Weekly Train Timetabling Approach for High-speed Railway Lines

Bowen Nie [a, b, c], Lei Nie [b], Zhiyuan Lin [a], Ronghui Liu [a]*, and Huiling Fu [b]

[a] *Institute for Transport Studies, University of Leeds, Leeds, UK;* [b] *School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China;* [c] *Scientific & Technological Information Research Institute, China Academy of Railway Sciences Corporation Limited, Beijing, China*

*Corresponding author: Ronghui Liu; Email: r.liu@its.leeds.ac.uk; Phone: +44(0)113 343 5338; Location: Room 2.05, Institute for Transport Studies (34-40 University Road), Leeds, UK

Weekly Train Timetabling Approach for High-speed Railway Lines

Traditional train timetabling methods typically schedule trains for periods ranging from half an hour to a single day, often neglecting the fluctuations in passenger demand over an entire week. To overcome this limitation, this study proposes a weekly train timetabling (WTT) model that schedules trains for the entire week. To effectively implement the WTT model under practical high-speed railway (HSR) scenarios, an Estimate-Generation-Evaluate (EGE) solution process is introduced, incorporating a customised hierarchical train generation strategy. Testing the EGE process on Chinese HSR lines demonstrates its superior performance improvement over CPLEX. Compared to manually generated timetables, the weekly timetable produced by EGE enhances passenger travel speeds and better aligns train schedules with passenger demand patterns. Further comparisons between solutions for two typical HSR lines verify the universality and robustness of the proposed approach.

Keywords: train timetabling; high-speed railway; weekly timetable; passenger demand

Subject classification codes: include these here if the journal requires them

Acknowledgement

This research was supported by Railway Joint Fund Project of the National Natural Science Foundation of China (No. U2468224); the Fundamental Research Funds for the Central Universities (No. 2022JBQY005); and the research plan of China Railway（K2023X031).

1. Introduction

As a key mode of transportation in modern society, aligning High-Speed Railway (HSR) services with passenger demand fluctuations has long been a fundamental goal of HSR train scheduling and optimisation. The HSR train optimisation process is typically divided into sequential subproblems (Schöbel, 2017). At the strategic level, the line planning problem involves determining the origin and terminal stations, travel routes, operation frequencies, and train stops based on passenger demand. At the tactical level, the train timetabling problem specifies the departure and arrival times at each station the train passes or stops, as well as the specific departure times and travel durations for passengers. Finally, at the operational level, the vehicle and crew scheduling problem assigns vehicle units and crew to serve the passengers. Among these planning processes, the timetabling problem plays a critical role in providing customised services for passengers.

Passenger demand fluctuations vary between different origin and destination (OD) pairs. On a macro scale, annual peaks in passenger volumes can typically be observed at the start and end of national holidays on most HSR lines. Although these significant peaks prompt railway companies to adjust timetables and add more trains, such occurrences are rare, typically happening only once a year. The more frequent and critical fluctuation, however, occurs across different days of the week and periods of the day.

Within a single day, commuter travel between city centres and surrounding areas often creates distinct morning and evening peaks on short-distance HSR lines. Business travellers, on the other hand, typically depart later in the morning, creating smaller peaks during periods such as 8:00 a.m. to 10:00 a.m. on HSR lines connecting major metropolitan areas. These demand features are commonly considered while scheduling trains among a day (e.g. Şahin et al., 2020; Martin-Iradi & Ropke, 2022; Yao et al., 2023). Weekly demand fluctuations, which are often overlooked, also play a significant role on many HSR lines. Notable differences can be seen between weekday and weekend demand characteristics. On weekdays, particularly from Tuesday to Thursday, passenger demand is relatively stable due to regular work schedules. While on weekends, increased commuting demand, especially from individuals who live and work in different locations, can be observed resulting in peaks on Friday and Sunday afternoons, as well as Saturday mornings. Additionally, tourist cities may also experience a surge in demand during the weekend as tourists travel to these locations.

Railway companies commonly adjust train frequencies and stops manually to align timetables with weekly demand fluctuations. For instance, on regional lines in European countries, weekday morning peak timetables often include more short-distance trains connecting smaller cities to major urban centres. In China, where passenger flow on HSR lines is often saturated, timetables are typically adjusted based on a saturated schedule designed for peak periods. However, OD pairs along an HSR line vary in distance and city attributes, leading to different proportions of commuter, tourist, and business travellers, each with distinct demand fluctuation patterns. As a result, simply increasing or decreasing the number of trains is not always sufficient to accommodate these varying demand patterns. Moreover, adjusting train stops could potentially disrupt the regularity of the timetable and diminish the service quality for other OD pairs.

The limitations of manual adjustments urgently require weekly train timetabling methods, while previous train timetabling studies generally focus on periodic timetabling with fixed train cycles (Zhang & Nie, 2016; Yao et al., 2023), flexible train cycles (Yan & Rob, 2019), non-periodic timetabling (Cacchiani et al., 2010; Robenek et al., 2016), and combined periodic and non-periodic timetabling (Şahin et al., 2020) that consider intraday demand fluctuations. The typical timetabling models in these studies are based on integer decision variables (Zhang et al., 2021; Zhao et al. 2021; Yuan et al., 2023) for the specific departure and arrival times, or time-space networks with binary decision variables to the time-space routes of trains (Brännlund et al., 1998; Liao et al., 2021; Zhang et al., 2022). Headway constraints between departure and arrival times of any two trains (Tian & Niu, 2020) and overtaking forbidding constraints between two neighbouring stations (Caprara et al., 2002) are always included in the models to guarantee the timetable feasibility. These studies provide a classic framework for modelling the train timetabling problem, including decision variables and critical constraints, upon which our approach is built and customised. While previous literature has widely explored periodic and non-periodic trains, our model incorporates both types along with a third category—**weekly trains**, which operate only on selected days of the week. This feature introduces intentional variation between daily timetables, allowing better alignment with weekly fluctuations in passenger demand, which constitutes a key distinction from classical timetabling approaches.

Many studies have tailored their methods to address demand variations across different origin-destination (OD) pairs or periods of the day. For instance, some studies introduce direct service frequency constraints to guarantee service level for each OD (Yan & Rob, 2019), robust constraints to address demand uncertainty (Cacchiani et al., 2020), and train capacity constraints according to demand fluctuation (Wang et al., 2020). In terms of optimisation objectives, passenger travel indicators are often incorporated to enhance the alignment between train services and demand fluctuations, such as minimising passenger travel time (Kaspi & Raviv, 2013; Borndörfer et al., 2016), passenger waiting time at stations (Niu & Zhou, 2013), and intermediate train dwelling time (Yue et al., 2016). Minimising the difference between obtained timetable and an ideal timetable generated totally based on passenger preference (Cacchiani et al., 2010; Cacchiani et al., 2015) can also be regarded as objective. Building on research that estimates passenger travel choice behaviours (Fischer, 2020; Wu et al, 2022), some studies form the integrated optimisation models to jointly schedule trains and determine passenger travelling routes (Schmidt & Schöbel, 2014; Xu et al., 2021; Martin-Iradi & Ropke, 2022) to improve the demand-matching level of obtained timetable. Rolling stock schedules (Zhou et al., 2022; Liu et al., 2024) and stop plans (Dong et al, 2020; Yuan et al., 2023; Yao et al., 2023) are also jointly optimised in some studies to further enhance passenger service. Additionally, other measures, such as flow control (Hu et al., 2025) and pricing and seat allocation (Yuan et al., 2025), have been considered in recent research to improve service efficiency and demand satisfaction. Building on the various methods proposed to enhance service, this study integrates an estimation of passenger behaviours before timetable generation, followed by an evaluation of the generated timetable based on detailed passenger preferences after timetable generation. This estimation-generation-evaluation approach further improves the alignment between train schedules and passenger preferences.

To solve the train timetabling models under larger-scale practical scenarios, if it is constructed based on time-space networks, Lagrangian relaxation (earlier proposed in Brännlund et al., 1998, and widely used by recent researches, e.g., Yue et al., 2016; Xu et al., 2021; Liao et al., 2021) or alternating direction method of multipliers (ADMM, Zhang et al., 2022; Yao et al., 2023) is commonly applied to decompose it into submodels for routing individual trains. Developing heuristic algorithms is also effective if customised search rules are designed, such as inserting or deleting trains (Dong et al. 2020), adjusting headway time intervals between adjacent trains (Zhou et al., 2022; Yuan et al., 2022), and changing the running time of trains (Shi et al., 2023). These tailored solution strategies significantly reduce the computational effort required to solve large-scale timetabling problems. In this study, we draw upon various methods from the existing literature and propose a customized heuristic strategy to generate train schedules following a hierarchical approach, with its efficiency in solving the weekly train timetabling problem demonstrated through extensive case studies.

While previous studies have established classic methods for modelling and solving train timetabling problems, the consideration of complex weekly demand fluctuations presents the following challenges that limit the efficiency of these widely used approaches: First, the increased number of trains over seven days results in an extensive number of decision variables and constraints, posing significant challenges to find solutions. Second, widely applied demand modelling approaches struggle to manage the larger scale of passenger data and the more complex fluctuation characteristics observed over a week. Using a single passenger demand indicator in the objective function inevitably introduces bias, while an integrated optimisation of timetabling and passenger estimation results in models that are too large to solve efficiently.

To address these challenges, we propose a customised Weekly Train Timetabling (WTT) approach that targets weekly demand fluctuations. Our earlier research (Nie et al., 2022) focused on the weekly line planning problem and proposed a genetic algorithm. Building on this, the current study introduces a WTT model using an Estimation-Generation-Evaluation (EGE) method to determine the specific departure and arrival times for all trains at all stations over a week. To enhance solution efficiency, a hierarchical train generation strategy is incorporated into the EGE. The contributions of this research are summarised as follows

- **Proposing weekly train operational modes based on weekly passenger demand fluctuations:** Previous studies have typically focused on single periods or days, with trains operating in either periodic or non-periodic modes. In this study, we introduce weekly operational modes for trains: periodic trains that operate consistently across all periods and days; daily trains that run once or several times each day; and weekly trains that operate only on specific days of the week. Among these three operational modes, periodic and daily trains ensure the timetable regularity, while weekly trains provide the flexibility needed to accommodate fluctuations in weekly demand.

- **Developing an EGE method for solving the WTT problem:** The EGE method accounts for refined passenger travel behaviours by estimating the passenger attraction for each train before timetabling. Once the weekly timetable is determined, an evaluation model assesses the solution quality comprehensively. Numerical experiments demonstrate that the resulting weekly timetable is superior to those manually generated by HSR operators, as evidenced by indicators such as passenger travel speed, the number of intermediate stops, and passenger departure times.
- **Incorporating a hierarchical train generation strategy in the EGE:** The large number of trains in real-world scenarios poses significant challenges in generating a weekly timetable. To address this, we introduce a customised hierarchical train generation strategy, designed according to the weekly operational modes, and incorporate it into the EGE framework. This strategy prioritises different train categories and follows a three-stage method to generate the weekly timetable. Case studies verify that this strategy significantly reduces computation time while maintaining stable solution quality compared to the commercial solver CPLEX.

The remainder of this paper is organised as follows: Section 2 describes the WTT problem and explains the weekly train operational modes in weekly timetables. Section 3 details the mathematical models used in the EGE method to solve the WTT problem, while Section 4 discusses the hierarchical train generation strategy. Section 5 presents numerical experiments based on a Chinese HSR case study, and Section 6 concludes the study and suggests future research directions.

2. Problem description

The Weekly Train Timetabling (WTT) problem focus on the planning period of a whole week, the continuous daily train operation time is divided into specific periods, such as 8:00 a.m. to 10:00 a.m. and 10:00 a.m. to 12:00 p.m. Train schedules within different days of the week should vary to align with weekly passenger fluctuations, such as the weekly commuting peaks at Friday, Saturday, and Sunday mentioned in Section 1. Additionally, timetables across different periods of the day should reflect daily demand variations.

While scheduling trains for a week, a certain level of regularity must also be maintained. To ensure the consistency of daily timetables, a proportion of trains within each period should retain identical schedules to facilitate passenger convenience and regular transfers. Furthermore, timetables between adjacent days should not vary significantly, enabling passengers to easily recall train schedules.

To formulate this complex train scheduling problem, the trains in the weekly timetable are classified into three operational modes based on the regularity of their schedules across different periods and days:

- **Periodic trains:** These trains operate in every period each day, with the interval between their departure and arrival times at the same station equal to the length of the period. For example, a periodic train might depart at 8:05 a.m., 9:05 a.m., 10:05 a.m., and so forth, continuing at the same times throughout each day of the week. Periodic trains have been the primary focus in recent studies on periodic timetabling.
- **Daily trains:** These trains operate every day but only during specific periods. For example, a daily train might depart at 8:05 a.m. each day of the week. These trains are considered non-periodic in research related to periodic train timetabling.

- **Weekly trains:** These trains operate only on specific days, such as at 8:05 a.m. from Monday to Friday. Their schedules can be flexibly adjusted without any regular restrictions.

At critical time slots—such as integral time points like 8:00 a.m. and 9:00 a.m. within each period—faster trains with fewer stops, connecting two terminals of an HSR line, are scheduled to ensure regular direct services between distant major cities. These slots are referred to as **mandatory slots**, and the trains operating at these slots are termed **mandatory trains**.

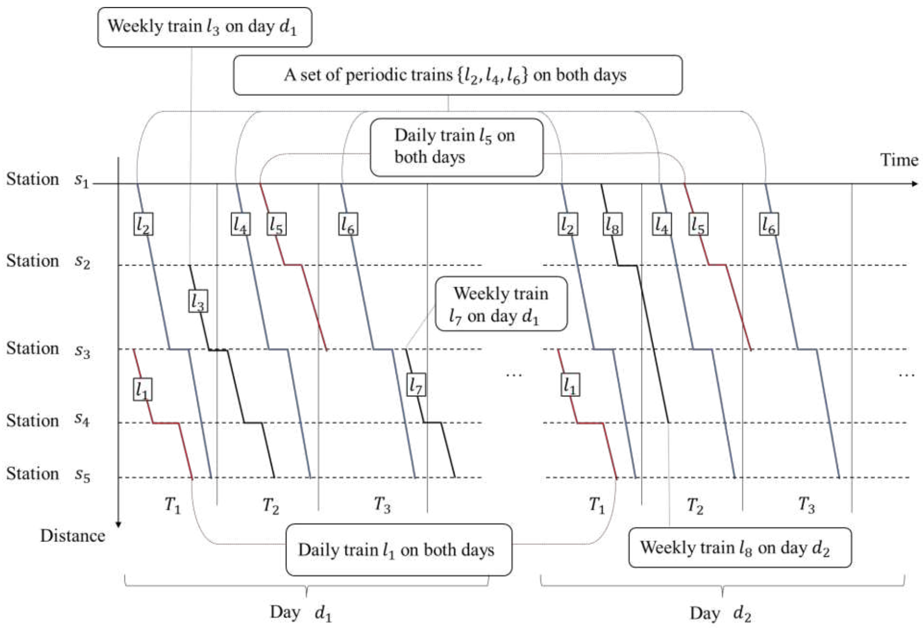

Figure 1. The periodic trains, daily trains, and weekly trains in a weekly timetable.

Figure 1 illustrates a typical weekly timetable, with trains labelled from l_1 to l_8. Trains l_2, l_4, and l_6 on both days d_1 and d_2 constitute a set of periodic trains. The departure or arrival time difference between these trains in adjacent periods, such as l_2 in period T_1 and l_4 in period T_2, is equal to the length of the period. Daily trains l_1 and l_2 run on both days d_1 and d_2, during period T_1 and T_2, respectively. Weekly trains l_3, l_7, and l_8 operate in different periods and on different days.

Based on these concepts, the WTT problem takes as input a **weekly line plan**, developed in our earlier work (Nie et al., 2022). In that study, a customised genetic algorithm was applied to select trains from a candidate set, resulting in a weekly line plan that determines the origin and terminal stations, intermediate stops, service frequencies, and operational modes of all trains across each day and period. With these elements fixed, the WTT problem then focuses on determining the precise departure and arrival times of the trains at all stations, thereby generating their specific time–space trajectories.

To formulate the mathematical model and algorithm for WTT problem, the following assumptions are presented. Assumptions 1, 2, and 3, established based on practical train operation rules, simplify the decisions regarding train travel speeds between adjacent stations and train dwelling times at various stations. Assumptions 4 and 5 are derived from passenger preference surveys.

Assumption 1. All the trains are of the same maximum speed, which depends on the technical condition of the HSR line. Therefore, with determined train stops in input weekly line plan, running times between any two adjacent stations are considered as fixed values. This assumption simplifies the problem by eliminating the need to optimise running times between stations.

Assumption 2. Stations along the HSR line are classified into different levels based on the volume of trains originating, terminating, and stopping at the station, as well as the passenger demand. The train dwelling time at these stations varies within specific upper and lower bounds according to the station levels. At higher-level stations with sufficient track resource, faster trains are allowed to overtake slower trains. This bounded dwelling time reduces the solution space and helps to control the total number of trains stopping at each station by limiting the maximum allowable dwelling time.

Assumption 3. The seating capacities of trains are fixed, and overloading is prohibited. This assumption establishes a constraint linking passenger flow with train schedules during the optimisation process.

Assumption 4. When the weekly timetable is established, passengers are assigned to their most preferred trains. Indicators related to their travel routes are used to evaluate the quality of the weekly timetable. Passenger preferences for train services are influenced by factors such as departure times, intermediate dwelling times, and transfer time consumption. Passengers are allowed to shift their departure to adjacent periods if no service is available in their desired departure period but are not permitted to depart on different days.

Assumption 5. Before solving the WTT problem, passengers are also assigned to train lines based on the input weekly line plan. In this context, passenger-related factors for each train in the line plan, such as the load factor, are used to represent the passenger attraction of each train, serving as input for train scheduling. Since the specific departure and arrival times of trains are not yet determined, passenger preferences are influenced by departure periods, intermediate stops, and the need for transfers. This passenger assignment approach was also examined in our earlier work (Nie et al., 2022).

Assumption 6. Passengers travelling between a specific origin-destination pair (OD) within the same period constitute a passenger group, and the preferences of passengers in a same group are of the same. Assumptions 4, 5, and 6 are designed to model passenger behaviours.

Assumption 7. Periodic trains of the same set operate only once per period. If the frequency of a periodic train is more than one in weekly line plan, multiple distinct periodic train sets should be established.

Assumption 8. Trains in input weekly line plan may be cancelled or depart in adjacent periods when their original departure periods become oversaturated, in this case a penalty is incorporated into the objective value. Assumptions 7 and 8 are introduced to ensure that train schedules adhere to their respective weekly operational modes.

Based on the description of the weekly timetable and the associated assumptions, the Weekly Train Timetabling (WTT) problem is defined as follows: Given the conditions of the HSR line—including the length of each period, the maximum speed of trains, the distance between stations, the categorisation of stations, and the upper and lower bounds of dwell times at various stations—alongside the weekly line plan, which specifies the origin and terminal stations, routes and stops, weekly operational modes, operational periods and days, and the operational frequencies of the trains, as well as the passenger demand data, the objective is to generate a weekly timetable that schedules trains across all seven days of the week and all periods within each day.

The generated weekly timetable must adhere to several constraints to ensure its feasibility: The departure and arrival time constraints for periodic trains within the same set, the consistency of schedules for daily trains within the same set, the safety headway constraints between the departure and arrival times of any two trains, and the running and dwell time constraints for each train.

3. Mathematical models for Estimation-Generation-Evaluation method (EGE)

In the weekly timetable, the schedules for periodic or daily trains remain identical across different days, allowing these trains to be consolidated into a single entity with only one set of decision variables for their departure and arrival times. In contrast, weekly trains, which operate on specific days, are represented as distinct entities with separate sets of decision variables for each day they operate. This consolidation effectively reduces the Weekly Train Timetabling (WTT) problem to a one-day timetabling problem, albeit with additional special headway constraints that govern the relationships between periodic trains, daily trains, and weekly trains.

As depicted in Figure 2, headway constraints are essential between the schedules of any two periodic or daily trains, as these trains operate every day, ensuring safe and efficient spacing between trains on the same line. However, such headway constraints are not necessary between the schedules of two weekly trains that operate on different days, for instance, one operating on weekdays and the other on weekends. By simplifying the WTT problem in this manner, it becomes more manageable while still accommodating the necessary constraints and maintaining the integrity of the weekly schedule.

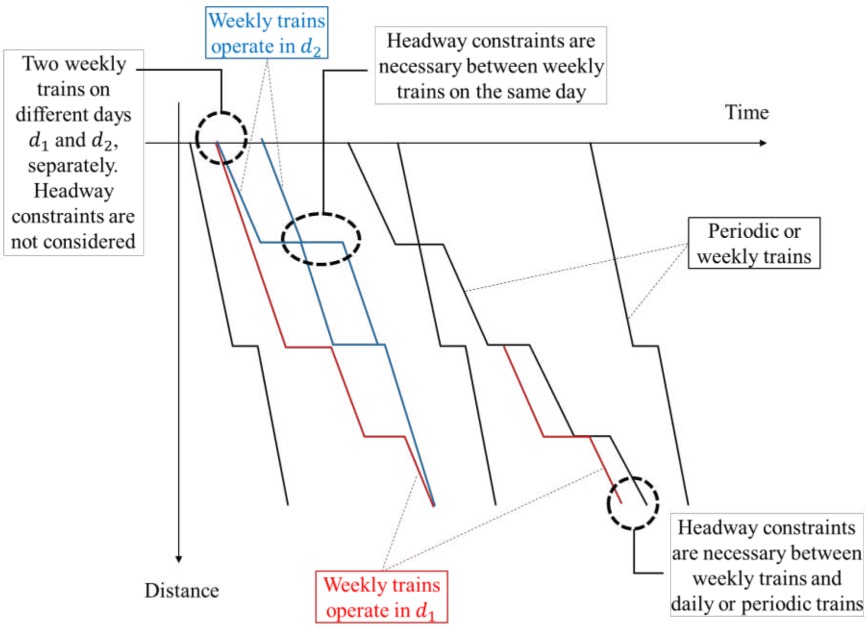

Figure 2. Special headway constraints in consolidated timetable.

Based on the consolidated timetable, this section proposes three models to form the Estimation-Generation-Evaluation (EGE) method for solving the WTT problem. The related mathematical expressions are shown in Table 1.

- **Estimation:** The passenger estimation model uses weekly line plan as input and initially assesses passenger travel behaviour based on the stops and periods of the trains before the weekly timetable is determined. It outputs the estimated passenger attraction factor for each train.
- **Generation:** The weekly timetabling model takes the weekly line plan and the estimated passenger attraction factors for each train as input to generate the weekly timetable using the modelling approach of the consolidated timetable.
- **Evaluation:** The timetable evaluation model takes the weekly timetable as input and simulates passenger travel behaviour. Indicators such as passenger travel speeds are used to evaluate the service of weekly timetable.

Table 1. Notations for mathematical models of the Estimation-Generation-Evaluation (EGE) method.

Notation	Definition
S	The set of stations s along the railway line.
L	The set of all the trains l (or its corresponding train line in the weekly line plan).
W	The set of passenger groups w. $W_{s_i s_j}$ denotes the set of passenger groups between station s_i and s_j.
T	The set of time unites t of each day, typically in minutes.
Q	The set of mandatory train operation slots q.
AL	The set of trains operate in generated weekly timetable. Then $(L - AL)$ denotes the set of cancelled trains.
S_l	The set of station that the train l passes through. Let $S_l = \{s_0, s_1, \dots, s_k\}$ denotes the sequence of stations indexed in the order.
T_l	The departure period of train l determined in weekly line plan.
OR_l	The feasible departure time of train l at its originating station, it includes period T_l and its adjacent periods.
DR_{ls}	The range of departure time of train l at station s.
AR_{ls}	The range of arrival time of train l at station s.

WR_{ls}	The range of dwelling time of train l at station s. If l not stop at s，$WR_{ls} = \{0\}$.
$T_0, T_1, T_2, \dots$	The sequence of different periods indexed in the order, e.g., 6:00-8:00, 8:00-10:00, …
T_w	The original departure period for passenger group w.
L_q	The set of alternative mandatory trains that can operate at mandatory slot q.
$L^P_{s_i s_{i+1}}, TL^P_{s_i s_{i+1}}$	The set of direct and transfer routes in passenger estimation model that pass the segment between s_i and s_{i+1}.
$L^T_{s_i s_{i+1}}, TL^T_{s_i s_{i+1}}$	The set of direct and transfer routes l in timetable evaluation model that pass the segment between s_i and s_{i+1}.
τ^T	The length of each period.
τ_q	The specific timestep of mandatory slot q.
DH_s	The safety headway for departure times between adjacent trains at station s.
AH_s	The safety headway for arrival times between adjacent trains at station s.
σ_l	The original station of train l
ρ_l	The terminal station of train l.
d_l	The number of days that train l operates. For example, it equals 7 if l is a periodic or daily train that runs every day of the week.
$\tau_{lss'}$	The travel time for train l between two adjacent stations s and s'.
$\theta_{ll'}$	Whether the train l and l' operate in the same day ($\theta_{ll'} = 1$) or not ($\theta_{ll'} = 0$).
VL_w	The volume of passenger group w.
$VL_{ls_i s_{i+1}}$	The seating capacity of train l between station s_i and s_{i+1}.
$u_{w\eta}, u_{wl}$	The matching utility of passenger group w travel through transfer route η or direct route l.
TR_w	The range of rational transfer time for passenger group w.
DT_i	The i-th set of daily trains. DT denotes the set of all the daily trains.
PT_i	The i-th set of periodic trains. PT denotes the set of all the periodic trains
LF_l	The estimated passenger attraction factor of train l.
PA_l	The penalty for the cancellation of train l.
PD_l	The penalty per minute for dwelling time of train l.
PS_l	The penalty for train l switching to another period for departure.
M	A maximum value

3.1 Passenger estimation model

The passenger estimation model is constructed similarly to the passenger routing model in our previous research (Nie et al., 2022), which was originally designed to evaluate the quality of a weekly line plan. It estimates the passenger attraction of trains within the line plan based on their departure periods, stops, and compositions. Output indicators, such as the load factors of trains, serve as evaluations of the passenger attraction ability of the trains and are used as inputs when scheduling trains to reflect their importance in the weekly timetable.

According to Assumption 6, the set W includes all passenger groups across different periods and days, covering all OD pairs. Each passenger group w consists of passengers travelling between a specific pair of stations, departing during a specific period on a specific day. To formulate the routes for different passenger groups, two types of decision variables are involved: m_{wl}^{P} are continuous variables representing the volumes of passengers of group w travelling along direct path $l \in L_w^P$, and $n_{w\eta}^{P}$ are continuous variables representing the volumes of passengers of group w travelling along transfer path $\eta \in TL_w^P$. The direct and transfer paths for passengers are determined based on the input weekly line plan. The set L_w^P includes all trains that stop at both the origin and destination stations and operate on the same day that passenger group w departs. The set TL_w^P includes all transfer routes. A transfer route can be formed by two trains only if operate on the same day as group w, and the connection at the transfer station is sufficiently close in time, for example, within the same period. This path formation process significantly reduces the number of alternative paths, making the model more computationally efficient.

The objective of the passenger estimation model is to maximise the total matching utility of all passengers, as formulated in (p1). Constraint (p2) set the maximum volume of travelling passengers, while constraint (p3) ensures that passenger volume does not exceed the seating capacity of the train. A detailed discussion of this model can be found in Nie et al. (2022).

$$\text{Max: } Z = \sum_{w,l \in L_w^P} m_{wl}^P \cdot u_{wl}^P + \sum_{w,\eta \in TL_w^P} n_{w\eta}^P \cdot u_{w\eta}^P \quad \text{(p1)}$$

$$\sum_{l \in L_w^P} m_{wl}^P + \sum_{\eta \in TL_w^P} n_{w\eta}^P \le VL_w \qquad \forall w \in W \quad \text{(p2)}$$

$$\sum_{w,l \in L_{s_i s_{i+1}}^P \cap L_w^P} m_{wl}^P + \sum_{w,\eta \in TL_{s_i s_{i+1}}^P \cap TL_w^P} n_{w\eta}^P \le VL_{l s_i s_{i+1}} \qquad \forall l \in AL;\ s_i, s_{i+1} \in S_l \quad \text{(p3)}$$

$$m_{wl}^P \in [0, p_w] \qquad \forall w \in W;\ l \in L_w^P \quad \text{(p4)}$$

$$n_{w\eta}^P \in [0, p_w] \qquad \forall w \in W;\ \eta \in TL_w^P \quad \text{(p5)}$$

3.2 Weekly timetabling model

The decision variables of weekly timetabling model are listed as follows:

- x_{ls} is an integer variable representing the arrival time of train l at station s.
- y_{ls} is an integer variable representing the departure time of train l at station s.
- z_{lq} is a binary variable, equals 1 if train l is served as mandatory train at the mandatory slot q.
- r_l is a binary variable, equals 1 if train l is cancelled.
- p_l is a binary variable, equals 1 if train l departs within its original departure period.

The range of departure and arrival times of train l at each station is restricted by the feasible originating time range $OR_l = DR_{l\sigma_l} = [\text{Min}(T_l) - \tau_T, \text{Max}(T_l) + \tau_T]$. With fixed upper and lower dwelling times at intermediate stations, the upper and lower departure and arrival times at stations can be determined as shown in Figure 3. Assume that train l originates at the earliest time in OR_l and dwells for the shortest time, the earliest departure and arrival times can be computed. Assume that train l originates at the latest time in OR_l and dwells for the longest time at each station s, the latest departure and arrival times can be computed. This computation method is formulated as (g1) and (g2).

$$DR_{ls_i} = \left[\text{Min}(OR_l) + \sum_{j=0}^{i-1} \tau_{ls_js_{j+1}} + \sum_{j=1}^{i} \text{Min}\left(WR_{ls_j}\right), \text{Max}(OR_l) + \sum_{j=0}^{i-1} \tau_{ls_js_{j+1}} + \sum_{j=1}^{i} \text{Max}\left(WR_{ls_j}\right)\right] \quad i \in [1, k-1] \qquad \text{(g1)}$$

$$AR_{ls_i} = \left[\text{Min}(OR_l) + \sum_{j=0}^{i-1} \tau_{ls_js_{j+1}} + \sum_{j=1}^{i-1} \text{Min}\left(WR_{ls_j}\right), \text{Max}(OR_l) + \sum_{j=0}^{i-1} \tau_{ls_js_{j+1}} + \sum_{j=1}^{i-1} \text{Max}\left(WR_{ls_j}\right)\right] \quad i \in [1, k] \qquad \text{(g2)}$$

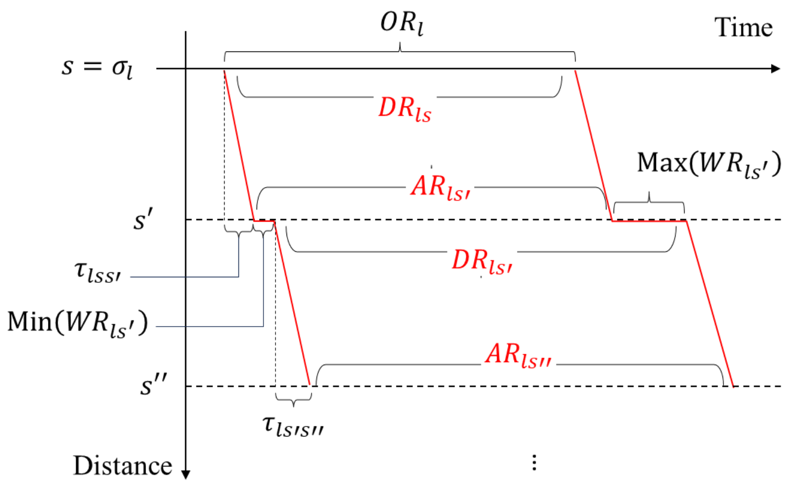

Figure 3. Computing the ranges for train departure and arrival times.

The objective value of weekly timetabling model is formulated as (g3), which consists of three penalty components: The penalty for intermediate dwelling times (PD_l per minute); the penalty for train cancellation (PA_l per train); and the penalty for train departing at adjacent periods (PS_l per train). The parameter d_l represents the penalty weights assigned to trains based on the number of days it operates during the week. Constraints are formulated as (g4) ~ (g19). The restricted range of dwelling times and travelling times are formulated as (g4) and (g5). The departure and arrival headway constraints are initially formulated as (g6) and (g7) and are linearised as described in

Appendix 1. Constraints (g8) and (g9) restrict the departure and arrival times of periodic trains of the same periodic set. Constraint (g10) ensure that only one train is selected for each mandatory slot, and constraints (g11) and (g12) restrict the departure time of train l if it is selected as a mandatory train. The rationale of the M-method applied in (g6), (g7), (g11), and (g12) is discussed in **Appendix 2**. Constraints (g13) and (g14) restrict the departure time of train l within T_l if its departure period is not changed ($p_l = 1$). Constraint (g15) ensures that if one periodic train within a set is cancelled, all trains in that set are also cancelled to maintain the periodicity of the weekly timetable. Formulations (g16) through (g20) set the lower and upper bounds for the decision variables.

$$\text{Min: } Z = \sum_{l,s\in(S_l-\sigma_l-\rho_l)} d_l(y_{ls} - x_{ls}) + PA_l \cdot \sum_l d_l \cdot r_l + PS_l \cdot \sum_{l\in L-PT} d_l(p_l - 1) \quad \text{(g3)}$$

$$y_{ls} - x_{ls} \in WR_{ls} \qquad \forall l, s \in (S_l - \sigma_l - \rho_l) \quad \text{(g4)}$$

$$x_{ls_{i+1}} - y_{ls_i} = \tau_{ls_is_{i+1}} \qquad \forall l;\ s_i, s_{i+1} \in S_l \quad \text{(g5)}$$

$$|y_{ls} - y_{l's}| + M(r_l + r_{l'} + \theta_{ll'}) \geq DH_s \quad \text{(g6)}$$

$$|x_{ls} - x_{l's}| + M(r_l + r_{l'} + \theta_{ll'}) \geq AH_s \quad \text{(g7)}$$

$$x_{l_{j+1}s} - x_{l_js} = \tau_T \quad \forall PT_i;\ l_j, l_{j+1} \in PT_i \quad \text{(g8)}$$

$$y_{l_{j+1}s} - y_{l_js} = \tau_T \quad \forall PT_i;\ l_j, l_{j+1} \in PT_i \quad \text{(g9)}$$

$$\sum_{l\in L_q} z_{lq} = 1 \quad \forall q \in Q \quad \text{(g10)}$$

$$x_{l\sigma_l} + M \cdot (1 - z_{lq}) \geq \tau_q \quad \forall q \in Q, l \in L^Q \quad \text{(g11)}$$

$$x_{l\sigma_l} - M \cdot \left(1 - z_{lq}\right) \leq \tau_q \quad \forall q \in Q, l \in L^Q \tag{g12}$$

$$x_{l\sigma_l} + M \cdot p_l \geq \text{Min}(T_l) \quad \forall l \in L - PT \tag{g13}$$

$$x_{l\sigma_l} - M \cdot p_l \leq \text{Max}(T_l) \quad \forall l \in L - PT \tag{g14}$$

$$r_{l_j} = r_{l_{j+1}} \quad \forall PT_i;\ l_j, l_{j+1} \in PT_i \tag{g15}$$

$$x_{ls} \in DR_{ls} \quad \forall l, s \in (S_l - \rho_l) \tag{g16}$$

$$y_{ls} \in DR_{ls} \quad \forall l, s \in (S_l - \sigma_l) \tag{g17}$$

$$z_{lq} \in \{0,1\} \qquad \forall q \in Q, l \in L^Q \tag{g18}$$

$$r_l \in \{0,1\} \qquad \forall l \tag{g19}$$

$$p_l \in \{0,1\} \qquad \forall l \in L - PT \tag{g20}$$

3.3 Timetable evaluation model

The obtained weekly timetable forms direct routes L_w^T and transfer routes TL_w^T for each passenger group w in a similar way as L_w^P and TL_w^P in Section 3.1. Since the departure and arrival times of trains are determined in weekly timetable, transfer routes are formed according to the specific transfer times additionally. Routes with infeasible transfer times are rejected.

Method to compute the passenger travelling utilities u_{wl}^T and $u_{w\eta}^T$ is described as follows: Set $\hat{u}_{wl}$ as the sum of all the intermediate dwelling times (in minute) of train l between original and terminal stations, and $\hat{u}_{w\eta}$ as the sum of all the intermediate dwelling times (in minute) of train l and l' that form the transfer route η and include the transfer time of η (in minute, with a larger weight due to the more inconvenience caused by transferring). Next, add a penalty to both $\hat{u}_{wl}$ and $\hat{u}_{w\eta}$ if the departure time of route is not within the original departure period of passenger group w. This penalty can either be uniform across all non-preferred periods or vary according to the deviation from the original departure period. In this study, we apply a significant penalty regardless of the period difference, as our objective is to ensure that all passengers depart within their original preferred periods. Finally, $u_{wl}^{T} \in DU_{w}$ is determined as the normalised value of $-\hat{u}_{wl}$ for different l, and $u_{w\eta}^{T} \in TU_{w}$ is determined as the normalised value of $-\hat{u}_{w\eta}$ for different η. Generally, $\mathrm{Min}(DU_{w}) \geq \mathrm{Max}(TU_{w})$.

Decision variables m_{wl}^{T} are continuous one representing the volumes of passengers of group w travelling along direct path $l \in L_{w}^{T}$, and $n_{w\eta}^{T}$ are continuous variables representing the volumes of passengers of group w travelling along transfer path $\eta \in TL_{w}^{T}$. The objective, and constraints of timetable evaluation model are formulated as (t1) ~ (t5), which is similar to the passenger estimation model in Section 3.1.

$$\text{Max: } Z = \sum_{w,l \in L_{w}^{T}} m_{wl}^{T} \cdot u_{wl}^{T} + \sum_{w,\eta \in TL_{w}^{T}} n_{w\eta}^{T} \cdot u_{w\eta}^{T} \quad \text{(t1)}$$

$$\sum_{l \in L_{w}^{T}} m_{wl}^{T} + \sum_{\eta \in TL_{w}^{T}} n_{w\eta}^{T} \leq VL_{w} \qquad \forall w \in W \quad \text{(t2)}$$

$$\sum_{w,l \in L_{s_i s_{i+1}}^{T} \cap L_{w}^{T}} m_{wl}^{T} + \sum_{w,\eta \in TL_{s_i s_{i+1}}^{T} \cap TL_{w}^{T}} n_{w\eta}^{T} \leq VL_{l s_i s_{i+1}} \qquad \forall l \in AL;\ s_i, s_{i+1} \in S_l \quad \text{(t3)}$$

$$m_{wl}^{T} \in [0, p_{w}] \qquad \forall w \in W;\ l \in L_{w}^{T} \quad \text{(t4)}$$

$$n_{w\eta}^{T} \in [0, p_{w}] \qquad \forall w \in W;\ \eta \in TL_{w}^{T} \quad \text{(t5)}$$

4. EGE incorporating hierarchical train generation strategy

In the Estimation-Generation-Evaluation (EGE) method, solving the weekly timetabling model is most challenging due to the inclusion of integer decision variables, while in other models all the decision variables are continuous. To address this challenge a hierarchical train generation strategy is incorporated into EGE. The outline of EGE is illustrated in Figure 4.

First, the passenger estimation model route passengers based on the input weekly line plan, the values of decision variables are denoted as $\overline{m}_{wl}^{P}$ and $\overline{n}_{w\eta}^{P}$. To measure the passenger attraction of each train line, the load factors are served as passenger attraction factor LF_l. The distance between station s_1 and s_2 is denoted as $DIS_{s_1 s_2}$, and the computation of LF_l is formulated as (e1).

$$LF_l = \sum_{s_i, s_{i+1} \in S_l} \frac{DIS_{s_i s_{i+1}}}{DIS_{\sigma_l \rho_l} \cdot VL_{l s_i s_{i+1}}} \cdot \left(\sum_{w, l \in L_{s_i s_{i+1}}^{P} \cap L_w^P} \overline{m_{wl}^{P}} + \sum_{w, \eta \in TL_{s_i s_{i+1}}^{P} \cap TL_w^P} \overline{n_{w\eta}^{P}} \right) \quad \text{(e1)}$$

The passenger attraction factor LF_l, computed in the passenger estimation model, is then used as critical input in the weekly timetabling model. The penalty values for different trains l in the objective (g3) of weekly timetabling model are computed based on LF_l, formulated as $PD_l = \mathrm{F}^{PD}(LF_l)$, $PA_l = \mathrm{F}^{PA}(LF_l)$, and $PS_l = \mathrm{F}^{PS}(LF_l)$. These penalties reflect the potential risk of passenger service reduction associated with the train adjustments and cancellation. Sections 4.1, 4.2, and 4.3 describe the hierarchical train generation process, which decomposes the original optimisation problem—considering all trains as input—into multiple stages. In each stage, a smaller subset of trains is selected as input and scheduled using commercial solvers to solve the optimisation model presented in Section 3.2, with different parameters applied.

Finally, through evaluating the output weekly timetable, demand-related indicators are used to measure passenger service level. For example, the total matching utility, which is the value of the objective function, indicating the extent to which passengers can select the optimal travel paths. Other detailed indicators, such as passenger travel speed, can also be implemented into discussion. An iterative method could be developed as follows, though it is not included in this study due to its current instability in efficiency: The weekly timetable evaluation can also generate LF_l values for scheduled trains, which can then be used as inputs for the weekly timetabling model in the next iteration.

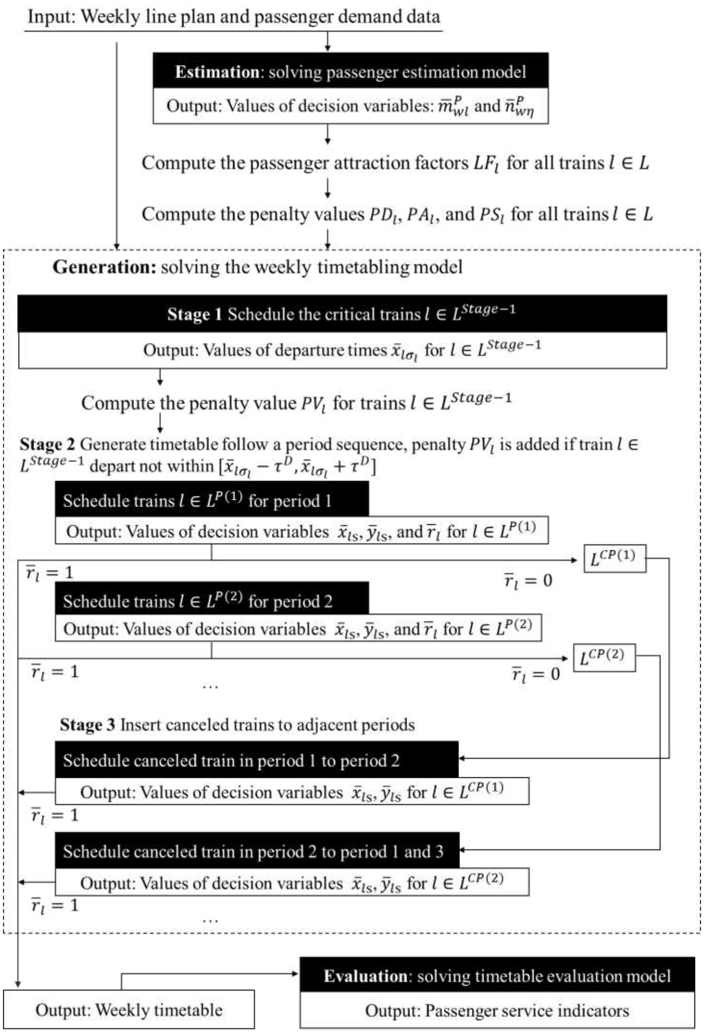

Figure 4. The procedure for Estimation-Generation-Evaluation (EGE) method incorporating hierarchical train generation strategy.

4.1 Scheduling the critical trains

The first stage of hierarchical train generation is scheduling the critical trains in weekly timetable, which containing three types of trains: all alternative mandatory trains for each mandatory slot, all periodic trains, and daily or weekly trains with higher passenger attractions. The steps of scheduling the critical train are given as follows.

Step 1 Let $L^{Stage-1}$ represent the set of critical trains that are scheduled in this Stage-1. And select trains from L into $L^{Stage-1}$. Order the trains $l \in L$ according to LF_l, and for each l:

if l is the alternative mandatory train for a mandatory slot $q \in Q$ ($l \in L_q$), add l into $L^{Stage-1}$;

if l is a periodic train ($l \in L^P$), add l into $L^{Stage-1}$;

if l is a daily train ($l \in L^D$) or weekly train ($l \in L^W$), and the number of trains in $L^{Stage-1}$ is less than a pre-given value ($|L^{Stage-1}| < VT^{Stage-1}$), add l into $L^{Stage-1}$.

Step 2 Compute the penalties PD_l, PA_l, and PS_l of trains $l \in L^{Stage-1}$ with linear formulations (e2), (e3), and (e4).

$$PD_l = \mathrm{F}^{PD}(LF_l) = a_l^{PD} \cdot LF_l + b_l^{PD} \tag{e2}$$

$$PA_l = \mathrm{F}^{PA}(LF_l) = a_l^{PA} \cdot LF_l + b_l^{PA} \tag{e3}$$

$$PS_l = \mathrm{F}^{PS}(LF_l) = a_l^{PS} \cdot LF_l + b_l^{PS} \tag{e4}$$

Step 3 Regarding the trains $l \in L^{Stage-1}$ and their penalties PD_l, PA_l, and PS_l as input, solve the weekly timetabling model described in Section 3.2 with commercial solver. The values of decision variables $x_{l\sigma_l}$ is recorded as $\bar{x}_{l\sigma_l}$, which is the expected departure values for critical trains.

In this process, the set $L^{Stage-1}$ must include at least all periodic trains, while the maximum number of trains $VT^{Stage-1}$ in the critical set should not be excessively large. An overly large critical train set may prevent the commercial solver from solving the weekly timetabling problem within a reasonable time frame, resulting in suboptimal scheduling of critical trains. This balance is crucial for maintaining computationally feasibility while effectively covering essential and high-demand services in the timetable. Values of a_l^{PD}, a_l^{PA}, and a_l^{PS}as well as b_l^{PD}, b_l^{PA}, and b_l^{PS} for different trains l and are set according to the distance and the number of stations they pass and stop at to measure the potential impact of their adjustments on passengers.

4.2 Scheduling the trains following the period sequence

This stage schedule all the trains by solving local timetables for each period. This method neglects the global information and leads to only local optimal solutions, which is the motivation of scheduling critical trains firstly, with demand fluctuation across different periods considered.

In this section, critical trains in $L^{Stage-1}$ are assumed to departure within a certain time range according to their schedules determined in Section 1, or a new penalty PV_l will be added in the objective of the weekly timetabling model, which is formulated as (e5). New type of decision variables e_l is incorporated for critical trains $l \in L^{Stage-1}$. When e_l equals 1, train l departs within a range of $[\bar{x}_{l\sigma_l} - \tau^D, \bar{x}_{l\sigma_l} + \tau^D]$, where τ^D represents the flexibility for adjusting the departure time of critical trains. Additional constraints (e6) and (e7) are included in the weekly timetabling model to restrict the relationship between e_l and $x_{l\sigma_l}$. The rationale of the M-method is discussed in **Appendix 2**. These constraints ensure that critical trains are positioned within time slots determined based on the global passenger demand information, aiming to achieve a higher level of demand matching.

$$\text{Min: } Z = PD_l \cdot \sum_{l,s\in(S_l-\sigma_l-\rho_l)} d_l(y_{ls} - x_{ls}) + PA_l \cdot \sum_l d_l r_l - PV_l \cdot \sum_{l\in L^{Stage-1}} d_l e_l \quad \text{(e5)}$$

$$x_{l\sigma_l} - M \cdot (1 - e_l) \le \bar{x}_{l\sigma_l} + \tau^D \qquad \forall l \in L^{Stage-1} \qquad \text{(e6)}$$

$$x_{l\sigma_l} + M \cdot (1 - e_l) \geq \bar{x}_{l\sigma_l} - \tau^D \qquad \forall l \in L^{Stage-1} \tag{e7}$$

The process for scheduling trains following the period sequence is given as follows.

Step 1. Compute the penalties PD_l, PA_l, and PS_l of trains $l \in L$ with (e2), (e3), and (e4), respectively. Compute the penalties PV_l for trains $l \in L^{Stage-1}$ with (e8).

$$PV_l = \mathrm{F}^{PV}(LF_l) = a_l^{PV} \cdot LF_l + b_l^{PV} \tag{e8}$$

Step 2. Order the periods according to a certain rule (for example, from the latest to the earliest within a day). Let $L^{P(i)}$ represents the set of trains operating in the period indexed by i in this sequence. For example, $L^{P(1)}$ corresponds to the last period of the day.

Steps 3 to 7 describe the procedure to schedule all the trains in the period indexed by i:

Step 3. For each train $l \in L^{P(i)}$, if l is also included in L_j^{PT}. Add the trains $l' \in L_j^{PT} \cap L^{P(i-1)}$ and $l' \in L_j^{PT} \cap L^{P(i+1)}$ to $L^{P(i)}$ if they have not been scheduled yet.

Step 4. Compute the departure and arrival time ranges DR_{ls} and AR_{ls} with (g1) and (g2). If $i \geq 2$, adjust the upper bounds of DR_{ls} and AR_{ls} are adjusted according to (e9) and (e10) based on the schedules of trains in adjacent period indexed $(i-1)$.

$$\mathrm{Max}(AR_{ls}) = \mathrm{Min}_{l \in L^{P(i-1)}}(\bar{x}_{ls}) - AH_s \qquad \forall l, s \in (S_l - \rho_l) \tag{e9}$$

$$\mathrm{Max}(DR_{ls}) = \mathrm{Min}_{l \in L^{P(i-1)}}(\bar{y}_{ls}) - DH_s \qquad \forall l, s \in (S_l - \sigma_l) \tag{e10}$$

Step 5. If a mandatory slot q falls within the time range of this period ($\tau_q \in T^{P(i)}$), add the trains $l \in L_q$ to $L^{P(i)}$and include decision variables z_{lq} along with constraints (g11) and (g12) for slot q.

Step 6. With the trains $l \in L^{P(i)}$ and their penalties PD_l, PA_l, PS_l, and PV_l (only for $l \in L^{Stage-1} \cap L^{P(i)}$) as input, a commercial solver is applied to solve the weekly timetabling model with formulation (e5) as the objective and include additional constraints (e6) and (e7). In this step, all trains within the selected period are scheduled in the output solution. The resulting values of the decision variables for train l at station s are denoted as $\bar{x}_{ls}$ and $\bar{y}_{ls}$. These values remain fixed and will not be adjusted when scheduling trains in subsequent periods. If the timetable capacity within the period is insufficient to accommodate all planned trains, a subset of trains in this period will be cancelled. The set of cancelled trains is denoted as $L^{CP(i)}$.

Step 7. For each train $l \in L^{P(i)} - L^{CP(i)}$, if l is also included in L_j^{PT}. Determine the schedules for other trains $l' \in L_j^{PT} \cap L^{P(k)}$ with (e11) and (e12)

$$\bar{x}_{l's} = \bar{x}_{ls} - (k - i) \cdot \tau^T \qquad \forall l \in L_j^{PT} \cap L^{P(i)};\ l' \in L_j^{PT} \cap L^{P(k)} \qquad \text{(e11)}$$

$$\bar{y}_{l's} = \bar{y}_{ls} - (k - i) \cdot \tau^T \qquad \forall l \in L_j^{PT} \cap L^{P(i)};\ l' \in L_j^{PT} \cap L^{P(k)} \qquad \text{(e12)}$$

Potential conflicts while scheduling trains following period sequence is described as follows. Corresponding methods for addressing them during the procedure is also outlined.

First, given that some trains may extend into adjacent periods, scheduling periodic trains at the earliest times of a period may violate the safety headway intervals between the latest departing trains of the current period and the earliest departing trains of the next period. For instance, in Figure 5, where trains l_1 and l_2 are scheduled to depart early in the current solving period and next period, respectively. Train l_3, however, departs late in the current solving period and extends to the next period. In this case, safety headway should be satisfied between trains l_2 and l_3. To handle this potential conflict, in **Step 3** when solving a period concluding periodic trains, other periodic trains departing in adjacent periods should also be scheduled simultaneously, with their safety headway constraints considered in weekly timetabling model.

Second, model may fail to schedule mandatory trains due to the other trains scheduled close to the mandatory slot. For example, in Figure 6, if train l_1 is scheduled too close to the mandatory slot q, it may violate the headway constraints between any potential mandatory trains l_2 and l_3 at this slot. Therefore, in **Step 5** if a single mandatory slot may span two adjacent periods, its mandatory train should be determined in the prioritised period in the sequence.

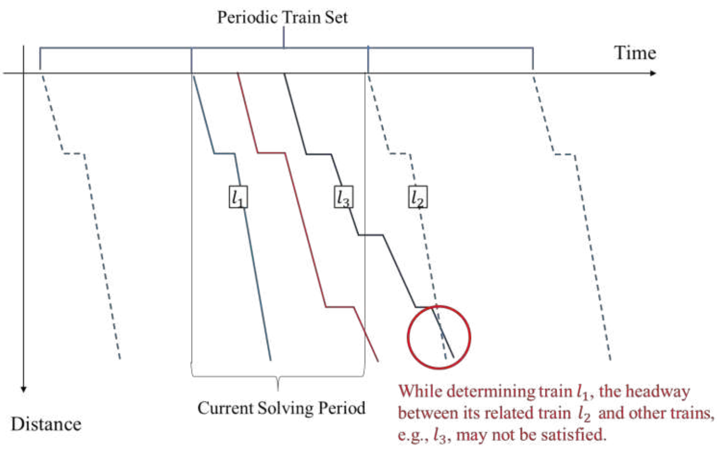

Figure 5. Potential conflicts between two periods while solving local timetables.

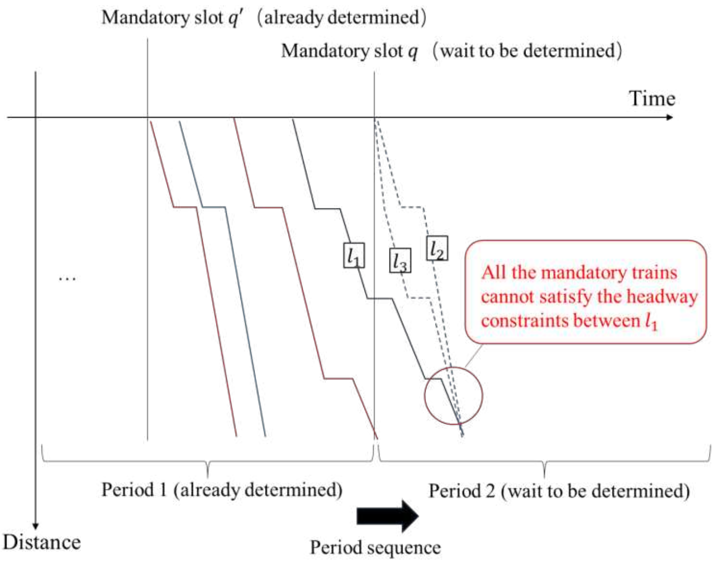

Figure 6. Potential conflicts of mandatory slots while solving local timetables.

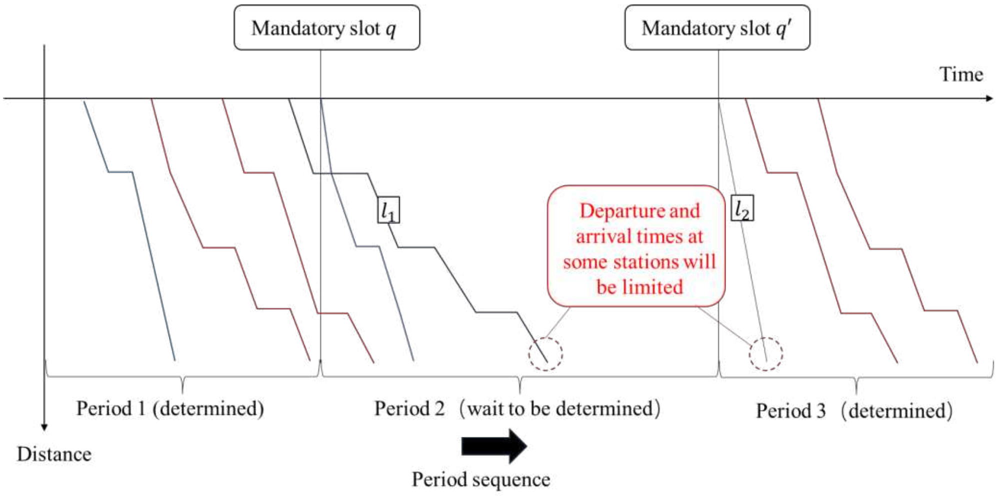

Figure 7. The train operation time range of a period may be limited by scheduled trains in adjacent periods.

Third, if trains in adjacent periods are both scheduled, the range for departure and arrival times at some stations may be reduced or even eliminated. For example, in Figure 7 the latest train l_3 in period 1 significantly intrudes into the departure time range of period 2. However, trains in period 2 cannot extend into period 3 to achieve a larger operating time range due to the scheduled mandatory train l_3, leading to a significant compression of the departure and arrival time range of period 2 trains. Therefore, in **Step 2** if there are fewer trains in later time periods, the period sequence should be constructed from the earliest to the latest one of a day. Conversely, if there are fewer trains in earlier periods, or if trains departing in the latest periods are prone to intruding into subsequent maintenance windows, the sequence could be constructed from the latest to the earliest one.

4.3 Rescheduling the cancelled trains

This stage aims to reschedule trains that are cancelled in previous stage described in Section 4.2. By expanding the range of decision variables for departure and arrival times, feasible routes may be found in adjacent periods. The original periods and its adjacent periods of train l is added into OR_l, then DR_{ls} and AR_{ls} can be computed by (g1) and (g2). The process for rescheduling the trains cancelled in previous stage is given as follows.

Step 1. Select the period indexed by i if its scheduled trains number less than a certain value ($\left|L^{P(i)}\right| - \left|L^{CP(i)}\right| \leq NL$).

Following steps 2 to 4 reschedule cancelled trains departing at the adjacent periods indexed by $(i-1)$ and $(i+1)$.

Step 2. Let the set $L^{S3P(i)} = L^{P(i)} - L^{CP(i)} + L^{CP(i-1)} + L^{CP(i+1)}$, in which $(L^{CP(i-1)} + L^{CP(i+1)})$ represents the cancelled trains in adjacent periods that are ready to be scheduled in period i in stage 3, and $(L^{P(i)} - L^{CP(i)})$ represents the already scheduled trains.

Step 3. Reset the departure and arrival time ranges DR_{ls} and AR_{ls} for trains $l \in L^{S3P(i)}$:

If $l \in L^{P(i)} - L^{CP(i)}$, $DR_{ls} = \bar{x}_{ls}$ and $AR_{ls} = \bar{y}_{ls}$. The scheduled trains won't change its departure and arrival times.

If $l \in L^{CP(i-1)}$, use (g1) and (g2) to compute DR_{ls} and AR_{ls}, in which $OR_l = [\text{Min}(T_{P(i)}), \text{Max}(T_{P(i-1)})]$. $T_{P(i)}$ represents the range of timesteps for period indexed by i in the period sequence.

If $l \in L^{CP(i+1)}$, use (g1) and (g2) to compute DR_{ls} and AR_{ls}, in which $OR_l = [\text{Min}(T_{P(i+1)}), \text{Max}(T_{P(i)})]$.

Step 4. With the trains $l \in L^{S3P(i)}$ and adjusted DR_{ls} and AR_{ls} as input, solve the weekly timetabling model. The values of decision variables are denoted as $\bar{x}_{ls}$, $\bar{y}_{ls}$, and $\bar{r}_l$. If $\bar{r}_l > 0$, this operated train is moved out of $L^{CP(i)}$.

Step 5. After all the periods in the sequence are computed, all scheduled trains are output as the weekly timetable.

5. Numerical experiment

Based on the typical Shanghai-Nanjing high-speed railway (HSR) line, we first construct smaller-scale scenarios to compare the solution quality between the Estimation-Generation-Evaluation (EGE) method and the commercial solver CPLEX, demonstrating the solution efficiency of EGE. Subsequently, practical passenger demand data and the weekly line plan from Nie et al. (2022) are used as inputs for a large-scale practical scenario. This allows us to compare the EGE-generated weekly timetable with a manually generated timetable and assess EGE's performance. Finally, the EGE is applied to the Beijing-Shanghai HSR line, which represents a typical long-distance route distinct from the Shanghai-Nanjing line. By comparing the solutions for these two HSR lines, we demonstrate the universality of the EGE method.

The railway condition of Shanghai-Nanjing HSR is illustrated in Figure 8. Stations are classified into three levels: major stations of level 1, larger intermediate stations of level 2, and smaller intermediate stations of level 3. The test computer runs on Windows 11 operating system, with an CPU of AMD R9 7490H, a main frequency of 4.00GHz, 8 cores, 16 threads, and 16GB of memory.

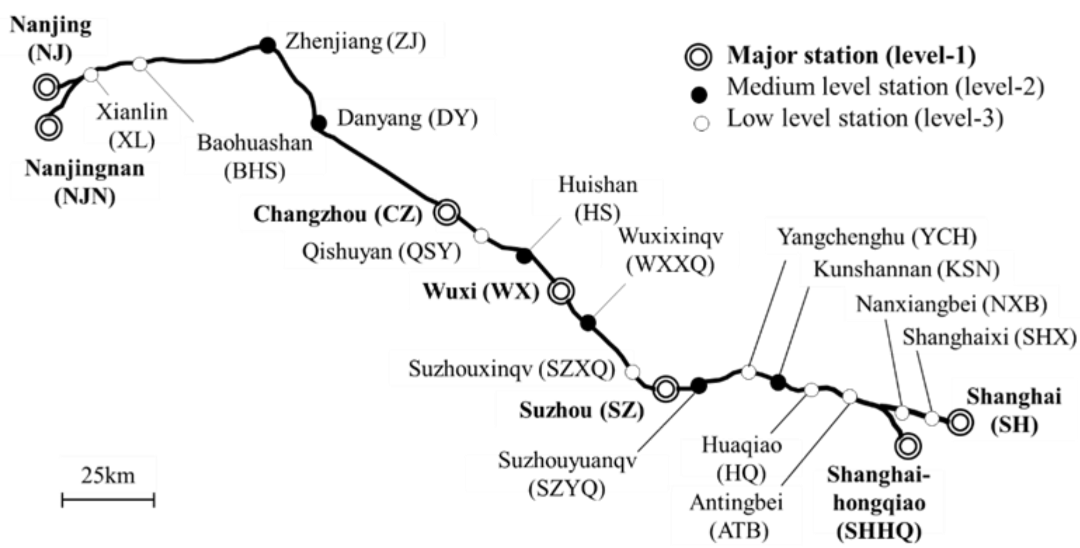

Figure 8. Condition of Nanjing-Shanghai high-speed railway line.

5.1 Solution analysis based on small-scale scenarios

Five smaller-scale scenarios containing from 50 to 100 trains are constructed The input weekly line plans for these scenarios are derived by removing a portion of trains from the original plan developed in Nie et al. (2022). Efforts were made to maintain similar proportions of periodic, daily, and weekly trains across the scenarios; however, these manually generated line plans do not match passenger demand as effectively as the original. The primary objective of comparing these scenarios is to evaluate the performance of the "Generation" stage of the EGE method, which employs a hierarchical train generation strategy. The commercial solver CPLEX is used as benchmark, with a maximum solution time to 7,200 seconds. The objectives and solution gaps for the hierarchical strategy and CPLEX are documented in Table 2.

While solving the smallest scenario (scenario 1), CPLEX can obtain the optimal solution with negligible time consumption. However, as the number of trains increase, the solution performance of CPLEX decreases significantly. In scenario 3 and 4, CPLEX fails to find solution with a small gap within 7,200-second limit. And in scenario 5, CPLEX cannot find any feasible solution within 7,200 seconds and only a lower bound value can be estimated.

For the hierarchical train generation strategy in the EGE, the computation times for all the scenarios are shorter than that of CPLEX. However, the hierarchical strategy divides the original model into multiple stages, leading to only local optimal solutions instead of the global ones. Therefore, while solving smaller scenarios 1 and 2, the solution gaps times of hierarchical strategy are larger than those of CPLEX. As the number of trains increases, the solution gaps become much smaller compared to those obtained by CPLEX, demonstrating a significant efficiency improvement while solving larger-scale weekly timetabling models.

Table 2. Solution comparison between hierarchical strategy and CPLEX

Scenarios		1	2	3	4	5
Num. of trains in the consolidated timetable		51	86	103	124	150
Num. of periodic trains in the consolidated timetable		6	18	18	24	30
Num. of daily trains in the consolidated timetable		21	23	28	29	43
Num. of trains in original weekly line plan		244	394	457	538	700
CPLEX	Computation time for the first solution with gap less than 5%	117s	2,159s	>7,200s	>7,200s	>7,200s
	Total computation time	356s	7,200s	7,200s	7,200s	--
	Obj. value of the best solution	4,000	6,455	23,297	72,178	--
	Gap of the best solution	0.00%	1.04%	67.57%	88.19%	--
	Estimated lower bound value	4,000	6,388	7,555	8,587	12,343
Hierarchical train generation strategy	Computation time	2s	29s	158s	612s	3,525s
	Obj. value	4,261	75,67	9,063	10,532	17,236
	Gap	6.12%	15.58%	16.63%	18.47%	28.39%

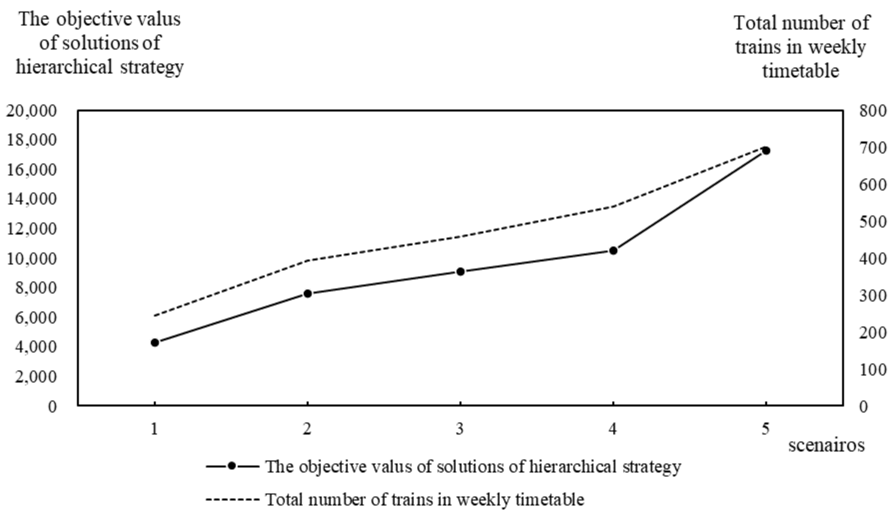

Figure 9. The changing trend of objective values of HTG strategy across different scenarios.

Since the hierarchical strategy cannot directly estimate the lower bound of the solution, gaps in Table 2 are computed based on the lower bounds estimated by CPLEX. Therefore, in scenario 5 when CPLEX fails to find good solutions, this estimation of lower bound may also be biased. Figure 9 illustrates the comparison between the train number and the changes in the objective values obtained by the hierarchical strategy across the five scenarios. The objective of the weekly timetabling model is strongly correlated with the number of trains. Hence, in scenarios 1 to 4, the trends of the two lines in Figure 9 are similar, indicating consistent solution quality. Due to an additional penalty value of totally 8 trains rescheduled to adjacent periods in scenario 5, a significant increase in the objective value can be observed between scenarios 4 and 5.

Combining the trends observed in the curves depicted in Figure 9 with the gaps for the solutions in Table 2, it becomes evident that the hierarchical strategy consistently outperforms CPLEX when solving weekly timetabling model in the process of EGE method under larger-scale scenarios. Moreover, the solution quality of the hierarchical strategy remains consistent across different scales of cases.

5.2 Solution analysis based on large-scale piratical scenarios

With the practical passenger demand data and weekly line plan of the Nanjing-Shanghai HSR line as inputs, the implementation details of the hierarchical train generation strategy are discussed in Section 5.2.1 and 5.2.2. In section 5.2.3, the EGE-obtained weekly timetable is compared to the manually generated timetable, which is currently applied in this practical scenario.

5.2.1 Implementation details related to stage 1 in the hierarchical strategy

With the implementation of stage 1 in hierarchical strategy, two parameters playing critical roles are analysed in detail:

(1) The schedule adjustment duration τ^D for critical trains in stage 2.

(2) The penalty value PV_l, which will be added into objective if the critical trains l departure not in the given time range. The value of PV_l can be set as a certain proportion of the penalty PA_l.

In Table 3, a timetable generated without involving stage 1 is recorded as ID 1. The other solutions are generated with the values of PV_l/PA_l set as 0.7, 0.5, 0.3, 0.1, and 0.05; and with the values of τ^D set as 0, 10, 20, 40, and 80. Generally, involving stage 1 leads an increase in computation time. However, if appropriate values are set for PV_l and τ^D, the passenger matching utility of solution with stage 1 incorporated can be 17.23% (the data of ID 20) higher than the solution without stage 1. This increase can be attributed to the following aspects.

(1) **Increase in the number of trains and seat kilometres.** This trend is particularly evident in IDs 15 and 20 in Table 3. Without stage 1, periodic trains will be scheduled at suboptimal times in stage 2, making it difficult to insert enough number of trains in the subsequent periods. While by introducing the stage 1 and setting appropriate values for PV_l and τ^D, both the number of trains and the total seat kilometres increase. This offers passengers more diverse travel options and greater transport capacity thus increase the total passenger matching utility.

(2) **Increase in passengers travelling within their original departure periods.** This trend is particularly evident in IDs 9, 20, 21, and 26 in Table 3. Although

maintaining passengers' original departure periods is not explicitly defined as an objective in the weekly timetabling model, it is achieved by prioritizing critical trains with higher passenger demand in the hierarchical strategy. When critical trains are scheduled to operate within their original periods, a higher proportion of passengers are accommodated within their preferred departure times, leading to a higher passenger matching utility.

(3) **Slight increase in penalties for the solutions.** This trend is particularly evident in IDs 4 and 9 in Table 3. Since the train penalties is the objective of weekly timetabling model, a large proportion of solutions with higher matching utilities typically reflect lower train penalties. However, in figure 10 which prints the solutions generated with various PV_l and τ^D (the one generated without the stage 1 is marked in red as a benchmark), solutions can both achieve high values of demand matching utility when the penalties for the solutions are lower (below 2.5×10^4) or higher (above 5.0×10^4). Most solutions with higher matching utilities (above 2.8×10^6) fall within the penalty values of 2.0×10^4 to 3.0×10^4, as indicated by the red dashed circles. This suggests a certain degree of bias between the specific passenger service level and the objective of the model. Therefore, the directly obtained optimal solution not necessary being the solution with best passenger matching utility. While the hierarchical strategy proposed in this study can further enhance the demand matching utility when appropriate PV_l and τ^D values are applied, with the cost of a slightly increased in the objective value.

Table 3. Solution indicators with different values of PV_l and τ^D.

ID	Whether stage 1 is included	PV_l $/PA_l$	τ^D	Computation time (min)	Total train penalty ($\times 10^4$)	Total demand matching utility ($\times 10^6$)	Total number of trains	Total seat kilometres ($\times 10^8$)	Number of passengers departing at original period

1	not included	--	--	86	2.04	2.62	759	1.71	64.3%
2			0	105	4.11	2.30	728	1.64	64.1%
3			10	107	4.68	2.50	734	1.66	67.6%
4		0.7	20	101	5.32	**2.69**	762	1.72	65.4%
5			40	101	2.81	2.62	754	1.71	66.3%
6			80	102	2.15	2.65	770	1.75	63.8%
7			0	110	3.52	2.44	735	1.67	64.1%
8			10	104	5.37	2.54	750	1.69	66.7%
9		0.5	20	103	3.11	**2.67**	758	1.72	69.9%
10			40	111	2.44	**2.83**	795	1.80	67.0%
11			80	112	2.08	2.60	758	1.71	64.3%
12			0	102	4.06	2.38	737	1.66	64.5%
13			10	102	2.61	2.53	763	1.72	66.0%
14	included	0.3	20	114	2.92	2.32	713	1.63	64.0%
15			40	110	2.36	**3.01**	803	1.82	66.9%
16			80	106	3.45	2.18	711	1.61	58.8%
17			0	102	3.38	2.59	753	1.69	66.1%
18			10	103	4.69	2.41	749	1.69	66.9%
19		0.1	20	102	4.93	2.56	733	1.71	68.7%
20			40	110	2.26	**3.08**	804	1.83	69.3%
21			80	105	2.03	**2.89**	784	1.78	68.0%
22			0	102	3.62	2.50	738	1.68	65.0%
23			10	111	4.49	2.51	750	1.71	68.8%
24		0.05	20	102	2.28	2.48	730	1.64	67.6%
25			40	105	7.00	2.04	687	1.58	63.3%
26			80	106	2.42	**2.79**	788	1.78	68.5%

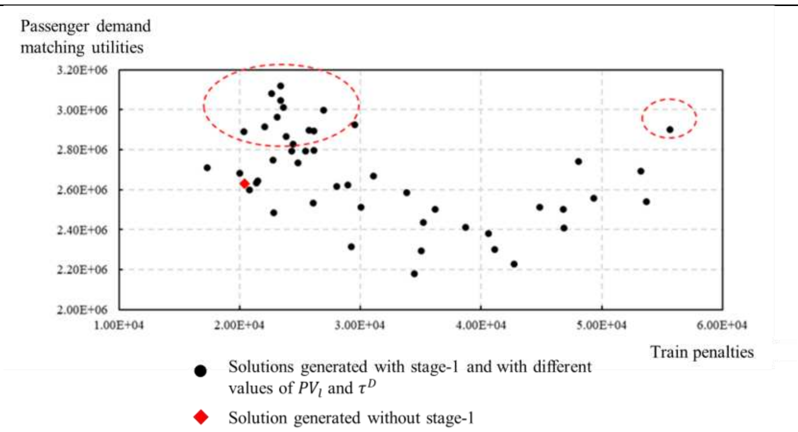

Figure 10. Train penalties and demand matching utilities of solutions with different values of PV_l and τ^D.

The records in Table 3 reveal the crucial influence of PV_l and τ^D on the optimisation effort of the hierarchical strategy. In practical scenarios, PV_l/PA_l is typically less than 1, and τ^D should be less than the length of each period. According to Table 3, solutions generally exhibit higher quality when PV_l/PA_l is between 0.1 and

0.5. Therefore, the values of 0.1, 0.3, and 0.5 for PV_l/PA_l is tested, and the values of τ^D is set between 0 and 90, the changing trends of passenger demand matching utilities are illustrated in Figure 11.

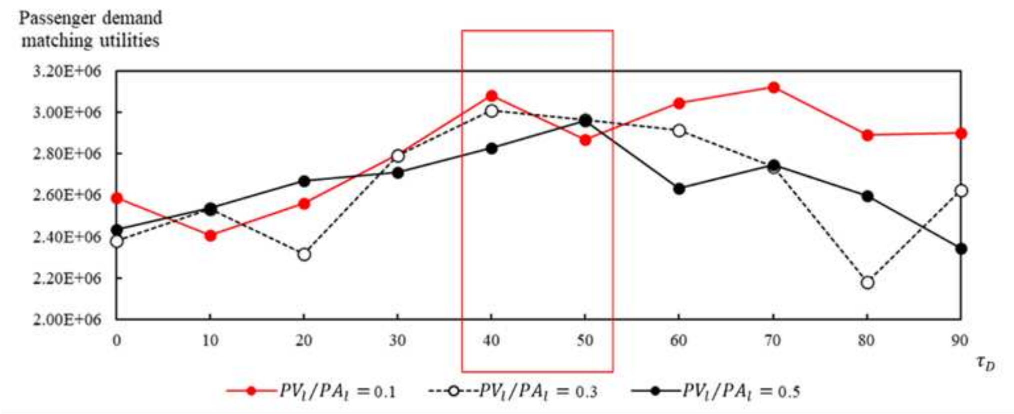

Figure 11. Passenger demand matching utilities changing trends with different values of PV_l and τ^D.

When τ^D is less than 30, the demand matching utilities of the three series are generally low and exhibit an upward trend. When τ^D =40 or 50, all three series achieve good solutions. As τ^D increases further, the demand matching utilities for PV_l/PA_l =0.3 and 0.5 series tend to decrease overall, while the $PV_l/PA_l = 0.1$ series further achieved the highest demand matching utility at $\tau^D = 70$. Referring to the overall trend changes in other series, $PV_l/PA_l = 0.1$ and $\tau^D = 40$ are set as considered reasonable values and will be utilised in subsequent experiments.

5.2.2 Optimisation effort related to stage 3 in the hierarchical strategy

To reveal the optimization effort of stage 3, two alternative strategies are compared. The first is the "2-stage strategy", in which stage 3 is omitted and trains cancelled in stage 2 are not reconsidered. The second, called the "Improved 2-stage strategy", also omits stage 3 but take into account cancelled trains in adjacent periods if these periods have already been solved.

With more trains included in Stage 2, the computation time for the Improved 2-stage strategy increases by approximately 16–23 minutes compared to the 2-stage strategy. However, when Stage 3 is introduced, the computation time of Stage 2 is reduced, resulting in a slight overall decrease in total computation time—approximately 5–7 minutes less than the Improved 2-stage strategy.

While the Improved 2-stage strategy demonstrates a 9.7% higher demand matching utility compared to the 2-stage strategy, the 3-stage hierarchical strategy further improves demand matching utility by 18.9% over the Improved 2-stage strategy. This clearly highlights the optimisation benefits of incorporating stage 3.

To further demonstrate the differences among these strategies, Figure 12 shows the number of trains cancelled in each period. The hierarchical strategy results in significantly fewer cancellations than the other two strategies. Notably, the 2-stage and Improved 2-stage strategies both cancel 91 trains between 16:00 and 18:00, while the peak cancellation for the hierarchical strategy is only 54 trains, indicating a better balance of service.

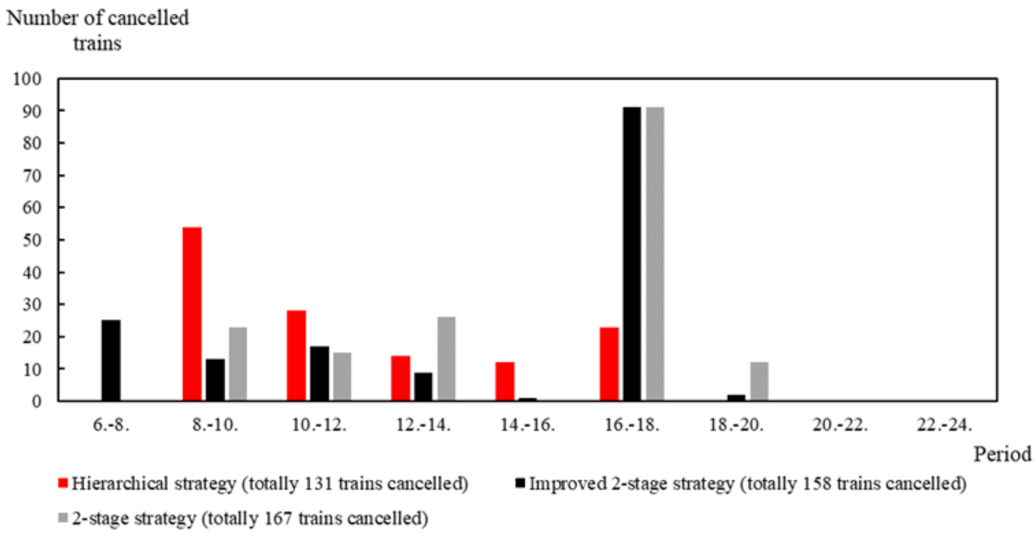

Figure 12. The cancellation distribution across periods in the three methods regarding stage 3 differently.

5.2.3 Solution comparison: weekly timetable and manually generated timetable

The optimisation of the proposed weekly timetabling approach is demonstrated by comparing it with the manually generated timetable, as indicated by the metrics recorded in Table 4. The passenger matching utility of the manually generated timetable is higher, as the trains generally operate at slower speeds. This allows passengers to more easily select routes that approximate the optimal ones, thereby yielding higher matching utilities. However, a closer examination of the detailed indicators reveals several advantages associated with the weekly timetable.

(1) **The increase in train travel speed is both significant and hierarchically distributed.** As illustrated in Figure 13, the peak distribution of train travel speeds differs markedly between the manual and weekly timetables. Trains in the weekly timetable predominantly operate at speeds exceeding 200 km/h, whereas those in the manual timetable are primarily within the 120-160 km/h range. This significant improvement in speed enhances the service level for passengers, particularly at major stations.

Furthermore, the weekly timetable exhibits a hierarchical distribution of travel speeds. A secondary peak is observed in the 140-160 km/h range, and it includes a greater number of trains operating below 100 km/h compared to the manual timetable. This hierarchical structure categorises trains into “last direct trains”, “common trains”, and “slower all-stopping trains”, thereby accommodating the diverse travel needs of passengers across various ODs.

Table 4. Indicator comparison between weekly timetable and manually generated timetable.

	Manually generated timetable	Weekly timetable
Passenger demand matching utility ($\times 10^6$)	3.42	3.08

Average travel speed of trains (km/h)	142.17	210.71
Average number of intermediate stops of passengers	1.97	0.89
Average travel speed of passengers (km/h)	158.41	211.03
Average train load factor	68.23%	71.10%
Number of mandatory trains per day	10	15
Average load factor of mandatory trains	95.08%	97.13%
Proportion of passengers travelling in their original departure periods	70.14%	69.26%
Proportion of passengers travelling in their original departure periods and adjacent periods	84.42%	93.13%
Total number of trains	882	804

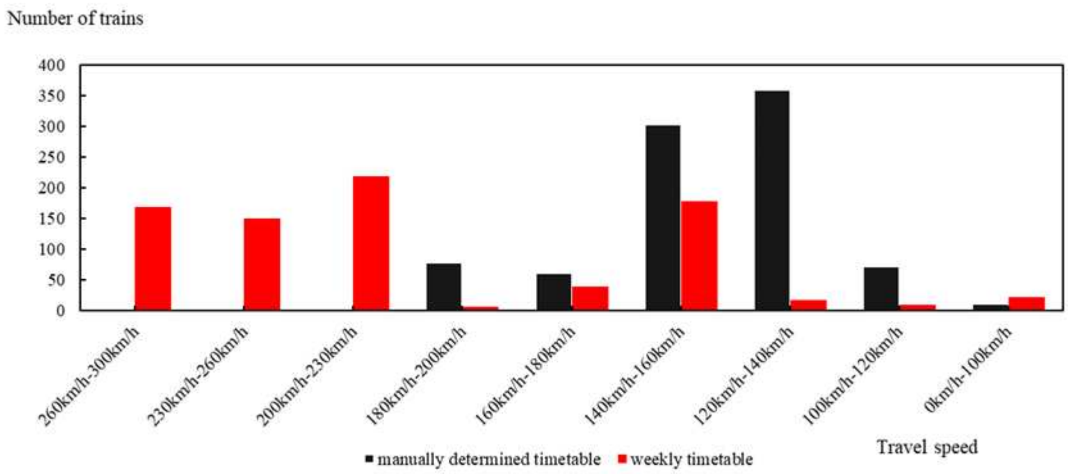

Figure 13. The travel speed distribution of trains in manual and weekly timetables.

(2) **Passenger attraction for both fast trains and mandatory trains is significantly greater.** The average load factors of trains across various speed ranges are depicted in Figure 14. In the weekly timetable, trains operating at speeds exceeding 260 km/h exhibit the highest load factors, indicating the strongest passenger attraction. With the introduction of differentiated pricing strategies, these high-speed trains could be designated as premium services, allowing for higher ticket prices to enhance the profitability of railway companies while offering additional services to passengers. Trains operating at

speeds of 160-200 km/h in the manual timetable also show elevated load factors, although their speeds are considerably lower.

The weekly timetable includes a greater number of mandatory trains, most of which achieve a 100% load factor, surpassing those in the manual timetable. These essential trains in the weekly timetable could also be marketed as premium services in mandatory slots, facilitating higher ticket prices and enhanced passenger offerings.

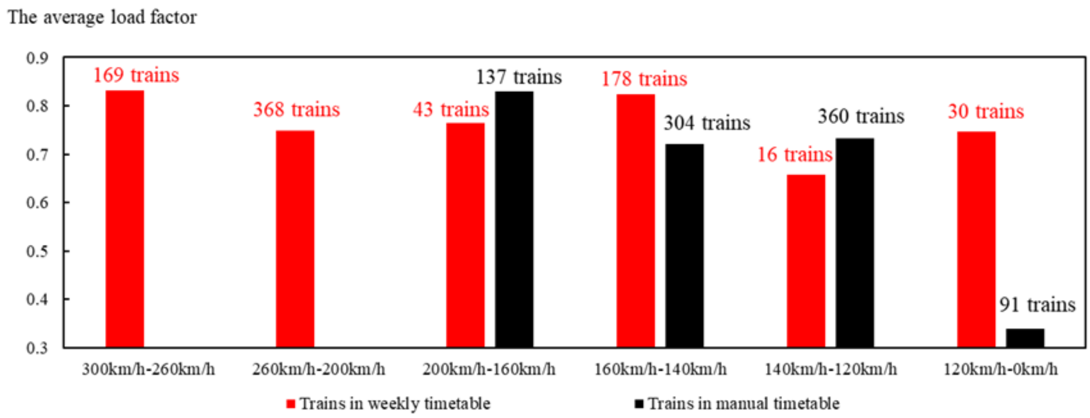

Figure 14. The average load factors of trains with different speed in the manual and weekly timetables.

(3) **Enhanced alignment between passenger kilometres and seat kilometres.** The frequency of train operations more accurately reflects passenger demand due to the optimisation of train stops in the weekly line plan. Figure 15 illustrates the ratio of total passenger kilometres to total seat kilometres—referred to as the "overall load factor"—for each period and day, highlighting the alignment between passenger demand and train capacity in both the weekly and manually generated timetables.

The weekly timetable exhibits smaller fluctuations in the overall load factor, indicating a closer alignment between seat kilometres and passenger kilometres. For example, in the manual timetable, 120 trains operate between 6:00 a.m. and 8:00 a.m., resulting in significantly lower overall load factors compared to the 78 trains operating in the same period in the weekly timetable. A similar pattern is observed between 4:00 p.m. and 6:00 p.m., where the manual timetable runs 136 trains compared to 110 trains in the weekly timetable.

Although the overall load factor in the weekly timetable increases during peak periods, it remains lower than that in the manual timetable, demonstrating its capacity to provide sufficient services during high-demand periods. During low-demand periods, the overall load factor in the weekly timetable decreases but often surpasses that of the manual timetable, underscoring its effectiveness in reducing resource wastage.

Additionally, as the demand input in this study is derived from corrections and predictions based on actual demand, the overall load factor in certain periods may approach or exceed 1. In such instances, excess passengers are likely to be transferred to adjacent periods.

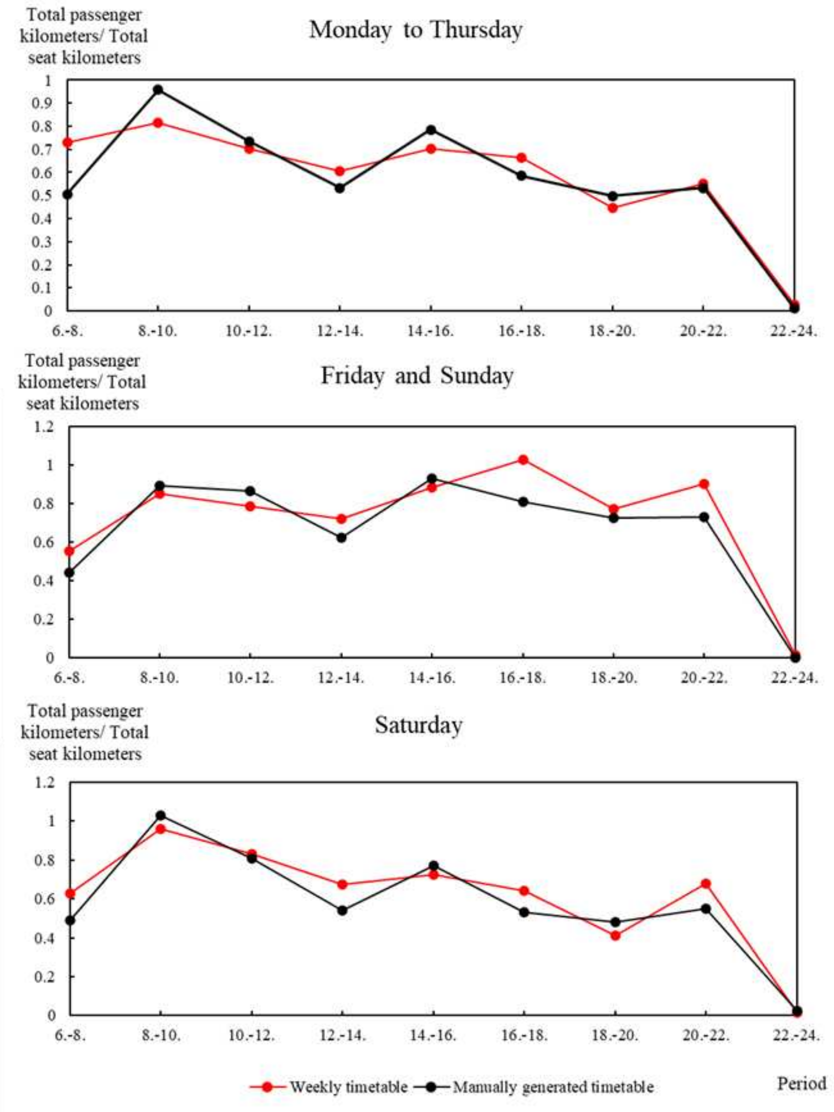

Figure 15. The changes in the ratio of total passenger kilometres to total seat kilometres across periods.

(4) **Trains exhibit a hierarchical and more regular departure sequence.** The hierarchical travel speeds of trains generally result in wider time intervals between departures, particularly between fast trains with few stops and slower trains with multiple stops. These intervals create triangular blank areas at the end of each period between the slower trains departing late in the current period and the faster trains departing early in the next period, as illustrated in Figure 16, which depicts the weekly timetable for the Nanjing-Shanghai HSR.

To mitigate imbalances in train services at stations within these blank triangular areas, the weekly timetable schedules trains with similar stop patterns to depart consecutively. For instance, as shown in Table 5, several fast trains depart consecutively between 8:00 a.m. and 8:24 a.m., forming a “faster train group” in the weekly timetable. The last two fast trains in this period depart at 9:42 a.m. and 9:44 a.m., creating another fast train group with those departing at the beginning of the next period. In contrast, this phenomenon is not observed in the manual timetable due to the smaller differences in train speeds (with the exception of two mandatory trains). The regular and balanced departure sequence of trains in the weekly timetable enhances the efficiency and predictability of train services.

Table 5. Long-distance train departure sequence between 8:00 a.m. and 10:00 a.m. in manual and weekly timetables.

Manually generated timetable		Weekly timetable		
Departure time	Travel speed (km/h)	Departure time	Travel speed (km/h)	Operating day
8:00	182 (Mandatory train)	8:00	267 (Mandatory train)	
9:09	110	8:08	230	
8:10	146	8:18	196	
8:40	136	8:21	274	
8:47	143	8:24	274	
9:00	182 (Mandatory train)	8:33	158	
9:01	140	8:35	162	
9:10	131	8:40	205	
9:17	123	8:46	155	Fri.
9:28	142	8:48	133	Mon. to Thu.
9:42	126	8:53	171	
9:55	135	9:00	230 (Mandatory train)	
9:56	130	9:03	160	Sat.
		9:03	166	Mon. to Thu.
		9:04	163	Sun.
		9:06	157	Fri.
		9:08	180	Sat.
		9:10	146	Fri.
		9:14	225	Mon. to Thu.
		9:17	157	
		9:42	239	Sat.
		9:44	230	Mon. to Thu.

5.3 Solution comparison: Nanjing-Shanghai and Beijing-Shanghai HSR lines

The Nanjing-Shanghai HSR, analysed in Sections 5.1 and 5.2, represents a typical shorter-distance commuter line between metropolitan areas. The weekly timetable for this HSR line is depicted in Figure 16. In contrast, the Beijing-Shanghai HSR, spanning 1,318 km and exemplifying a typical long-distance line connecting multiple metropolitan areas, is examined as another case study in this section. Due to the larger number of periodic trains on the Beijing-Shanghai HSR, a greater number of trains are included in stage 1 of the hierarchical strategy. Additionally, given the longer distance, more trains depart in earlier periods, leading to local timetables being solved from the earliest to the latest period in stage 2. The resulting weekly timetable is shown in Figures 17 and 18. The key differences between the Nanjing-Shanghai and Beijing-Shanghai HSR lines are as follows.

(1) **Alternating operation of trains with different speeds and stops.** Unlike the hierarchical departure sequence on the Nanjing-Shanghai HSR, the Beijing-Shanghai HSR passes through a greater number of major stations. As a result, even the fast trains have varying stopping patterns, making it impractical to reduce time intervals between schedules solely by having fast trains depart consecutively. Instead, on the Beijing-Shanghai HSR, trains with different speeds and stops operate alternately. To maintain travel speeds, more frequent overtakes occur between adjacent trains with differing stop patterns.

(2) **Less pronounced triangle blank areas between adjacent periods.** As analysed in Section 5.2.3, the hierarchical departure sequence on the Nanjing-Shanghai HSR typically creates triangular blank areas between two adjacent periods. Additionally, short-distance trains are operated only between intermediate stations and Shanghai Hongqiao (SHHQ) or Shanghai (SH),

leaving the triangular areas near Nanjing (NJ) and Nanjing South (NJN) unfilled. In contrast, the alternating operation of trains on the Beijing-Shanghai HSR results in smaller triangular blank areas. Short-distance trains operate along all segments of the route, with trains terminating and originating at various major stations along the HSR line. These short-distance trains fill the triangular blank areas, ensuring more balanced and consistent service across the entire route.

These differences underscore the necessity for distinct scheduling strategies tailored to the specific characteristics of each HSR line, demonstrating the flexibility and adaptability of our weekly timetabling approach in accommodating diverse operational scenarios.

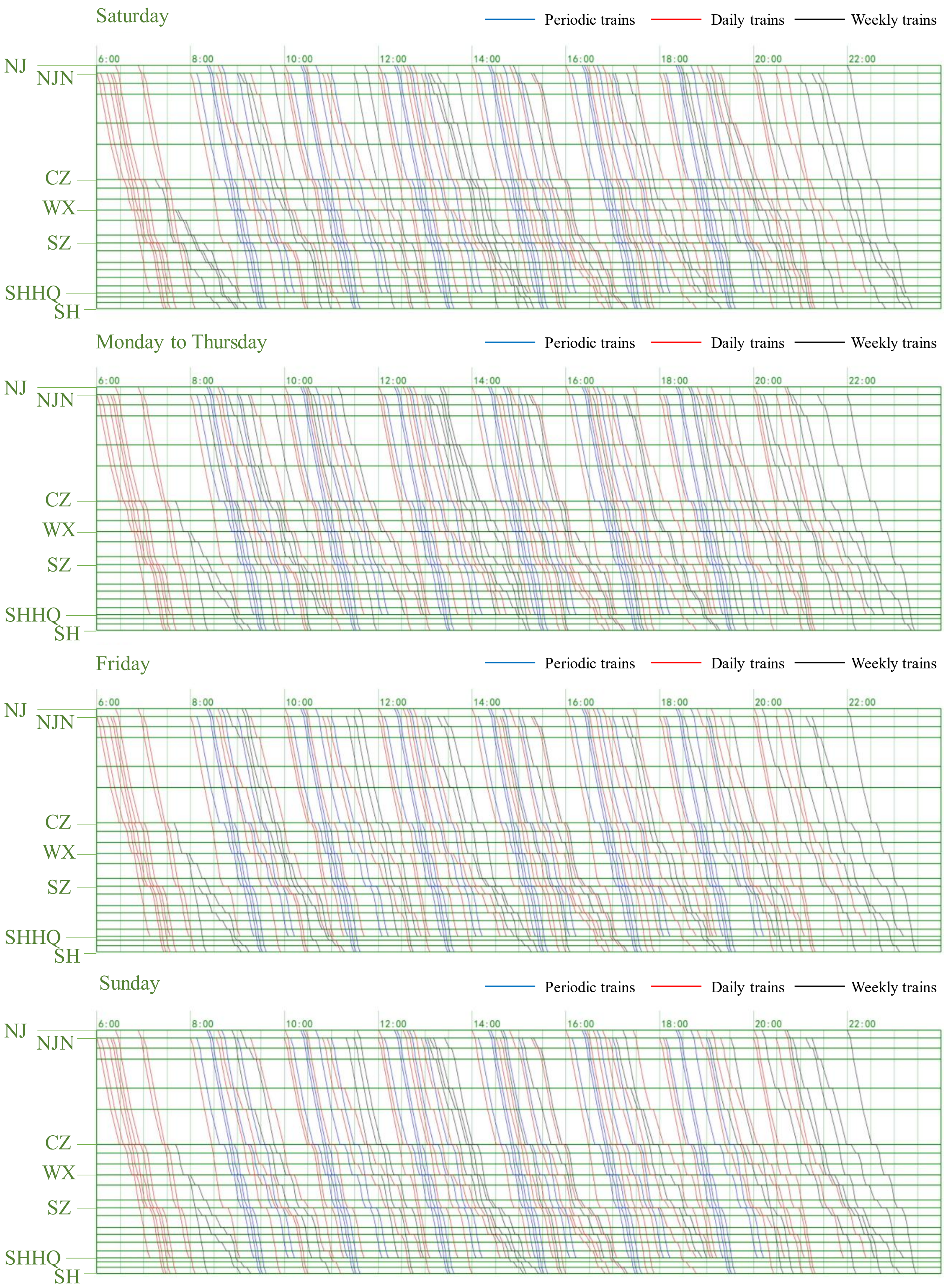

Figure 16. Weekly timetable for Nanjing-Shanghai high-speed railway line.

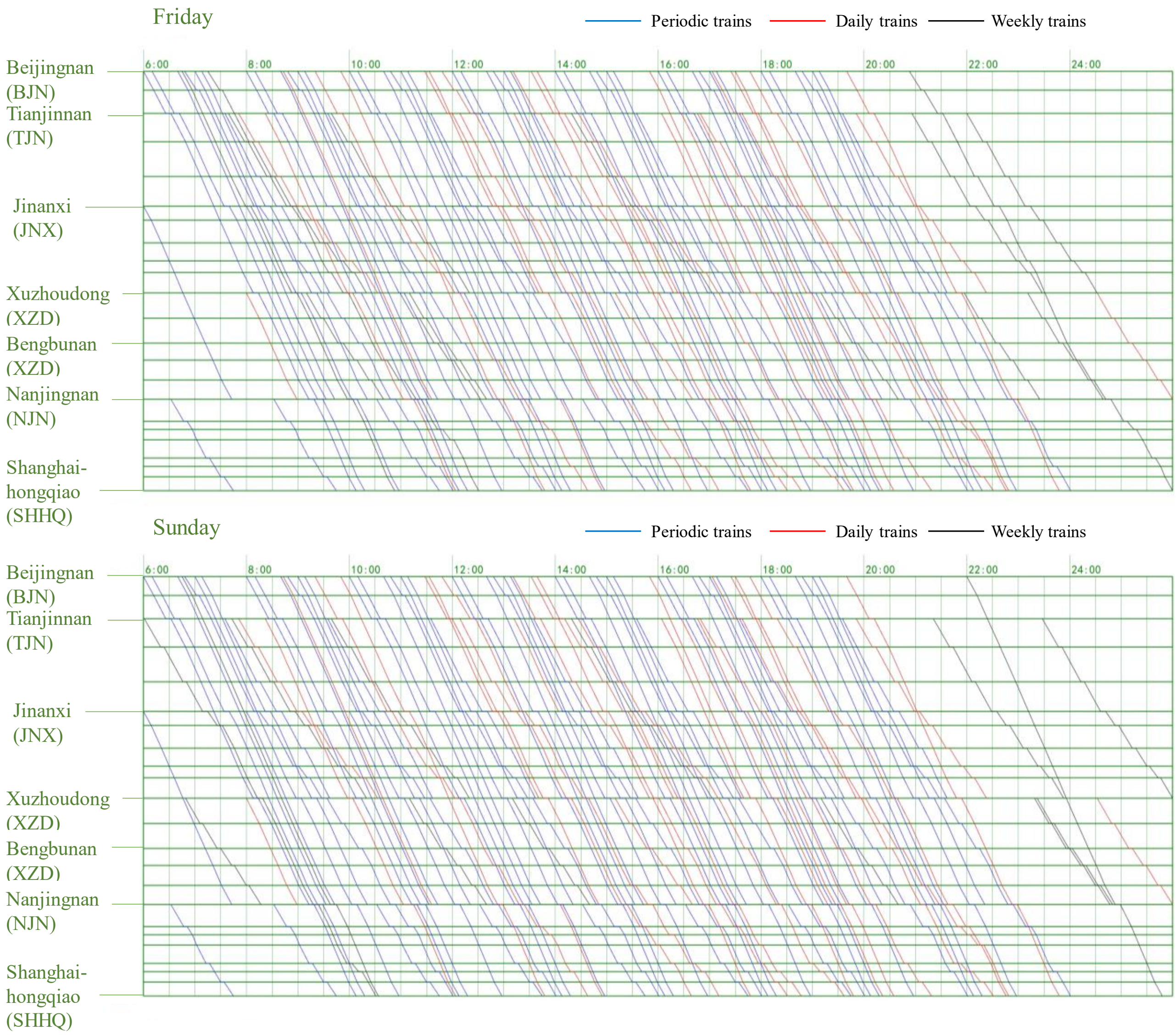

Figure 17. Weekly timetable for Beijing-Shanghai high-speed railway line (Friday and Sunday).

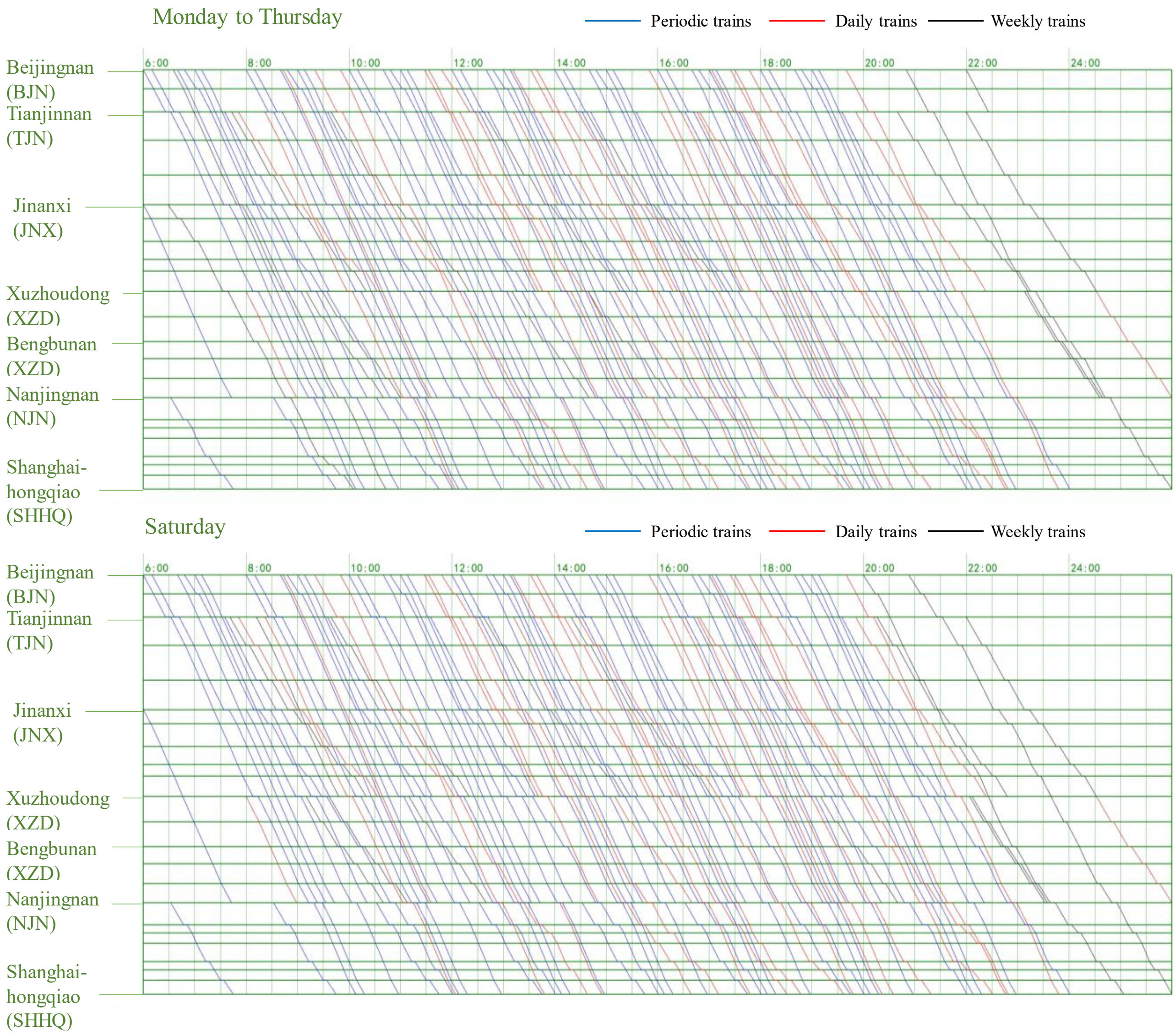

Figure 18. The weekly timetable for Beijing-Shanghai high-speed railway line (Monday to Thursday, and Saturday).

6. Conclusion

This study addresses the weekly train timetabling (WTT) problem for high-speed railways (HSR), aiming to determine train schedules for all seven days of the week while accommodating fluctuations in passenger demand across different periods of the day and varying days of the week. The study utilises detailed passenger demand data over a week and the weekly line plan obtained in previous research (Nie et al., 2022) as inputs. To effectively solve the WTT problem under practical real-world scenarios, an Estimation-Generation-Evaluation (EGE) method is proposed, incorporating three mathematical models.

To enhance the efficiency of the EGE method, a three-stage customised hierarchical train generation strategy is introduced during the "Generation" step. In stage 1, critical trains with higher passenger attraction are scheduled first. In stage 2, local timetables for each period are determined sequentially, while efforts are made to preserve the departure times of critical trains established in stage 1. In stage 3, trains cancelled in stage 2 are rescheduled in adjacent periods near their original departure times.

To evaluate the performance of the EGE method, the Nanjing-Shanghai HSR is selected as the primary case in the numerical experiments. First, the solution quality of the hierarchical strategy is compared against CPLEX in a series of small-scale scenarios. The results confirm the stability and efficiency of the proposed hierarchical strategy. Subsequently, using the entire weekly line plan as input, solutions derived from different implementations of stages 1 and 3 are compared to demonstrate the improvement in solution quality due to the hierarchical strategy. By comparing the EGE-generated weekly timetable with a manually generated timetable across various indicators, the optimisation effectiveness of the EGE method is verified. Finally, a case study of the Beijing-Shanghai HSR is conducted to demonstrate the flexibility and universality of the weekly train timetabling approach.

The potential advancements of this research can be explored from two perspectives. First, the weekly timetabling process could be enhanced by integrating vehicle scheduling, as varying timetables may result in more empty train movements to depots between days. Second, the staged solution approach inherent in the hierarchical strategy tends to produce locally optimal solutions rather than globally optimal ones. Therefore, further research is warranted to explore train stop adjustments or timetable modifications based on a comprehensive timetable evaluation. Additionally, an effective iterative methodology could be developed, incorporating an adjustment mechanism guided by the evaluation of the weekly timetable.

Reference

Borndörfer, Ralf, Heide Hoppmann, and Marika Karbstein. 2016. "Passenger routing for periodic timetable optimization." Public Transport 9 (1-2):115-35. doi: 10.1007/s12469-016-0132-0.

Brännlund, Ulf, Per Olov Lindberg, A. Nõu, and Jan-Eric Nilsson. 1998. "Railway timetabling using Lagrangian relaxation." Transportation Science 32(4):358-369. doi: 10.1287/trsc.32.4.358.

Cacchiani, Valentina, Alberto Caprara, and Paolo Toth. 2010. "Non-cyclic train timetabling and comparability graphs." Operations Research Letters 38 (3):179-84. doi: 10.1016/j.orl.2010.01.007.

Cacchiani, Valentina, Laura Galli, and Paolo Toth. 2015. "A tutorial on non-periodic train timetabling and platforming problems." EURO Journal on Transportation and Logistics 4 (3):285-320. doi: 10.1007/s13676-014-0046-4.

Cacchiani, Valentina, Jianguo Qi, and Lixing Yang. 2020. "Robust optimization models for integrated train stop planning and timetabling with passenger demand uncertainty." Transportation Research Part B: Methodological 136:1-29. doi: 10.1016/j.trb.2020.03.009.

Caprara, Alberto, Matteo Fischetti, and Paolo Toth. 2002. "Modeling and Solving the Train Timetabling Problem." Operations Research 50 (5):851-61. doi: 10.1287/opre.50.5.851.362.

Dong, Xinlei, Dewei Li, Yonghao Yin, Shishun Ding, and Zhichao Cao. 2020. "Integrated optimization of train stop planning and timetabling for commuter railways with an extended adaptive large neighborhood search metaheuristic approach." Transportation Research Part C: Emerging Technologies 117. doi: 10.1016/j.trc.2020.102681.

Şahin, Güvenç, Amin Ahmadi Digehsara, Ralf Borndörfer, and Thomas Schlechte. 2020. "Multi-period line planning with resource transfers." Transportation Research Part C: Emerging Technologies 119. doi: 10.1016/j.trc.2020.102726.

Hu, Lu, Dongjie Li, and Zhicheng Ren. 2025. "Collaborative train timetabling and passenger flow control in oversaturated metro lines considering state-dependence." Expert Systems with Applications 261. doi: 10.1016/j.eswa.2024.125436.

Fischer, Samuel M. 2020. "Locally optimal routes for route choice sets." Transportation Research Part B: Methodological 141:240-66. doi: 10.1016/j.trb.2020.09.007.

Kaspi, Mor, and Tal Raviv. 2013. "Service-Oriented Line Planning and Timetabling for Passenger Trains." Transportation Science 47 (3):295-311. doi: 10.1287/trsc.1120.0424.

Liao, Zhengwen, Haiying Li, Jianrui Miao, and Francesco Corman. 2021. "Railway capacity estimation considering vehicle circulation: Integrated timetable and vehicles scheduling on hybrid time-space networks." Transportation Research Part C: Emerging Technologies 124. doi: 10.1016/j.trc.2020.102961.

Liu, Jie, David Canca, Hongxia Lv, and Shaoquan Ni. 2024. "Demand-adapted train timetabling with coupling-decoupling operations on a bidirectional intercity railway line." Computers & Industrial Engineering 189. doi: 10.1016/j.cie.2024.109999.

Martin-Iradi, Bernardo, and Stefan Ropke. 2022. "A column-generation-based matheuristic for periodic and symmetric train timetabling with integrated passenger routing." European Journal of Operational Research 297 (2):511-31. doi: 10.1016/j.ejor.2021.04.041.

Nie, Bowen, Lei Nie, Huiling Fu, and Xin Wu. 2022. "Weekly line planning to improve the match between rail supply and passenger demand." Computer-Aided Civil and Infrastructure Engineering. doi: 10.1111/mice.12941.

Niu, Huimin, and Xuesong Zhou. 2013. "Optimizing urban rail timetable under time-dependent demand and oversaturated conditions." Transportation Research Part C: Emerging Technologies 36:212-30. doi: 10.1016/j.trc.2013.08.016.

Robenek, Tomáš, Yousef Maknoon, Shadi Sharif Azadeh, Jianghang Chen, and Michel Bierlaire. 2016. "Passenger centric train timetabling problem." Transportation Research Part B: Methodological 89:107-26. doi: 10.1016/j.trb.2016.04.003.

Schmidt, Marie, and Anita Schöbel. 2014. "Timetabling with passenger routing." OR Spectrum 37 (1):75-97. doi: 10.1007/s00291-014-0360-0.

Schöbel, Anita. 2017. "An eigenmodel for iterative line planning, timetabling and vehicle scheduling in public transportation." Transportation Research Part C: Emerging Technologies 74:348-65. doi: 10.1016/j.trc.2016.11.018.

Shi, Jungang, Jing Yang, Lixing Yang, Lefeng Tao, Shengjie Qiang, Zhen Di, and Junhua Guo. 2023. "Safety-oriented train timetabling and stop planning with time-varying and elastic demand on overcrowded commuter metro lines." Transportation Research Part E: Logistics and Transportation Review 175. doi: 10.1016/j.tre.2023.103136.

Tian, Xiaopeng, Huimin Niu, Yuxing Jiang, and Hetian Chai. 2024. "A variable-splitting Lagrangian decomposition for train timetabling and skip-stopping with train-type decision." Transportation Research Part C: Emerging Technologies 163. doi: 10.1016/j.trc.2024.104645.

Wang, Yizhen, Dewei Li, and Zhichao Cao. 2020. "Integrated timetable synchronization optimization with capacity constraint under time-dependent demand for a rail transit network." Computers & Industrial Engineering 142. doi: 10.1016/j.cie.2020.106374.

Wu, Runfa, Feng Shi, Shuo Zhao, Guangming Xu, and Hai Yang. 2022. "A Dantzig-Wolfe decomposition-based algorithm for capacitated passenger assignment problem with time-varying demand in high-speed railway networks." Transportation Research Part C: Emerging Technologies 145. doi: 10.1016/j.trc.2022.103909.

Xu, Xiaoming, Chung-Lun Li, and Zhou Xu. 2021. "Train timetabling with stop-skipping, passenger flow, and platform choice considerations." Transportation Research Part B: Methodological 150:52-74. doi: 10.1016/j.trb.2021.06.001.

Yan, Fei, and Rob M. P. Goverde. 2019. "Combined line planning and train timetabling for strongly heterogeneous railway lines with direct connections." Transportation Research Part B: Methodological 127:20-46. doi: 10.1016/j.trb.2019.06.010.

Yang, Lixing, Jianguo Qi, Shukai Li, and Yuan Gao. 2016. "Collaborative optimization for train scheduling and train stop planning on high-speed railways." Omega 64:57-76. doi: 10.1016/j.omega.2015.11.003.

Yang, Lin, Dewei Li, Hui Wang, and Yuan Gao. 2024. "Integrating stop planning, timetabling and rolling stock planning on high-speed railway lines: A multi-objective optimization approach." Expert Systems with Applications 237. doi: 10.1016/j.eswa.2023.121515.

Yao, Zhiyuan, Lei Nie, Yixiang Yue, Zhenhuan He, Yu Ke, Yuxin Mo, and Hongda Wang. 2023. "Network periodic train timetabling with integrated stop planning and passenger routing: A periodic time–space network construct and ADMM algorithm." Transportation Research Part C: Emerging Technologies 153. doi: 10.1016/j.trc.2023.104201.

Yuan, Jiawei, Yuan Gao, Shukai Li, Pei Liu, and Lixing Yang. 2022. "Integrated optimization of train timetable, rolling stock assignment and short-turning strategy for a metro line." European Journal of Operational Research 301 (3):855-74. doi: 10.1016/j.ejor.2021.11.019.

Yuan, Yin, Shukai Li, Ronghui Liu, Lixing Yang, and Ziyou Gao. 2023. "Decomposition and approximate dynamic programming approach to optimization of train timetable and skip-stop plan for metro networks." Transportation Research Part C: Emerging Technologies 157. doi: 10.1016/j.trc.2023.104393.

Yuan, Jiawei, Yuan Gao, Jiannan Ke, and Lin Yang. 2025. "Integrated train timetable with supply–demand interactions at the Chinese high-speed railway." Omega 131. doi: 10.1016/j.omega.2024.103185.

Yue, Yixiang, Shifeng Wang, Leishan Zhou, Lu Tong, and M. Rapik Saat. 2016. "Optimizing train stopping patterns and schedules for high-speed passenger rail corridors." Transportation Research Part C: Emerging Technologies 63:126-46. doi: 10.1016/j.trc.2015.12.007.

Zhang, Chuntian, Jianguo Qi, Yuan Gao, Lixing Yang, Ziyou Gao, and Fanting Meng. 2021. "Integrated optimization of line planning and train timetabling in railway corridors with passengers' expected departure time interval." Computers & Industrial Engineering 162. doi: 10.1016/j.cie.2021.107680.

Zhang, Yongxiang, Qiyuan Peng, Gongyuan Lu, Qingwei Zhong, Xu Yan, and Xuesong Zhou. 2022. "Integrated line planning and train timetabling through price-based cross-resolution feedback mechanism." Transportation Research Part B: Methodological 155:240-77. doi: 10.1016/j.trb.2021.11.009.

Zhang, Xin, and Lei Nie. 2016. "Integrating capacity analysis with high-speed railway timetabling: A minimum cycle time calculation model with flexible overtaking constraints and intelligent enumeration." Transportation Research Part C: Emerging Technologies 68:509-31. doi: 10.1016/j.trc.2016.05.005.

Zhao, Shuo, Hai Yang, and Yinghui Wu. 2021. "An integrated approach of train scheduling and rolling stock circulation with skip-stopping pattern for urban rail transit lines." Transportation Research Part C: Emerging Technologies 128. doi: 10.1016/j.trc.2021.103170.

Zhou, Housheng, Jianguo Qi, Lixing Yang, Jungang Shi, Hanchuan Pan, and Yuan Gao. 2022. "Joint optimization of train timetabling and rolling stock circulation planning: A novel flexible train composition mode." Transportation Research Part B: Methodological 162:352-85. doi: 10.1016/j.trb.2022.06.007.

Appendix 1. The linearisation of safety headway constraints in Section 3.2

The original formulation (g6) and (g7) are as follows

$$|y_{ls} - y_{l's}| + M(r_l + r_{l'} + \theta_{ll'}) \geq DH_s \tag{g6}$$

$$|x_{ls} - x_{l's}| + M(r_l + r_{l'} + \theta_{ll'}) \geq AH_s \tag{g7}$$

Introduce two auxiliary binary variables $\xi_{ll's}$ and $\zeta_{ll's}$. $\xi_{ll's}$ equals 1 if $y_{ls} - y_{l's} > 0$ and equals 0 if $y_{ls} - y_{l's} < 0$. $\zeta_{ll's}$ equals 1 if $x_{ls} - x_{l's} > 0$, equals 0 if $x_{ls} - x_{l's} < 0$. In this case (g6) is linearised into (a1) and (a2); and (g7) is linearised into (a3) and (a4).

$$y_{ls} - y_{l's} - M \cdot \xi_{ll's} - M(r_l + r_{l'} + \theta_{ll'}) \leq -DH_s \qquad \forall l, l';\ s \in (S_l - \rho_l) \cap (S_{l'} - \rho_{l'}) \quad \text{(a1)}$$

$$y_{ls} - y_{l's} - M \cdot \xi_{ll's} + M(r_l + r_{l'} + \theta_{ll'}) \geq DH_s - M \qquad \forall l, l';\ s \in (S_l - \rho_l) \cap (S_{l'} - \rho_{l'}) \quad \text{(a2)}$$

$$x_{ls} - x_{l's} - M \cdot \zeta_{ll's} - M(r_l + r_{l'} + \theta_{ll'}) \leq -AH_s \qquad \forall l, l';\ s \in (S_l - \sigma_l) \cap (S_{l'} - \sigma_{l'}) \quad \text{(a3)}$$

$$x_{ls} - x_{l's} - M \cdot \zeta_{ll's} + M(r_l + r_{l'} + \theta_{ll'}) \geq AH_s - M \qquad \forall l, l';\ s \in (S_l - \sigma_l) \cap (S_{l'} - \sigma_{l'}) \quad \text{(a4)}$$

The relationship between $\xi_{ll's}$ and $\zeta_{ll's}$ and x_{ls} and y_{ls} is formulated as (a5) and (a6).

$$y_{ls} - y_{l's} - M \cdot \xi_{ll's} \in (-M, 0) \qquad \forall l, l';\ s \in (S_l - \rho_l) \cap (S_{l'} - \rho_{l'}) \quad \text{(a5)}$$

$$x_{ls} - x_{l's} - M \cdot \zeta_{ll's} \in (-M, 0) \qquad \forall l, l';\ s \in (S_l - \sigma_l) \cap (S_{l'} - \sigma_{l'}) \quad \text{(a6)}$$

Appendix 2. Discussing the rationale of M-method in equations (g6), (g7), (g11), (g12), (e6) and (e7).

The original formulation (g6) and (g7) are as follows.

$$|y_{ls} - y_{l's}| + M(r_l + r_{l'} + \theta_{ll'}) \geq DH_s \quad \text{(g6)}$$

$$|x_{ls} - x_{l's}| + M(r_l + r_{l'} + \theta_{ll'}) \geq AH_s \quad \text{(g7)}$$

These constraints ensure that if trains l and l' depart on the same day ($r_l = 0$, $r_{l'} = 0$, and $\theta_{ll'} = 0$), the term $M(r_l + r_{l'} + \theta_{ll'})$ equals 0. In this case (g6) and (g7) are classical formulations to ensure the departure and arrival times of the two trains, x_{ls}, $x_{l's}$, y_{ls} and $y_{l's}$, maintain the required safety headways. However, if one of the trains is cancelled ($r_l = 1$ or $r_{l'} = 1$) or if both are weekly trains departing on different days ($\theta_{ll'} = 1$), the sum $r_l + r_{l'} + \theta_{ll'}$ will exceed 0, as all three variables are binary. In such cases, these constraints will no longer apply, ensuring that the left-hand sides of inequalities (g6) and (g7) are always greater than their right-hand sides. Consequently, the decision variables x_{ls}, $x_{l's}$, y_{ls} and $y_{l's}$ are no longer restricted.

The formulation (g11) and (g12) are as follows.

$$x_{l\sigma_l} + M \cdot (1 - z_{lq}) \geq \tau_q \quad \forall q \in Q, l \in L^Q \tag{g11}$$

$$x_{l\sigma_l} - M \cdot (1 - z_{lq}) \leq \tau_q \quad \forall q \in Q, l \in L^Q \tag{g12}$$

These constraints ensure that if train l is selected as a mandatory train and departs in slot q ($z_{lq} = 1$), the term $M \cdot (1 - z_{lq})$ equals 0. In this case its departure time at the first station, $x_{l\sigma_l}$, is restricted to τ_q, the designated time for slot q. However, if train l is not selected ($z_{lq} = 0$), then $1 - z_{lq} = 1$. In this case the left-hand side of constraint (g11) will always exceed the right-hand side, and the left-hand side of constraint (g12) will always be less than the right-hand side. Consequently, the decision variable $x_{l\sigma_l}$ is no longer restricted.

The formulation (e6) and (e7) are as follows.

$$x_{l\sigma_l} - M \cdot (1 - e_l) \leq \bar{x}_{l\sigma_l} + \tau^D \quad \forall l \in L^{Stage-1} \tag{e6}$$

$$x_{l\sigma_l} + M \cdot (1 - e_l) \geq \bar{x}_{l\sigma_l} - \tau^D \quad \forall l \in L^{Stage-1} \tag{e7}$$

These constraints are formulated in almost the same manner as (g11) and (g12). They use the variable e_l to measure whether $x_{l\sigma_l}$ lies within the range $[\bar{x}_{l\sigma_l} - \tau^D, \bar{x}_{l\sigma_l} + \tau^D]$. When $x_{l\sigma_l}$ exceeds $\bar{x}_{l\sigma_l} + \tau^D$, equation (e6) requires $e_l = 0$ to ensure that the term $M \cdot (1 - e_l)$ is positive. Conversely, when $x_{l\sigma_l}$ is smaller than $\bar{x}_{l\sigma_l} - \tau^D$, equation (e7) also sets $e_l = 0$. When $x_{l\sigma_l}$ lies in the range $[\bar{x}_{l\sigma_l} - \tau^D, \bar{x}_{l\sigma_l} + \tau^D]$, the term $M \cdot (1 - e_l)$ equals 0, implying that $e_l = 1$.