With  $\lambda_1$  increasing from 0 to 0.5, 0.8 and 1, the overall time sensitivity of the solutions overall decreases which implies that the time robustness of the solutions becomes better. As shown in Table II, the average time sensitivity under  $\lambda_1 = 1$  and  $\lambda_2 = 0$  is 1.7% which is lower than that of 1.8% under  $\lambda_1 = 0.5, \lambda_2 = 0.5$ , 1.9% under  $\lambda_1 = 0.8, \lambda_2 = 0.2$  and 2.0% under  $\lambda_1 = 1, \lambda_2 = 0$ . As shown in Fig. 25 (b), the energy sensitivities of the solutions by p-NSGA-II under  $\lambda_1 = 0$  and  $\lambda_2 = 1$  are overall lower than those of solutions under other weighting factors. With  $\lambda_2$  increasing from 0 to 0.2, 0.5 and 1, the overall energy sensitivity of the solutions decreases which implies that the energy robustness of the solutions becomes better. As shown in Table II, the average energy sensitivity under  $\lambda_1 = 0$  and  $\lambda_2 = 1$  is 15.4% which is lower than that of 15.5% under  $\lambda_1 = 0.5, \lambda_2 = 0.5$ , 15.6% under  $\lambda_1 = 0.8, \lambda_2 = 0.2$  and 15.8% under  $\lambda_1 = 1, \lambda_2 = 0$ . It is therefore confirmed that the proposed p-NSGA-II has shown to be capable of producing a set of robust Pareto solutions by adjusting the weighting factors based on the DM preference. The weighting factors of  $\lambda_1$  and

$\lambda_2$  dictate the journey time robustness and energy consumption robustness respectively. While in Fig. 25 (a) and (b), the time sensitivity and energy sensitivity of the solutions by NSGA-II are both relatively larger and distributed disorderly with an average time sensitivity 2.1% and an average time sensitivity 15.6%. The average time sensitivity of the solutions obtained by NSGA-II is 23.5% higher than that of the solutions under  $\lambda_1 = 1, \lambda_2 = 0$  obtained by p-NSGA-II. And the average time sensitivity of the solutions obtained by NSGA-II is 1.3% higher than that of the solutions under  $\lambda_1 = 0, \lambda_2 = 1$  obtained by p-NSGA-II. Even if compared to the solutions under  $\lambda_1 = 0.5, \lambda_2 = 0.5$ , the solutions obtained by the original NSGA-II have lower overall and average robustness of journey time and energy consumption. These again confirm the effectiveness of the proposed p-NSGA-II.

### V. CONCLUSION

This paper has proposed a new DM preference based multi-objective optimization algorithm namely p-NSGA-II for robust train speed trajectories under the uncertainty of passenger load. This has been formulated as a bi-objective problem of minimizing energy consumption and running time, with specific constraints. By adding the scheme robustness constraint, the non-scheme-robust solutions are filtered out and all solutions are guaranteed to be feasible. The p-NSGA-II method has been developed to further search the performance-robust solutions within the feasible solutions. Two case studies have been conducted to verify the effectiveness of the proposed robust optimization approach. It is revealed that, compared to the optimization results obtained without considering performance robustness, the proposed approach p-NSGA-II can effectively produce a well spread-out and robust set of driving schemes with up to 40.59% average robustness improvement. Note that the proposed p-NSGA-II proves to be adept among the preference function-based multi-objective optimization problems, and thus can be easily applied to a number of other related engineering optimization problems.

### REFERENCES

- [1] J. Yin, T. Tang, L. Yang, J. Xun, Y. Huang, and Z. Gao, "Research and development of automatic train operation for railway transportation systems: A survey," *Transportation Research Part C: Emerging Technologies*, vol. 85, pp. 548–572, 2017.
- [2] G. M. Scheepmaker, R. M. Goverde, and L. G. Kroon, "Review of energy-efficient train control and timetabling," *European Journal of Operational Research*, vol. 257, no. 2, pp. 355–376, 2017.

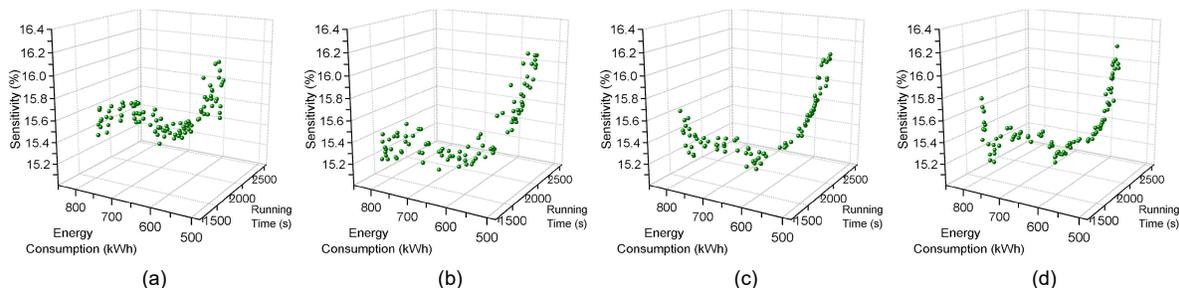


Fig. 21. Iterations of p-NSGA-II under  $\lambda_1 = 0, \lambda_2 = 1$ . in case 2 (a) 20 iterations; (b) 50 iterations; (c) 75 iterations; (d) 100 iterations.

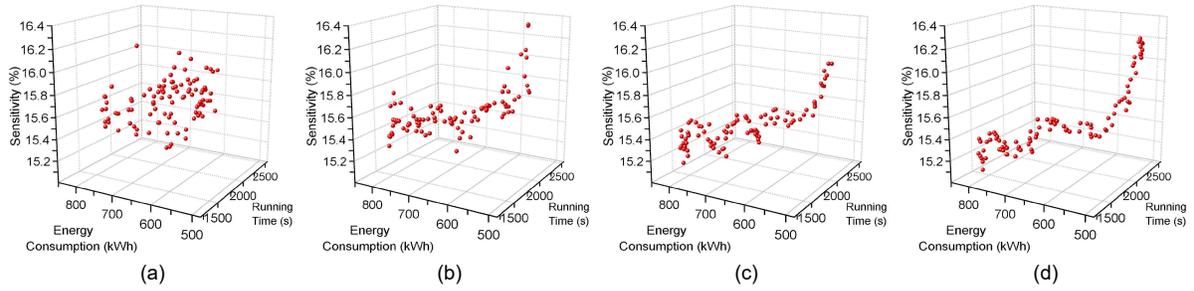


Fig. 22. Iterations of p-NSGA-II under  $\lambda_1 = 0.5$ ,  $\lambda_2 = 0.5$ . in case 2 (a) 20 iterations; (b) 50 iterations; (c) 75 iterations; (d) 100 iterations.

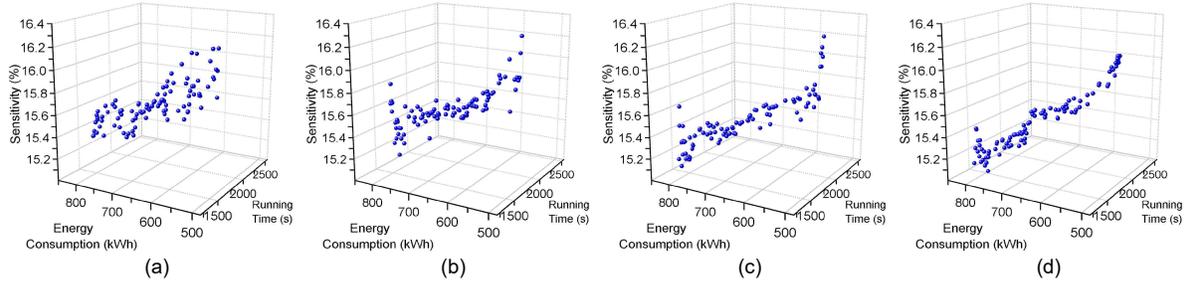


Fig. 23. Iterations of p-NSGA-II under  $\lambda_1 = 0.8$ ,  $\lambda_2 = 0.2$ . in case 2 (a) 20 iterations; (b) 50 iterations; (c) 75 iterations; (d) 100 iterations.

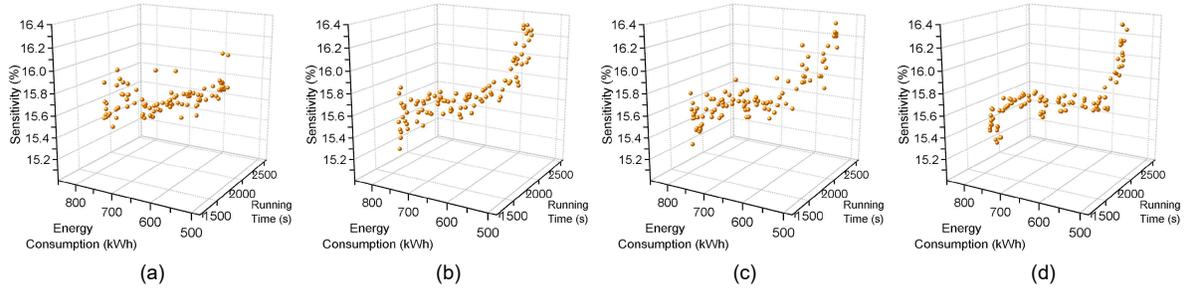


Fig. 24. Iterations of p-NSGA-II under  $\lambda_1 = 1$ ,  $\lambda_2 = 0$ . in case 2 (a) 20 iterations; (b) 50 iterations; (c) 75 iterations; (d) 100 iterations.

- [3] X. Yang, B. Ning, X. Li, and T. Tang, "A two-objective timetable optimization model in subway systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 1913–1921, 2014.
- [4] X. Yang, X. Li, Z. Gao, H. Wang, and T. Tang, "A cooperative scheduling model for timetable optimization in subway systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 1, pp. 438–447, 2012.
- [5] S. Su, X. Li, T. Tang, and Z. Gao, "A subway train timetable optimization approach based on energy-efficient operation strategy," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 2, pp. 883–893, 2013.
- [6] J. Liao, G. Yang, S. Zhang, F. Zhang, and C. Gong, "A deep reinforcement learning approach for the energy-aimed train timetable rescheduling problem under disturbances," *IEEE Transactions on Transportation Electrification*, vol. 7, no. 4, pp. 3096–3109, 2021.
- [7] C. Wu, S. Lu, F. Xue, L. Jiang, M. Chen, and J. Yang, "A two-step method for energy-efficient train operation, timetabling, and onboard energy storage device management," *IEEE Transactions on Transportation Electrification*, vol. 7, no. 3, pp. 1822–1833, 2021.
- [8] C. Wu, S. Lu, F. Xue, L. Jiang, and M. Chen, "Optimal sizing of on-board energy storage devices for electrified railway systems," *IEEE Transactions on Transportation Electrification*, 2020.
- [9] J. A. Aguado, A. J. S. Racero, and S. de la Torre, "Optimal operation of electric railways with renewable energy and electric storage systems," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 993–1001, 2016.
- [10] C. Wu, W. Zhang, S. Lu, Z. Tan, F. Xue, and J. Yang, "Train speed trajectory optimization with on-board energy storage device," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 11, pp. 4092–4102, 2018.
- [11] P. Arboleya, B. Mohamed, and I. El-Sayed, "Dc railway simulation including controllable power electronic and energy storage devices," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5319–5329, 2018.
- [12] S. de la Torre, A. J. Sánchez-Racero, J. A. Aguado, M. Reyes, and O. Martínez, "Optimal sizing of energy storage for regenerative braking in electric railway systems," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1492–1500, 2014.
- [13] H. Yang, W. Shen, Q. Yu, J. Liu, Y. Jiang, E. Ackom, and Z. Y. Dong, "Coordinated demand response of rail transit load and energy storage system considering driving comfort," *CSEE Journal of Power and Energy Systems*, vol. 6, no. 4, pp. 749–759, 2020.
- [14] C. Wu, B. Xu, S. Lu, F. Xue, L. Jiang, and M. Chen, "Adaptive eco-driving strategy and feasibility analysis for electric trains with onboard energy storage devices," *IEEE Transactions on Transportation Electrification*, vol. 7, no. 3, pp. 1834–1848, 2021.
- [15] X. Yang, X. Li, B. Ning, and T. Tang, "A survey on energy-efficient train operation for urban rail transit," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 1, pp. 2–13, 2015.
- [16] S. Lu, M. Q. Wang, P. Weston, S. Chen, and J. Yang, "Partial train speed trajectory optimization using mixed-integer linear programming," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 10, pp. 2911–2920, 2016.
- [17] Z. Tan, S. Lu, F. Xue, and K. Bao, "A speed trajectory optimization model for rail vehicles using mixed integer linear programming," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2017, pp. 1–6.
- [18] A. Cucala, A. Fernández, C. Sicre, and M. Domínguez, "Fuzzy optimal schedule of high speed train operation to minimize energy consumption

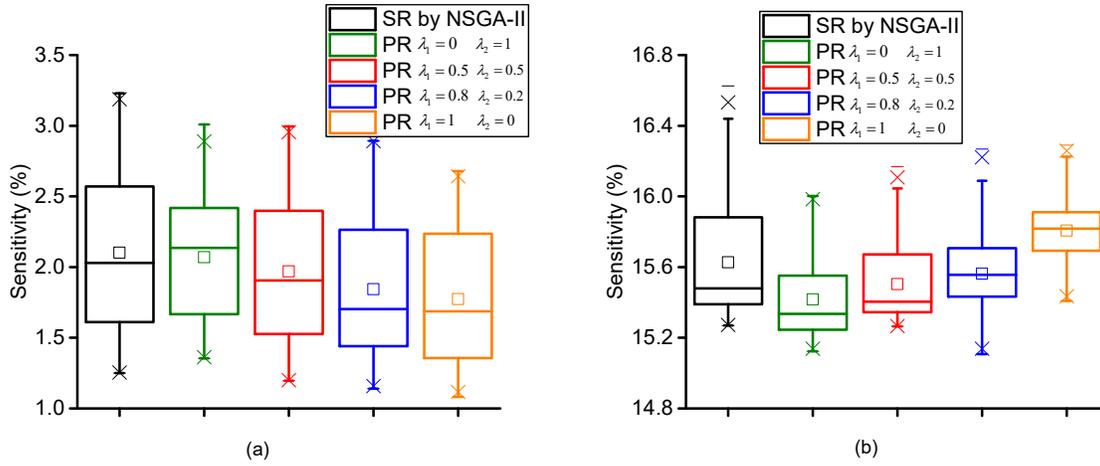


Fig. 25. Comparisons of the Pareto fronts and solutions distribution under different situations in case 2. (a) Time sensitivity; (b) Energy sensitivity.

TABLE II  
COMPARISONS OF DIFFERENT ALGORITHMS

|                      | Energy distribution (kWh)         | Running time distribution (s) | Computational time (s) | Average energy sensitivity (%) | Average time sensitivity (%) |            |
|----------------------|-----------------------------------|-------------------------------|------------------------|--------------------------------|------------------------------|------------|
| Case 1               |                                   |                               |                        |                                |                              |            |
| NSGA-II (non-robust) | 94.5-301.0                        | 524-1436                      | 152.58                 | 14.3                           | 1.8                          |            |
| NSGA-II (SR)         | 94.5-301.0                        | 524-1436                      | 489.51                 | 14.2                           | 2.0                          |            |
| p-NSGA-II            | $\lambda_1 = 0 \lambda_2 = 1$     | 94.5-296.6                    | 525-1436               | 993.21                         | <b>13.7</b>                  | 2.5        |
|                      | $\lambda_1 = 0.5 \lambda_2 = 0.5$ | <b>94.5-301.0</b>             | <b>524-1436</b>        | 985.60                         | 13.9                         | 1.8        |
|                      | $\lambda_1 = 1 \lambda_2 = 0$     | 94.5-296.6                    | 525-1436               | 953.35                         | 14.4                         | <b>1.6</b> |
| RO                   | $\eta = 0.02$                     | 94.5-109.7                    | 965-1436               | 513.29                         | 15.7                         | 0.8        |
|                      | $\eta = 0.03$                     | 94.5-134.2                    | 775-1436               | 498.64                         | 15.6                         | 1.1        |
|                      | $\eta = 0.05$                     | 94.5-194.8                    | 622-1436               | 501.54                         | 15.3                         | 1.4        |
|                      | $\eta = 0.06$                     | 94.5-238.4                    | 568-1436               | 536.86                         | 15.1                         | 1.6        |
|                      | $\eta = 0.07$                     | 94.5-296.8                    | 525-1436               | 525.48                         | 14.3                         | 1.9        |
| SO                   | $\eta = 0.001$                    | 94.5-297.7                    | 525-1436               | 2790.30                        | 14.4                         | 1.8        |
| Case 2               |                                   |                               |                        |                                |                              |            |
| NSGA-II (SR)         | 525.4-807.7                       | 1379-2461                     | 958.52                 | 15.6                           | 2.1                          |            |
| p-NSGA-II            | $\lambda_1 = 0 \lambda_2 = 1$     | 523.1-802.5                   | 1384-2468              | 2618.65                        | <b>15.4</b>                  | 2.0        |
|                      | $\lambda_1 = 0.5 \lambda_2 = 0.5$ | 527.7-809.6                   | 1381-2453              | 2526.68                        | 15.5                         | 1.9        |
|                      | $\lambda_1 = 0.8 \lambda_2 = 0.2$ | 527.2-809.6                   | 1381-2455              | 2568.56                        | 15.6                         | 1.8        |
|                      | $\lambda_1 = 1 \lambda_2 = 0$     | 527.1-771.6                   | 1395-2457              | 2683.52                        | 15.8                         | <b>1.7</b> |

with uncertain delays and driver's behavioral response," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 8, pp. 1548–1557, 2012.

- [19] X. Li and H. K. Lo, "An energy-efficient scheduling and speed control approach for metro rail operations," *Transportation Research Part B: Methodological*, vol. 64, pp. 73–89, 2014.
- [20] C. Sicre, A. P. Cucala, A. Fernández, and P. Lukaszewicz, "Modeling and optimizing energy-efficient manual driving on high-speed lines," *IEEE Transactions on Electrical and Electronic Engineering*, vol. 7, no. 6, pp. 633–640, 2012.
- [21] N. Zhao, Z. Tian, L. Chen, C. Roberts, and S. Hillmansen, "Driving strategy optimization and field test on an urban rail transit system," *IEEE Intelligent Transportation Systems Magazine*, 2020.
- [22] D. He, G. Lu, and Y. Yang, "Research on optimization of train energy-saving based on improved chicken swarm optimization," *IEEE Access*, vol. 7, pp. 121 675–121 684, 2019.
- [23] S. Lu, S. Hillmansen, T. K. Ho, and C. Roberts, "Single-train trajectory optimization," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 2, pp. 743–750, 2013.
- [24] Z. Pan, M. Chen, S. Lu, Z. Tian, and Y. Liu, "Integrated timetable optimization for minimum total energy consumption of an ac railway system," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 3641–3653, 2020.
- [25] X.-H. Zhao, B.-R. Ke, and K.-L. Lian, "Optimization of train speed curve for energy saving using efficient and accurate electric traction models on the mass rapid transit system," *IEEE Transactions on Transportation Electrification*, vol. 4, no. 4, pp. 922–935, 2018.
- [26] Z. Xiao, Q. Wang, P. Sun, B. You, and X. Feng, "Modeling and energy-optimal control for high-speed trains," *IEEE Transactions on transportation electrification*, vol. 6, no. 2, pp. 797–807, 2020.
- [27] Z. Xiao, Q. Wang, P. Sun, Z. Zhao, Y. Rao, and X. Feng, "Real-time energy-efficient driver advisory system for high-speed trains," *IEEE Transactions on Transportation Electrification*, vol. 7, no. 4, pp. 3163–3172, 2021.
- [28] X. Zhu, R. Zhang, W. Dai, Z. Zhang, and J. Li, "Performance and safety assessment of ato systems in urban rail transit systems in china," *Journal of transportation engineering*, vol. 139, no. 7, pp. 728–737, 2013.
- [29] S. Gao, H. Dong, Y. Chen, B. Ning, G. Chen, and X. Yang, "Approximation-based robust adaptive automatic train control: An approach for actuator saturation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1733–1742, 2013.
- [30] M. Domínguez, A. Fernández, A. Cucala, and P. Lukaszewicz, "Optimal

- design of metro automatic train operation speed profiles for reducing energy consumption,” *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 225, no. 5, pp. 463–474, 2011.
- [31] Q. Pu, X. Zhu, J. Liu, D. Cai, G. Fu, D. Wei, J. Sun, and R. Zhang, “Integrated optimal design of speed profile and fuzzy pid controller for train with multifactor consideration,” *IEEE Access*, vol. 8, pp. 152 146–152 160, 2020.
- [32] A. Fernandez-Rodriguez, A. Fernández-Cardador, A. P. Cucala, M. Domínguez, and T. Gonsalves, “Design of robust and energy-efficient atp speed profiles of metropolitan lines considering train load variations and delays,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 2061–2071, 2015.
- [33] M. Domínguez, A. Fernández-Cardador, A. P. Cucala, T. Gonsalves, and A. Fernández, “Multi objective particle swarm optimization algorithm for the design of efficient atp speed profiles in metro lines,” *Engineering Applications of Artificial Intelligence*, vol. 29, pp. 43–53, 2014.
- [34] W. ShangGuan, X.-H. Yan, B.-G. Cai, and J. Wang, “Multiobjective optimization for train speed trajectory in ctcs high-speed railway with hybrid evolutionary algorithm,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 2215–2225, 2015.
- [35] Z. Tian, N. Zhao, S. Hillmansen, C. Roberts, T. Dowens, and C. Kerr, “Smartdrive: traction energy optimization and applications in rail systems,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 7, pp. 2764–2773, 2019.
- [36] H. Cui, Y. Guan, and H. Chen, “Rolling element fault diagnosis based on vmd and sensitivity mckd,” *IEEE Access*, vol. 9, pp. 120 297–120 308, 2021.
- [37] X. Ran, X. Zhou, M. Lei, W. Tepsan, and W. Deng, “A novel k-means clustering algorithm with a noise algorithm for capturing urban hotspots,” *Applied Sciences*, vol. 11, no. 23, p. 11202, 2021.
- [38] H. Cui, Y. Guan, H. Chen, and W. Deng, “A novel advancing signal processing method based on coupled multi-stable stochastic resonance for fault detection,” *Applied Sciences*, vol. 11, no. 12, p. 5385, 2021.
- [39] P. Wang, A. Trivella, R. M. Goverde, and F. Corman, “Train trajectory optimization for improved on-time arrival under parametric uncertainty,” *Transportation Research Part C: Emerging Technologies*, vol. 119, p. 102680, 2020.
- [40] J. Yin, T. Tang, L. Yang, Z. Gao, and B. Ran, “Energy-efficient metro train rescheduling with uncertain time-variant passenger demands: An approximate dynamic programming approach,” *Transportation Research Part B: Methodological*, vol. 91, pp. 178–210, 2016.
- [41] J. Yin, D. Chen, and L. Li, “Intelligent train operation algorithms for subway by expert system and reinforcement learning,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 6, pp. 2561–2571, 2014.
- [42] X. Yang, A. Chen, B. Ning, and T. Tang, “A stochastic model for the integrated optimization on metro timetable and speed profile with uncertain train mass,” *Transportation Research Part B: Methodological*, vol. 91, pp. 424–445, 2016.
- [43] J. Szlapczynska and R. Szlapczynski, “Preference-based evolutionary multi-objective optimization in ship weather routing,” *Applied Soft Computing*, vol. 84, p. 105742, 2019.
- [44] Q. Kang, S. Feng, M. Zhou, A. C. Ammari, and K. Sedraoui, “Optimal load scheduling of plug-in hybrid electric vehicles via weight-aggregation multi-objective evolutionary algorithms,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 9, pp. 2557–2568, 2017.
- [45] H. Li, B. Wang, Y. Yuan, M. Zhou, Y. Fan, and Y. Xia, “Scoring and dynamic hierarchy-based nsga-ii for multiobjective workflow scheduling in the cloud,” *IEEE Transactions on Automation Science and Engineering*, 2021.
- [46] J. Molina, L. V. Santana, A. G. Hernández-Díaz, C. A. C. Coello, and R. Caballero, “g-dominance: Reference point based dominance for multiobjective metaheuristics,” *European Journal of Operational Research*, vol. 197, no. 2, pp. 685–692, 2009.
- [47] S. Jiang and S. Yang, “A strength pareto evolutionary algorithm based on reference direction for multiobjective and many-objective optimization,” *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 3, pp. 329–346, 2017.
- [48] Y. Hua, Q. Liu, K. Hao, and Y. Jin, “A survey of evolutionary algorithms for multi-objective optimization problems with irregular pareto fronts,” *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 2, pp. 303–318, 2021.
- [49] Q. Kang, X. Song, M. Zhou, and L. Li, “A collaborative resource allocation strategy for decomposition-based multiobjective evolutionary algorithms,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 12, pp. 2416–2423, 2018.
- [50] J. Tang, G. Liu, and Q. Pan, “A review on representative swarm intelligence algorithms for solving optimization problems: Applications and trends,” *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 10, pp. 1627–1643, 2021.
- [51] Z. Zhao, S. Liu, M. Zhou, and A. Abusorrah, “Dual-objective mixed integer linear program and memetic algorithm for an industrial group scheduling problem,” *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 6, pp. 1199–1209, 2020.
- [52] N. Zhao, C. Roberts, S. Hillmansen, and G. Nicholson, “A multiple train trajectory optimization to minimize energy consumption and delay,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2363–2372, 2015.
- [53] R. C. Purshouse, K. Deb, M. M. Mansor, S. Mostaghim, and R. Wang, “A review of hybrid evolutionary multiple criteria decision making methods,” in *2014 IEEE congress on evolutionary computation (CEC)*. IEEE, 2014, pp. 1147–1154.
- [54] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: Nsga-ii,” *IEEE transactions on evolutionary computation*, vol. 6, no. 2, pp. 182–197, 2002.
- [55] H. Edelsbrunner, D. Kirkpatrick, and R. Seidel, “On the shape of a set of points in the plane,” *IEEE Transactions on information theory*, vol. 29, no. 4, pp. 551–559, 1983.
- [56] K. Deb and H. Gupta, “Searching for robust pareto-optimal solutions in multi-objective optimization,” in *International conference on evolutionary multi-criterion optimization*. Springer, 2005, pp. 150–164.