Communications in Transportation Research 4 (2024) 100131

Contents lists available at ScienceDirect

Communications in Transportation Research

journal homepage: www.journals.elsevier.com/communications-in-transportation-research

Full Length Article

The effect of optimisation objectives on the outcome of line planning

Prasetyaning Diah Rizky Lestari [a,b,*], Ronghui Liu [a], Richard Batley [a]

[a] Institute for Transport Studies, University of Leeds, Leeds, LS2 9JT, UK
[b] National Research and Innovation Agency of Indonesia (BRIN), Jakarta, 10340, Indonesia

ARTICLE INFO

Keywords:
Line planning
Optimisation
Railway transportation
Stopping patterns
Various objectives

ABSTRACT

One of the core problems in the strategic planning of railway operations revolves around developing an optimal line plan. The line plan optimisation problem aims to build a workable line system that achieves specific objectives. Many models presented in existing literature typically focus on either maximising direct traveller numbers or minimising costs. In contrast, this paper introduces a model with diverse objectives for addressing line plan optimisation problems, allowing for variations in stopping patterns across different lines. Our model examines how setting different objectives can result in different line plan designs. This will be valuable for railway operators, offering diverse perspectives when selecting the most suitable design, particularly in the context of new railway service development, such as the introduction of a high-speed train. A case study of future semi high-speed rail in Indonesia is presented to test the model.

1. Introduction

The planning of the railway system consists of several decisions on different and successive planning stages, ranging from strategic network design and line planning, tactical level timetabling, crew scheduling, etc., to real-time operational management (Assad, 1980; Bussieck et al., 1997a; Goossens et al., 2004). Importantly, decisions made at one planning stage can significantly impact those made at subsequent stages. To illustrate these stages, Fig. 1, adapted from studies such as Ghoseiri et al. (2004), Lusby et al. (2011), Schöbel (2012), and Zhao et al. (2021), provides an overview. Fig. 1 outlines the different levels of railway planning commonly adopted by most railway companies, highlighting the main subtasks and their typical order of occurrence. At the strategic level, long-term planning involves network design such as resource acquisition, and infrastructure planning, followed by decisions regarding which routes or lines to be operated and how frequently—commonly known as line planning. The tactical level focuses on resource allocation within a presumed fixed infrastructure involving train timetabling, rolling stock scheduling, and crew scheduling. Finally, at the operational level, daily tasks are determined, often requiring adjustments to operational policies established at the tactical level due to unexpected disruptions—commonly known as real time operational management. Some essential tasks at the tactical level, such as train timetabling depend on the output from the strategic level, mainly line planning. Consequently, the efficiency of operations in the next stage is directly determined by the quality of timetabling from the tactical level (Wang and Li, 2022). The problem addressed in this paper, which is line planning, belongs to the strategic level.

Examining the line planning process in detail within this planning regime contributes by establishing service routes or lines in a railway network with corresponding stop and frequency patterns to meet passengers' demands. The inefficiency of optimising line planning for the entire network necessitates the reallocation and circulation of vehicles (Zhang et al., 2021), which can be exhausting. Thus, it is important to design optimal lines that have considered efficiency from the start, with careful stop planning. Two related line planning methods are commonly employed: (1) optimising train frequencies with predetermined stopping patterns and (2) simultaneously optimising train frequencies and stopping patterns (Park et al., 2013). Typically, although the all-stop operation is the simplest stop planning method for satisfying all passenger demands, it might potentially increase the total travel time for long-distance passengers (Yang et al., 2016). The choice of train-stopping patterns plays a crucial role in linking the demand and supply sides and providing high-quality transport services. Therefore, it should be established early in the operational planning of railways.

Moreover, line planning is often undertaken with the objectives of maximising passenger satisfaction and/or reducing operational costs (Schöbel, 2012). These objectives are directly related to the interests of relevant stakeholders. When considering the affected parties, objectives can be distinguished into two types of orientation: first, from the

* Corresponding author. Institute for Transport Studies, University of Leeds, Leeds, LS2 9JT, UK.
E-mail address: tspdrl@leeds.ac.uk (P.D.R. Lestari).

https://doi.org/10.1016/j.commtr.2024.100131
Received 23 December 2023; Received in revised form 4 April 2024; Accepted 4 April 2024

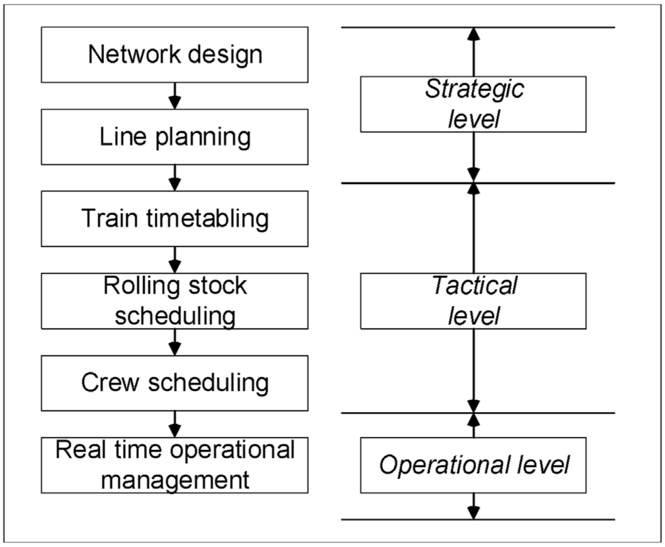


Fig. 1. Railway planning stages (adapted from Ghoseiri et al. (2004), Lusby et al. (2011), Schöbel (2012), and Zhao et al. (2021)).

passenger's point of view (passenger-oriented), and second, from the operator's point of view (operator-oriented). Delving further into operator types, such as privatised railway systems and nationalised railway systems, allows us to explore the application of these objective orientations in each system. In many privatised railway systems, the infrastructure may be owned and managed by a separate entity, often a government-owned one, while train operations may be handled by various private train operating companies (TOCs). In countries with nationalised railway systems, a single state-owned entity may oversee both the infrastructure and train operations. From the perspective of commercial TOCs, for instance in Japan's privatised rail industry, line planning primarily aims to target revenue growth (profitability) by optimising routes and services to attract more passengers. Commercial TOCs are free to decide on new routes and service improvements based on their own profit motivations (Song and Shoji, 2016). However, they need to provide better customer service to enhance profitability by maintaining and increasing ridership. When a state-owned company is responsible for both the infrastructure and train operations, the focus shifts to the idea of social welfare in the provision of railway services (Bracaglia et al., 2020). Social welfare consists of passenger benefits or consumers' surplus plus any profits generated by the operating agency. Some countries, such as China, France, and Indonesia, have state-owned railway companies overseeing both the infrastructure and train operations. The various foci of the objective of line planning will be considered in our study.

The problem considered in this paper consists of determining line plans with frequencies and stopping patterns to serve passenger demand and optimise a specified objective. Various objective functions based on passenger and/or operator interest orientation are proposed to provide railway operators, whether private or state-owned, with different perspectives when deciding which line designs to apply.

1.1. Literature review

In this review, we will briefly examine the existing literature dealing with railway line planning problems and introduce some terminology used throughout the text. Moreover, we will focus on papers that deal with different objective functions and how previous researchers have developed the modelling framework.

- Definition

Line planning involves determining the frequency and stopping patterns of train services needed on a line to meet a specified travel demand while reducing associated costs (Bussieck, 1998). The frequencies refer to how often service should be offered. Line planning is important in railway planning as it aims to balance the supply and demand sides. The outcome of line planning are the train numbers, routes, and stops in one day or a given period according to the time-space distribution of passengers. The line planning problem is typically formulated as an optimisation problem that maximises or minimises a given objective (Wang et al., 2011). Approaches to line planning problems can be categorised according to whether the objectives are passenger-oriented, operator--oriented, or passenger- and operator-oriented; they can be categorised according to the modelling approach, such as optimising train frequencies subject to predetermined stop patterns, or simultaneously optimising train frequencies and stop patterns.

- Objective function

Of the three types of objectives, it might be observed that the first two, passenger- and operator-oriented, might sometimes exhibit conflicting tendencies (Chang et al., 2000; Schöbel, 2012). For instance, when efforts are made to decrease the operational costs linked to running trains, including maintenance, staffing, and fuel expenses, operators may seek to compensate for these savings by modifying passenger fares. Another illustration is that although lowering speed and saving fuel can be economically advantageous for railway operators, it affects passenger satisfaction, travel time, and overall convenience. Nevertheless, these objectives are not mutually exclusive; they can coexist simultaneously without interfering with each other by substituting one objective with a relevant term to make them equivalent. The need is to reach a balance between passenger benefits and transportation costs, which lies at the heart of transport planning (Eliasson, 2021). For example, Yue et al. (2016) proposed a model to provide an optimal train service plan by simultaneously enhancing passenger satisfaction, improving fare revenues and profits for railway operators, and maximising line capacity utilisation. Recognising the importance of considering the benefits for both passengers and railway operators, some academics have sought to combine these two perspectives, referring to it as passenger- and operator-oriented, aiming for a favourable trade-off between the two. Further variants of these broad objectives in line planning models can be formed with either a passenger- or operator-oriented focus, or passenger- and operator-oriented.

1) Passenger-oriented objectives. Passenger-oriented objectives refer to line planning models that emphasise passenger benefits, i.e., maximum direct passengers, minimum number of transfers, or lost times of travellers. Many early research papers focused on maximising the number of direct travellers. One of the first studies is Bussieck (1998). He developed a mixed-integer linear programming (MILP) model and cutting planes method to solve the problem of maximising the number of direct travellers while ensuring sufficient capacity to accommodate all passengers. Remarking that Bussieck's approach would, however, result in more transfers for travellers without direct connections, Scholl (2006) and Schöbel and Scholl (2006) aimed to minimise the travel time of all passengers, including penalties for transfers, and formulated an integer program model in detail that was solved using the branch-and-bound and Dantzig–Wolfe decomposition algorithms. Recently, the passenger-centric line plan also focuses on improving passenger transfer efficiency, which is closely related to efficient train stop planning (Huang and Shuai, 2019). They applied a loop optimisation and passenger-oriented approach to obtain a better railway train stop plan by utilising an improved Frank–Wolfe algorithm until the plan achieves reasonable passenger transfer efficiency.
2) Operator-oriented objectives. Operator-oriented objectives pertain to line planning models that aim to minimise the operational costs of running trains. In the context of the operator-oriented approach,

previous academics have formulated different models to deal with this problem. Claessens et al. (1998) developed integer programming (IP) models to optimise lines by minimising the costs of a line system, which includes the costs of the number of cars being used, car kilometres, and train kilometres. This was solved by an algorithm based on a branch-and-bound procedure. Goossens et al. (2004) introduced a binary linear programming model to solve line planning problems that minimise the cost of a set of operational lines, employing a branch-and-cut approach. The costs are categorised into fixed and variable costs, where variable costs are hourly costs per kilometre associated with operating the line, and fixed costs are incurred for the availability of individual trains and carriages, including depreciation costs, for example. Goossens et al. (2006) introduced a novel approach to address stopping patterns, presenting distinct IP formulations for line planning. They developed a general model inspired by Bussieck (1998) and Claessens et al. (1998) to simultaneously solve operator-optimising line planning problems with multiple train types. Canca et al. (2016) proposed a profit-oriented model for line planning, aiming to maximise network profit, which is defined as the difference between revenue and total cost, from the operator's perspective. The authors employed a heuristic approach, utilising a combination of branch-and-bound and relaxed non-linear problems to solve the model.

3) Passenger- and operator-oriented objectives. Passenger- and operator-oriented objectives refer to line planning models that consider the benefits to both passengers and railway operators. However, when passengers and railway operators make decisions concerning transportation, they usually weigh only the benefits against their own transportation costs, often overlooking the broader social costs involved (Eliasson, 2021). Thus, this objective aims to balance the requirements between the passenger and the operator. Borndörfer et al. (2007) proposed a new model for line planning in public transport that allows generating lines dynamically and routing passengers freely to achieve minimum operating costs and passenger travel time. Wang et al. (2011) proposed a two-layer optimisation model to achieve an optimal stop-schedule set with the service frequencies by minimising the total operating cost and unserved passenger volume at the top layer and maximising the served passenger volume and minimising the total travel time for all passengers at the bottom layer. To address this problem, a genetic algorithm was proposed. Fu et al. (2015) took a customer/railway company perspective, creating an integrated hierarchical line planning strategy that utilised heuristic algorithms to address the resulting bi-level programming model. The first level aimed to minimise the total travel time of passengers, while the second level aimed to maximise the total passengers' kilometres. Profit can be considered a passenger- and operator-oriented objective when the revenue function is linked to travel time, as demonstrated in Li et al. (2011) and Liu et al. (2020). Li et al. (2011) addressed the design problem of a rail transit line with the goal of profit maximisation. They introduced two profit maximisation models that account for the effects of different transit pricing structures, including flat- and distance-based fare regimes. Similarly, Liu et al. (2020) applied dynamic fares dependent on transfer and detour patterns in lines chosen by passengers. These factors lead to lower ticket prices, subsequently affecting the operator's revenue—similar to the impact of more stops. Another approach to the passenger- and operator-oriented objective is proposed by Yan and Goverde (2019), using a multi-frequency line planning model to simultaneously minimise total travel time, empty seat – hours, and the number of train lines. Passengers may prioritise the shortest travel time, while the train operator aims to minimise empty seat – hours and the number of different train lines (stop patterns). The use of empty seat – hours in this model is unique, providing a direct view of capacity loss. Zhao et al. (2021) presented a model considering both cost and passenger-oriented line plans, minimising operating costs and improving passenger service levels using a simulated annealing algorithm. Zhou et al. (2021) designed a mixed-integer nonlinear programming model to determine line configuration, line frequency and passenger assignment in an urban rail transit network by minimising operational cost and total travel time. Furthermore, an outer approximation method was proposed to linearise the objective. In a recent study, Pu and Zhan (2021) proposed a two-stage optimisation model for line planning, considering passenger demand uncertainty. The Lagrangian relaxation algorithm and strengthening techniques were employed to minimise passengers' overall travel expenses and the operating costs of the railway company.

- Modelling framework

The line planning problem can be formulated as an IP problem which is to choose a set of lines and frequencies for the lines in such a way that there is enough transportation capacity to cover the aggregated demand on each link or sub-track of the transportation network (Bussieck et al., 1997b; Claessens, 1994; Torres et al., 2008). First, we will discuss this approach, which is based on a demand-covering model to determine the optimised frequencies that meet the leg traffic load (rather than origin-destination demands). The leg traffic load is associated with the passenger flow on a link or sub-track between two points. The line planning problem, which is based on a demand covering model, requires a priori distribution of the passenger flow on the arcs or sections of the transportation network. These aggregated demands are then covered by lines of sufficient capacity. Bussieck et al. (1997b) derived a mixed integer linear programming formulation to find the optimal or shortest choice of traffic lines that offer the adequate capacity to serve the known amount of traffic on the system by maximising the number of direct travellers. Possible lines in a railway network can be modelled by simple paths using an undirected graph with respect to some edge evaluation, i.e., travel time or travel distance. The objective is to fix the traffic load through the links of the railway network while assuming that all shortest paths are uniquely determined. The optimal solution will be a set of lines with their frequencies satisfying the line-frequency requirement for every link. They solved the problem using LP-relaxation and cutting planes method. Claessens (1994) developed a mathematical programming model to determine a set of operational lines, as well as which trains to assign to a specific route, the service frequencies, and the trains' lengths, by minimising the total costs of this set of operational lines. They used a demand-covering model by considering that all travellers can be transported on every sub-track. The problem was solved using a heuristic approach. The heuristic works by repeatedly solving the real-valued relaxation of the program and setting bounds on the line frequency variables until they are close to integers. Torres et al. (2008) used the concept of a flow-based model for bus line planning where each passenger has been routed along some specific directed path in a pre-processing step, such that an aggregated transportation demand on each link of the network can be computed. The problem can be formulated as an IP problem to minimise the total operating costs and solved using a direct solving method with the IP-solver SCIP (Solving Constraint Integer Programs). This study gave a set of all possible lines in a line pool.

Another approach is to consider origin-destination (*OD*) demand and optimal stopping patterns (Chang et al., 2000; Fu et al., 2015; Liu et al., 2020; Park et al., 2013; Wang et al., 2011). Chang et al. (2000) developed a model to determine the optimal train service plan for Taiwan's high-speed rail under a given set of stop schedules for a given travel demand. Their model sought to balance the requirements between the operator and the passenger by minimising the operator's total operating costs and the passenger's total travel time loss. It was solved using fuzzy mathematical programming to generate the best compromise train service plan, including the optimal stop-schedule plan, service frequency, and fleet size. In the study by Park et al. (2013), two scenarios of stopping patterns are examined in the context of the high-speed train service in Republic of Korea. The distinction is based on whether the patterns are determined within the model during the optimisation process or predetermined. Park's model is one of the simplest models that can optimally allocate *OD* demands to different lines within the network. Unlike many other studies that solely focus on determining the line frequency, Park's model accomplishes both tasks—assigning *OD* demands and establishing frequencies—in a single step. This model pursued the objective of minimising total operating cost and total passenger travel time and was solved using LP-relaxation based on Dantzig–Wolfe decomposition and branch-and-bound algorithm. As mentioned in the previous section regarding the passenger- and operator-oriented objectives, a study by Wang et al. (2011) also determined both the service frequency and the passenger assignment. However, their model was developed in a two-layer optimisation where the initial layer was dedicated to identifying the optimal stopping schedule and service frequency, while the subsequent layer determined the count of passengers opting for various trains based on the outcomes obtained from the first layer. The model was illustrated for the context of the Beijing–Shanghai high-speed railway line. Similarly, Fu et al. (2015) proposed an integrated hierarchical approach based on a bi-level programming model to establish line plans by specifying the stations and trains based on two categories (higher and lower classes) and solved it by heuristic method. The line plan model generation is based on daily passenger demand among any two stations. Liu et al. (2020) introduced a mathheuristic iterative strategy for line planning and frequency determination with a focus on maximising profit, specifically applied to a high-speed railway (HSR) network. The effectiveness of the suggested model was evaluated through practical tests on networks of different scales and with diverse randomly generated passenger demand situations. However, they only considered major stations with all-stop patterns. They left passenger flow in small stations in the pre-calculation phase. These studies show that having various planning scenarios in term of stop schemes for high-speed train services is important to increase service quality.

1.2. Research gaps

From the literature, it is evident that in the realm of optimisation problems, the formulation of objectives plays a pivotal role in shaping the eventual outcome of the optimisation process. Essentially, the objectives serve as the guiding principles that direct the algorithmic or mathematical processes towards achieving specific goals. The careful selection of objectives not only defines the success criteria for the optimisation task but also influences the entire trajectory of the solution-seeking process.

In line planning, these objectives are instrumental in shaping the quality of services offered by the operator. Moreover, the objectives in the strategic level of railway planning can be adjusted over time (Borndörfer et al., 2017), as it is a dynamic process influenced by various factors such as changes in demand, technological advancements, economic conditions, evolving priorities, and policy changes. As circumstances change, railway operators may find it necessary to reassess their objectives and make adjustments to better align with current needs and goals.

For instance, when introducing a new transportation mode, such as a high-speed train, planners might initially focus on maximising efficiency and minimising costs. However, as the high-speed railway system matures, there could be a shift in focus towards enhancing passenger satisfaction, accommodating changing travel patterns, or adopting new technologies to improve overall service quality. Another example is a country transitioning its railway system from a nationalised model to a privatised one. In the nationalised stage, the primary objective might be to provide universal access to rail services, emphasising widespread coverage even in less profitable, with a focus on social welfare. However, if the decision is made to privatise the railway system due to the financial burden faced by the state, objectives may shift in response to the change in ownership structure and market dynamics. In this stage, new objectives could include profit maximisation, cost efficiency, passenger satisfaction, innovation, and technology adoption, or adaptability to market demand, such as adjusting services and routes based on changing passenger needs.

In conclusion, flexibility in adjusting objectives in the line planning model is important as it allows railway operators to respond to emerging challenges, take advantage of new opportunities, and adapt to the evolving infrastructures and technologies. Regular reviews and adjustments to objectives ensure that the railway system remains aligned with the broader goals of providing efficient, sustainable, and passenger-centric transportation services. Therefore, comparing different objectives and examining their implications for line plan design is undoubtedly more revealing. Through this comparative analysis, we can assess how each choice influences the results by evaluating the estimated performance. Thus, this contribution forms the essence of our study.

We have selected specific line planning objectives for our study, identified through the literature, where each of them is important to consider, as the focus in the strategic planning of railways can be adjusted over time. The objective function is relatively straightforward, making it easily adjustable by railway planners to capture different interests. Table 1 presents objective options selected based on their group focus (passenger-oriented, operator-oriented, and passenger- and operator-oriented), and we explain the purpose of considering each. Some of these single objectives may encompass multiple factors (such as travel time, operating costs, and revenue).

In term of modelling approach, Table 2 summarises key literature on leg traffic loads and *OD*-based line planning approaches. While earlier studies consider input demand as either leg flow or *OD* demand, our study aims to determine the optimal line frequency that meets both specified leg traffic loads and *OD* demands. Whilst a high-speed train service needs various stop schemes to deliver the requisite service quality, an all-stop scheme may degrade the service because intermediate stops will also increase the travel time. Our model considers two-stop schemes, with and without predetermined stopping patterns. In the first model, stopping patterns are given in advance in the form of some possible combination. In the second model, the optimal stopping patterns will be searched through optimisation, wherein a distinct stopping pattern can be denoted by a distinct line. In Table 2, consideration of stopping patterns in the model will be given in two ways; considered and not considered. 'Not considered' means that the model does not consider stopping patterns-related variables when modelling the objective and constraints. The line plan results will only follow whichever route satisfies travel demand, even if it is a very short trip (only between two stations). The results will vary by start- and end-station because there is no need to define where a train should depart and end for this model. 'Considered' means that the model includes stopping patterns-related variables when developing the optimisation model. For example, trains running on a given route r will travel from the same start-station to the same end-station, and then the optimisation model will seek optimal stops for these trains—called 'lines'. For this given route r, there might be more than one line. Trains running on the same line will depart and arrive at the same two endpoint stations and stop at the same intermediate station(s).

2. Problem definition

2.1. Definition

Before outlining the model formulation, we need to define some terminology used in the model and describe the set of parameters and decision variables.

- Network

A physical railway network is made up of a number of stations $\in N$ and segments $\in \mathbb{Z}$. A 'segment' is the physical railway track between two stations. We differentiate between a segment and a link. A 'link' is a rail connection between two stations. It can be unidirectional or bidirectional, depending on whether trains can travel in one direction or

Table 1
Choices of line planning objectives considered in our study.

Objective function	Type	Purpose	Ref.
Operation cost + passenger travel time	Passenger- and operator-oriented	Balancing the customer and economic perspective by finding an optimal solution that benefits both the operator and passenger.	Borndörfer et al. (2007); Park et al. (2013); Wang et al. (2011); Zhou et al. (2021)
Passenger travel time	Passenger-oriented	Directly addressing the customer-centric aspect of transportation services. Passengers typically value shorter travel time, making it an essential component in designing a railway system that meets customer expectations.	Scholl (2006); Schöbel and Scholl (2006)
Operation cost	Operator-oriented	Ensuring that the system operates in an operator-effective manner, which is essential for long-term financial viability.	Claessens et al. (1998); Goossens et al. (2004, 2006)
Profit	Passenger- and operator-oriented	Balancing the customer and economic perspective by finding an optimal solution that benefits both the operator and passenger. The profit will be formulated to be linked with passenger travel time to reflect social welfare.	Canca et al. (2016); Li et al. (2011); Liu et al. (2020)
Empty seat	Operator-oriented	Ensuring that seats are occupied optimally to make the best use of available resources (minimising capacity loss). Operating trains with fewer empty seats directly correlate with higher revenue potential.	Yan and Goverde (2019)
Empty seat–kilometres	Operator-oriented	A similar purpose to empty seats objective, but also contributing to reducing the distance travelled with empty seats, thereby enhancing overall operator-effectiveness and improving service in terms of passenger—kilometres. This aligns with operator-oriented considerations by minimising unnecessary operational costs associated with underutilised capacity.	Yan and Goverde (2019)

Table 2
Related previous studies on line planning.

Ref.	Input demand type	Objective function	Stopping patterns	Character of stopping patterns	Optimisation formulation
Claessens (1994)	Leg load	Operation cost	Not considered	—	Non-linear programming
Bussieck et al. (1997b)	Leg load	Direct travellers	Not considered	—	Mixed integer linear programming
Torres et al. (2008)	Leg load	Operation cost	Considered	With predetermined stopping pattern	Integer linear programming
Chang et al. (2000)	*OD* demand	Operation cost, travelling time	Considered	Without predetermined stopping pattern	Multi objective linear programming
Park et al. (2013)	*OD* demand	Operation cost + travelling time	Considered	With and without predetermined stopping pattern	Integer linear programming
Wang et al. (2011)	*OD* demand	(1) Operation cost + unserved passenger volume; (2) Served passenger volume, travelling time	Considered	Without predetermined stopping pattern	(1) Non-linear programming; (2) Mixed integer linear programming
Fu et al. (2015)	*OD* demand	(1) Total passengers' travel time; (2) Kilometres of route length chosen by passengers	Considered	With predetermined stopping pattern	Integer linear programming
Liu et al. (2020)	*OD* demand	Profit	Considered	With predetermined stopping pattern	Mixed integer linear programming
Pu and Zhan (2021)	*OD* demand	(1) Total travelling time; (2) Operation cost	Considered	Without predetermined stopping pattern	Mixed integer linear programming
This study	Leg load and *OD* demand	Various distinct objective functions	Considered	With and without predetermined stopping pattern	Integer linear and non-linear programming

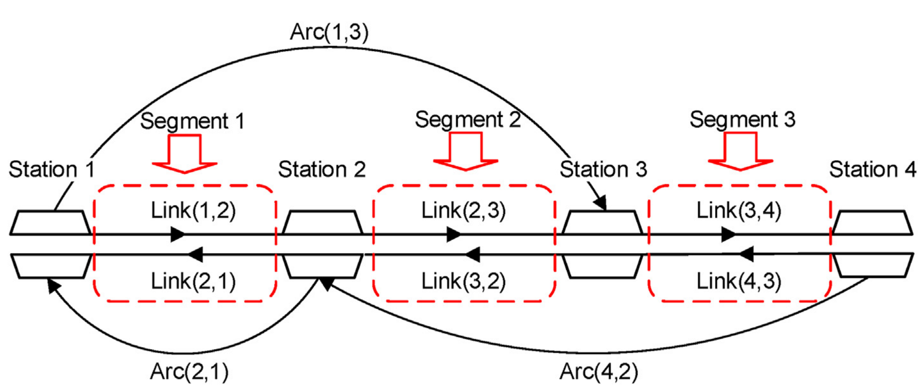


Fig. 2. Illustration of a small railway network with unidirectional double-track.

opposing directions on the associated track. We use the term link when we refer to other terms related to direction; otherwise, a segment is used. The railway track is the structure that provides trains with a surface for wheels to roll upon. The illustration to differentiate between a segment and a link can be seen in Fig. 2.

- Route

From the Commission Regulation (EU) No. 1305/2014, a route is a geographical way to be taken from a starting point to a point of destination (Official Journal of the European Union, 2014). Derived from that definition, in our research, we describe a route as a set with three elements or a triplet (s, t, A). In this context, s, t, and A denote the start-station, end-station, and set of links that trains go through, respectively. A link is represented by $(i,j) \in A$.

- Journey

From the Commission Regulation (EU) No. 454/2011 definition, a journey is a movement of a passenger (or several passengers travelling together) from location A to location B (Official Journal of the European Union, 2011). It denotes the spatial forwarding of a train from the origin station to the destination station. To perform a journey, a passenger can take one or more routes to complete their journey. People usually use the term 'journey' to refer to travelling a long distance, while 'trip' refers to a short distance. In this paper, both terms are identical.

- Line

A line refers to a path within the public transportation network (Schöbel, 2012). In railways, a line is one or more adjacent running tracks forming a route between two points. In our definition, we define a line as a route that initiates at the start-station (first point), terminates at the end-station (final point), and links various intermediary stations that a particular train can serve. In our problem, a line is a set of two (r, H), where r signifies the route and H denotes the set of stop stations. Along a given route, multiple lines with distinct stopping patterns may coexist. Whereas a route considers only the start-station, end-station, and geographical ways that a train travels between them, a line also considers intermediate stations as stop schemes. To illustrate, Fig. 3 shows how these definitions are distinct.

- Arc

An arc is used to represent nonstop travel on a route. An arc (i,j) with $i,j \in N$ exists if nonstop travel from i to j is possible. To avoid confusion between a link – that is defined earlier – and an arc, the following explanation is given. A link (i,j) is a physical rail that connects two consecutive stations i and j, while an arc (i,j) reflects a nonstop travel from stations i to station j, where i and j do not have to be sequential. Following Schlechte (2014), an arc $(i,j) \in A_k^r$ represents the consecutive processing of task j after i by a pairing. A_k^r comprises the set of arcs that symbolise nonstop travel within the k-th line on the r-th route.

- Origin and destination pairs and matrix

Traffic is broken down into *OD* pairs based on source and destination. The form of these components is usually represented as an *OD* matrix. Directional flow data between a number of locations are needed to build an *OD* matrix. For example, Fig. 4 represents movements from five locations (A, B, C, D, and E) to one or more locations among themselves. From this graph, a 5 × 5 *OD* matrix can be constructed where each *OD* pair becomes a cell of the matrix. Any pair of *ODs* without an observable flow is given a value of 0.

2.2. Set of parameters and variables

Following from the above terminology, we now define the set of indices, parameters, and decision variables used in the model. The input data required by the model are the following:

- Network data
  - N: Set of stations in the physical rail network, in which i,j represent station, $i,j \in N$.
  - $\mathbb{Z}$: Set of segments in the physical rail network, in which $(i,j), j = i+1$ represents segment.
  - Λ: Set of links in the rail network, in which (i,j) represents link.
  - R: Set of route indices $= \{1,2,\ldots\}, r \in R$ represents route.
  - $s(r), t(r)$: Start-station and end-station of r-th route, respectively.
  - L^r: Set of line indices on the r-th route $= \{1,2,\ldots,K^r\}$.
  - K^r: For a case without stopping patterns, K^r is the desired number of lines or stopping patterns on the r-th route.
  - $L^{r,OD}$: Set of line indices on the r-th route that accommodate the *OD* passenger flow $= \{1,2,\ldots\}$.
  - A_k^r: Set of arcs representing nonstop travel within the k-th line on the r-th route, in which (i,j) represents arcs, $(i,j) \in A_k^r, i \neq j$.
  - A^r: Set of arcs representing nonstop travel on the r-th route, in which (i,j) represents arcs, $(i,j) \in A^r, i \neq j$.
  - H_k^r: Set of stop stations for the k-th line on the r-th route, in which i,j represent station, $i,j \in N$.
  - R_{ij}: Set of routes which use the link (i,j), indexed by $\{1,2,\ldots\}$.
  - Ω^r: Length of the r-route (km).
  - d_{ij}: Distance between the nonstop arc (i,j) or from station i to the station j (km).

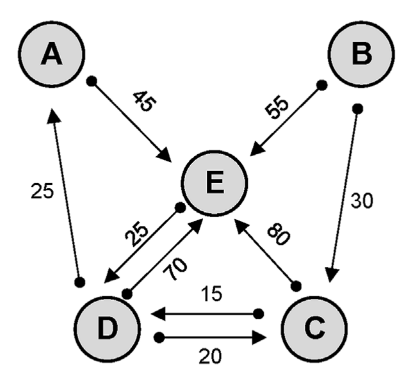


Fig. 4. Flow information between locations A, B, C, D, and E.

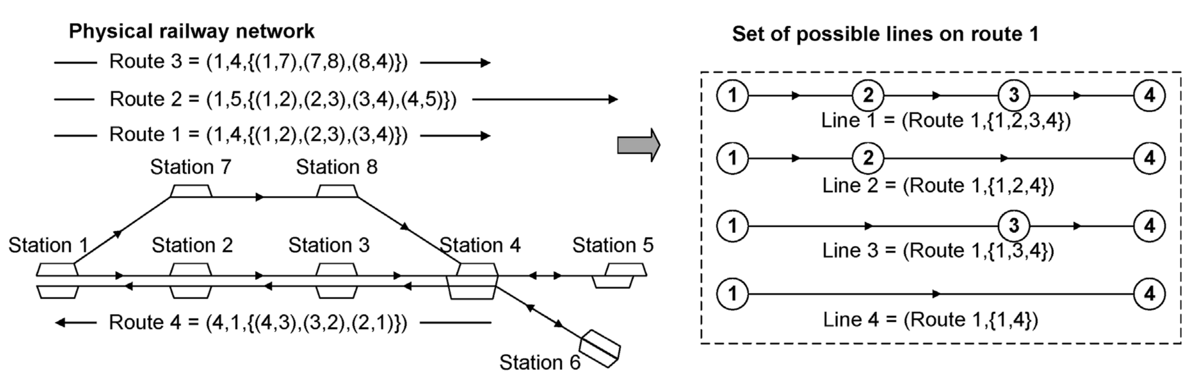


Fig. 3. Illustration of a railway network, route and line.

- Demand data
  D^{OD}: Passenger demand between station O and D on the r-th route.
  S: Set of OD pairs, indexed by $OD \in S$.
  S_{ij}^{r}: Set of possible OD pairs (passengers) who can travel in the nonstop arc (i,j) on the r-th route.
  F_{ij}: Number of passengers passing through each link, $(i,j) \in \Lambda$.
- Vehicle data
  C^{r}: Seating capacity of train on the r-th route.
  M_{ij}: Minimum number of trains which can run on the link (i,j) (derived from the total traffic leg load on the link (i,j)).
  T_{ij}: Maximum number of trains which can run on the link (i,j).
  N_k: Maximum number of trains running on the k-th line.
- Time related data
  $t_k^{r,OD}$: Travelling time between O and D for k-th line on the r-th route.
  t_{ij}^{r}: Dwell time at station i plus nonstop travelling time between i and j on the r-th route.
- Cost related data
  c_k^{r}: Cost for operating one train in the k-th line on the r-th route.
  $P_k^{r,OD}$: Ticket fares for each OD pair in the network within the k-th line on the r-th route.
  $P_{k,ij}^{r,OD}$: Ticket fares for each OD pair in the network uses the nonstop arc (i,j) within the k-th line on the r-th route.

To construct the model, variables are established according to two cases, i.e., with and without predetermined stopping patterns.

- With predetermined stopping patterns
  f_k^{r}: Frequency of trains on k-th line on the r-th route.
  $\lambda_k^{r,OD}$: OD passenger flow assignment on k-th line on the r-th route.
- Without predetermined stopping patterns
  f_k^{r}: Frequency of trains on k-th line on the r-th route.
  $x_k^{r} = 1$: if the k-th line on the r-th route is used, and $= 0$, otherwise.
  $x_{k,ij}^{r} = 1$: if the k-th line on the r-th route uses the nonstop arc (i,j), and $= 0$, otherwise.
  $\lambda_k^{r,OD}$: OD passenger flow assignment on k-th line on the r-th route.
  $\lambda_{k,ij}^{r,OD}$: OD passenger flow assignment on the nonstop arc (i,j) of k-th line on the r-th route.

2.3. Mathematical formulations

We now formally move on to the model formulation of the line planning problem using different objective scenarios. Our proposed model is motivated by Park et al. (2013), as their model is one of the simplest models that can optimally allocate OD demands to different lines within a specific supply network and determine the line frequency in a single-stage model. To enhance the quality of the line planning outcome, we introduce a new constraint aimed at satisfying leg traffic loads. Firstly, we provide an IP formulation for the problem with predetermined stopping patterns. Secondly, we will present an IP formulation for the problem where stopping patterns in each line are decided during the optimisation process. In our model, we will evaluate a series of single-objective models in turns to observe the implications for the line plan design. The differences highlighted in our two proposed line plan models are summarised in Table 3.

2.3.1. Case with predetermined stopping patterns

In the line planning problem, lines or stopping patterns can be either determined through modelling or predetermined based on the operational rules of the railway company. Our literature review emphasises the importance of considering various planning scenarios in high-speed train services, especially when introducing a new mode of transportation. This implies allowing trains on the same route to have different stopping patterns, ultimately reducing the total travel time.

Table 3
Highlights from the two proposed models.

Highlight	Predetermined stopping pattern	
	With	Without
Input		
A physical rail network (supply network)	✓	✓
A set of possible routes	✓	✓
A set of possible stopping patterns (lines)	✓	×
Desired number of possible stopping patterns (lines)	×	✓
OD demand information	✓	✓
Train capacity, minimum number of trains run in the network	✓	✓
Output		
Optimal number of trains (frequency) per line per day	✓	✓
Optimal stopping patterns (lines)	×	✓
Passenger flow assignment	✓	✓

This section presents the optimisation model for a case with predetermined stopping patterns. The inputs include a physical rail network, a set of possible routes, a set of possible stopping patterns (candidate of lines), OD demands, and other information related to train capacity design, minimum required number of trains running in the network, etc. The output aims to find the optimal frequency of trains on each line and the passenger flow assignment to achieve a specified objective. We introduce various objectives, as presented in Table 1, to be incorporated in our model. For simplicity, we focus on a single route (one direction) in the model ($r = 1$). The models are formulated as IP formulations as follows.

- Objective 1-P: Total operating cost and total passenger travel time

The first objective function, designated as objective 1-P, is passenger- and operator-oriented, incorporating two elements: the total operating cost and passenger journey time. The first element represents the operating cost incurred by the railway operator for running a specific number of trains (frequency) in the railway network. The second element indicates the overall passenger travel time, calculated as the sum of the products of travel time and passenger assignment at each origin and destination. The goal is to minimise the sum of the total operational costs and the overall passenger travel duration. Due to the difference in units between the first (monetary-unit) and the second element (time-unit) in the objective function, a multiplier is introduced to the second element to convert to monetary units. This multiplier is denoted as VoT, representing the value of time. VoT serves to convert travel time into a monetary cost. In transport economics, the value of time signifies the opportunity cost of the time that a traveller spends on their journey. In essence, it reflects the amount a traveller would be willing to pay to save time or the compensation they would accept for time lost.

$$\min \sum_{k\in L} c_k f_k + VoT \sum_{k\in L} \sum_{OD\in S} t_k^{OD} \lambda_k^{OD} \tag{1}$$

- Objective 2-P: Total passenger travel time

We will now delve into the model featuring the second objective function, denoted as objective 2-P, which is centred solely on passenger satisfaction. This objective is essentially the total journey time of all passengers, a component derived from the second element of the first objective 1-P. The journey or travel time of a passenger is defined as the duration between arriving at the origin and reaching the destination.

$$\min \sum_{k\in L} \sum_{OD\in S} t_k^{OD} \lambda_k^{OD} \tag{2}$$

- Objective 3-P: Total operating cost

In this objective function, we aim to minimise the operating cost incurred by the train operator – referred to as objective 3-P. Information about expenses for rolling stocks, drivers, etc., needs to be known to determine the costs associated with a line. However, in common practice, without detailing each cost item, these costs are often aggregated into some fixed cost per line, coupled with some variable cost contingent upon the frequency. We can exclude the fixed cost and focus solely on the variable cost. Parameter c_k^r represents the variable cost for the operational cost incurred in running a train in the network, and the value is associated with a specific line. Each line may have a different cost due to variations in stop patterns.

$$\min \sum_{k \in L} c_k f_k \tag{3}$$

- Objective 4-P: Total profit

We also incorporate an objective function that combines passenger- and operator-oriented goals, forming a profit-oriented objective – referred to as objective 4-P. The value of the profit objective function serves as a measure of the overall cost or benefit that an optimal solution would yield. Decision-makers can use profit as the objective function to find good solutions to the associated problem. Assuming passenger demand is given and fixed, the profit for the operator is computed by considering the overall operating cost and the income generated through the sale of train tickets. The operator costs are determined by Eq. (3). Following Liu et al. (2020), in order to consider the travel time in the ticket fare, the retail fare for the train tickets is assumed to decrease when passengers choose lines with more stop schemes for the same *OD* pair. For instance, the ticket fare for a direct train with no intermediate stops is higher than a train that stops at some intermediate stations. This is intended to compensate passengers who must wait at intermediate stations and encourage the railway operator to provide direct services as frequently as possible. Therefore, the first element of the profit objective function represents the total revenue from train operation (passenger tickets), while the second element accounts for the operating cost of running trains in the network. This objective function is formulated as a maximisation problem.

$$\max \sum_{k \in L} \sum_{OD \in S} P_k^{OD} \lambda_k^{OD} - \sum_{k \in L} c_k f_k \tag{4}$$

- Objective 5-P: Empty seats

Objective 5-P is developed from the railway planner's point of view. Although train stop planning is important in long-journey transportation (e.g., the high-speed train) to accommodate different passenger *OD* demands, some problems remain. For example, the service frequency of the intermediate station along the corridor is reduced due to an arbitrary train stop schedule plan. It could lead to low seat occupancy with commercial implications for operators (Wang et al., 2020). Consequently, effective management and operation of high-speed train services have become crucial for railway companies. In response to this issue, this objective aims to maximise the capacity utilisation of each train in the line design, ensuring as many seats as possible are occupied. The objective function can be formulated as follows: The first term represents the total seating capacity for all lines in the network, and the second term signifies the total number of passengers occupying the seats.

$$\min \sum_{k \in L} \sum_{ij \in A_k} \left(Cf_k - \sum_{OD \in S_{ij}} \lambda_k^{OD} \right) \tag{5}$$

- Objective 6-P: Empty seat–kilometres

We expand objective 5-P to minimise the empty seat–kilometres of trains, referred to as objective 6-P. To achieve this, we incorporate the total distance of a route, denoted as Ω, and the distance between every two stations, d_{ij}, into the model. Similar to objective 5-P, this objective aims to assess the effective management and operation of the high-speed train service through line plan designs. The empty seats in Eq. (5) represent a wasted opportunity at the departure that can be reduced during the optimisation. If we add a distance parameter to those empty seats, as shown in Eq. (6), it might reduce empty seats for longer distances. The railway company can then earn higher revenue and improve service in terms of passenger–kilometres and overall passenger load by maximising the utilisation of existing facilities. Consequently, this objective is anticipated to enhance seat allocation by optimising overall revenue and passenger load.

$$\min \sum_{k \in L} Cf_k \Omega - \sum_{k \in L} \sum_{ij \in A_k} \sum_{OD \in S_{ij}} \lambda_k^{OD} d_{ij} \tag{6}$$

- Constraints

The constraints are described as follows:

Capacity constraint:

$$\sum_{OD \in S_{ij}} \lambda_k^{OD} \le Cf_k, \forall (i,j) \in A_k, \forall k \in L \tag{7}$$

Equation (7) signifies the total of *OD* passenger flow assignment passing the (i,j) link within the k-th line must not exceed the overall seating capacity of the k-th line; the latter is calculated by multiplying the frequency by seating capacity C.

Passenger flow constraint based on *OD* passenger demand:

$$\sum_{k \in L^{OD}} \lambda_k^{OD} \ge D^{OD}, \forall OD \in S \tag{8}$$

Equation (8) ensures that the sum of *OD* passenger flow assignments across all lines is no less than the demand for that *OD* pair, D^{OD}. For objectives 4-P, 5-P, or 6-P, we change Eq. (8) into an equality form. Since the demand is fixed, the total passenger flow assignment for each *OD* pair across all lines should fulfil only the *OD* passenger demand. If this constraint remains an inequality while using objective 4-P, for example, the first term value of objective 4-P in Eq. (4), which consists of the passenger flow assignment for each *OD* pair, would be excessive due to the nature of the objective. A similar condition happens when using objectives 5-P and 6-P. The second term in these objectives – Eqs. (5) and (6) – involving the passenger flow assignment, might result in excessive values due to the nature of both objectives. In this context, achieving a better solution entail having the second term with a minus sign represent a significantly larger number, as a larger value is considered superior to a smaller one. Changing the 'more than' sign to an 'equal' sign in Eq. (8) limits the passenger flow to satisfying only the *OD* demand flow.

Passenger flow constraint based on leg traffic loads:

$$\sum_{k \in L} \sum_{OD \in S_{ij}} \lambda_k^{OD} \ge F_{ij}, \forall (i,j) \in \Lambda \tag{9}$$

Following Claessens (1994), Eq. (9) ensures that all passengers can be transported, similar to Eq. (8). However, in Claessens' model, the constraint was developed using frequency decision variables, where the frequency multiplied by train capacity should be greater than the number of passengers passing through each link. In our model, we formulate it using passenger flow assignment decision variables. This is expected to directly reduce the occurrence of empty seats for each link by utilising the actual passenger flow assignment.

Frequency constraint:

$$M_{ij} \le \sum_{k \in L} f_k \le T_{ij}, \forall (i,j) \in \mathbb{Z} \tag{10}$$

Equation (10) ensures that the total number of trains using the link (i,j) does not exceed a specified maximum, while also ensuring that it satisfies the minimum number of trains required to operate on the link (i, j). This minimum number can be calculated from the total leg traffic loads.

Variables limit and type constraints:

$$\lambda_k^{OD}, f_k \ge 0, \text{integer} \tag{11}$$

$$f_k \le N_k \tag{12}$$

Equation (11) ensures an integer line frequency and an integer passenger flow for each OD pair on the k-th line. For the special case when we know the limitation of trains run on each line design, we can add an upper bound to the frequency as Eq. (12).

Coupling constraint between non-exist nonstop arcs and passenger flows:

$$\lambda_k^{OD} = 0, \forall OD \notin S_{ij} \text{ where } (i,j) \in A_k, \forall k \in L \tag{13}$$

This coupling constraint helps the algorithm find the passenger flows related to the nonstop arc (i,j) that does not exist in a specific line to be set to zero passengers.

2.3.2. *Case without predetermined stopping patterns*

Traditionally, stop plans are pre-specified based on how stations are categorised during the line design process, which is our previous problem formulation. However, with the expansion of railway networks, especially for high-speed trains and the significant changes in travel demand among different origin-destination pairs, it becomes increasingly challenging to produce stops using this straightforward approach. Hence, the quality of rail passenger service is improved considerably by optimising stop plans for train lines. Consequently, the model will be more complex and NP-hard (Park et al., 2013; Wang et al., 2011).

In this model, the set of possible stopping patterns is not provided as input. However, some constraints remain identical to the case with predetermined stopping patterns. Optimal stopping patterns are determined as new decision variables in this modelling scheme (Table 3). The same six objectives from Eqs. (1)–(6) for the case with predetermined stopping patterns are retained here. The differences lie in the decision variables used in the objectives, specific sets utilised in the summation, and the number of summations for certain terms.

- Objective 1-NP: Total operating cost and total passenger travel time

The objective in this problem remains the same as the previous one, aiming to minimise two elements: The total operating cost and passenger journey time – we refer to it as objective 1-NP. However, this objective involves different units between the first (monetary) and the second element (time-unit).

$$\min \sum_{k \in L} c_k f_k + VoT \sum_{k \in L} \sum_{OD \in S} \sum_{ij \in A} t_{ij} \lambda_{k,ij}^{OD} \tag{14}$$

- Objective 2-NP: Total passenger travel time

Without predetermined stopping patterns, the total travel time of passenger is given below – we refer to it as objective 2-NP.

$$\min \sum_{k \in L} \sum_{OD \in S} \sum_{ij \in A} t_{ij} \lambda_{k,ij}^{OD} \tag{15}$$

- Objective 3-NP: Total operating cost

This objective is the same as the objective 3-P in Eq. (3) – we refer to it as objective 3-NP.

- Objective 4-NP: Total profit

Total profit for this case – we refer to it as objective 4-NP – can be seen as below.

$$\max \sum_{k \in L} \sum_{OD \in S} \sum_{ij \in A} P_{k,ij}^{OD} \lambda_{k,ij}^{OD} - \sum_{k \in L} c_k f_k \tag{16}$$

- Objective 5-NP: Empty seats

In this objective function – we refer to it as objective 5-NP – we change a set A_k from objective 5-P with A in the summation over arcs (i,j) representing nonstop travel. This objective function has a non-linear term, i.e., the first term containing two decision variables.

$$\min \sum_{k \in L} \sum_{ij \in A} \left[x_{k,ij} C f_k - \sum_{OD \in S_{ij}} \lambda_{k,ij}^{OD} \right] \tag{17}$$

- Objective 6-NP: Empty seat–kilometres

Similar to above, in this objective function – we refer to it as objective 6-NP – we change a set A_k from objective 6-P with A in the summation over arcs (i,j) representing nonstop travel on the r-th route.

$$\min \sum_{k \in L} C f_k \Omega - \sum_{k \in L} \sum_{ij \in A} \sum_{OD \in S_{ij}} \lambda_{k,ij}^{OD} d_{ij} \tag{18}$$

- Constraints

For a case without predetermined stopping patterns, we consider the following constraints.

Train flow conservation constraint:

$$\sum_{j:ij \in A} x_{k,ij} - \sum_{j:ji \in A} x_{k,ji} = \begin{cases} x_k & i = s \\ -x_k & i = t \\ 0 & \text{otherwise} \end{cases}, \forall i \in N, \forall k \in L \tag{19}$$

Equation (19) ensures that for each station, the sum of the train flow into the station must be equal to the flow out of the station in every line.

Passenger flow conservation constraint:

$$\sum_{j:ij \in A} \lambda_{k,ij}^{OD} - \sum_{j:ji \in A} \lambda_{k,ji}^{OD} = \begin{cases} \lambda_k^{OD} & i = o \\ -\lambda_k^{OD} & i = d \\ 0 & \text{otherwise} \end{cases}, \forall OD \in S, \forall k \in L \tag{20}$$

Equation (20) ensures that for the station used for stopping, the sum of the passenger flow into the station must be equal to the passenger flow out of that station in every line.

Capacity constraint:

$$\sum_{OD \in S_{ij}} \lambda_{k,ij}^{OD} \le C f_k, \forall (i,j) \in A, \forall k \in L \tag{21}$$

Equation (21) represents capacity constraint, where the sum of ODs demand passing the nonstop arc (i,j) within the k-th line should not exceed the total seating capacity of the k-th line, calculated as the product of frequency and seating capacity.

Coupling constraint between train frequencies and exist lines:

$$f_k \le M x_k, \forall k \in L \tag{22}$$

Coupling constraint between passenger assignments and exist lines:

$$\sum_{OD\in S_{ij}} \lambda_{k,ij}^{OD} \le M x_{k,ij}, \forall (i,j) \in A, \forall k \in L \quad (23)$$

Coupling constraint on passenger assignment variables:

In addition to the above coupling constraints, this following constraint is considered in our proposed model.

$$\lambda_{k,ij}^{OD} \le \lambda_k^{OD}, \forall (i,j) \in A, \forall k \in L \quad (24)$$

Passenger flow constraint based on *OD* passenger demand:

$$\sum_{k\in L} \lambda_k^{OD} \ge D^{OD}, \forall OD \in S \quad (25)$$

Passenger flow constraint based on leg traffic loads:

Equation (9) is applied.

Frequency constraint:

Equation (10) is applied.

Variable limit and type constraint:

$$x_k, x_{k,ij} \in \{0,1\}, f_k, \lambda_k^{OD}, \lambda_{k,ij}^{OD} \ge 0 \text{ and integer} \quad (26)$$

These constraints ensure an integer line frequency, an integer passenger flow for each *OD* pair on *k*-th line and a binary variable for x_k and $x_{k,ij}$. In common with the case of predetermined stopping patterns, for the particular case when we know the limitation of trains run on each line design, we can add an upper bound to the frequency like Eq. (12).

The problems defined above are all IP problems. They can be addressed using both linear and non-linear solvers, depending on the nature of the objective function and constraints. Solvers such as Gurobi, CPLEX, and MATLAB are effective for solving linear IP problems. When the objective function or constraints involve non-linear relationships, the problem becomes a non-linear IP. Solving such problems is generally more challenging, and specialised solvers like BARON or certain configurations of global optimisation solvers in MATLAB may be employed. Therefore, we will utilise MATLAB solvers to address the problems considered in this paper. To demonstrate the effectiveness of the non-linear solver embedded in MATLAB, typically a heuristic method, in contrast to its linear solver, we provide a benchmark utilising our case study (refer to Appendix A).

3. Numerical experiments

In order to illustrate the proposed approach, we model and solve the problem of line planning problem to determine frequencies with and without stopping patterns given in advance using a case study of future semi high-speed rail in Indonesia, Jakarta–Surabaya.

3.1. Data preparation

The Indonesian Government is considering replacing the existing Jakarta–Surabaya executive trains when operating the semi high-speed train on the new railway line. The Jakarta–Surabaya semi high-speed train project will connect Jakarta to Surabaya along 715 km via a new track on the northern route in Java Island. The new track will be built parallel to the existing track on the north side and provided exclusively for the Jakarta–Surabaya semi high-speed train. There will be some intermediate stations between Jakarta and Surabaya used as stopping stations for boarding and alighting, and operating stations or passing loops for train crossing and overtaking due to the new track being a single-track only. This new track is designed to have a maximum speed of 160 km/h, with an average target speed of 130 km/h, resulting in a travel time from Jakarta to Surabaya of 5.5 h. Fig. 5 shows the track plan of a semi high-speed train in the Jakarta–Surabaya corridor.

As shown in Fig. 5, the network is comprised of 5 major stations, i.e., Jakarta Manggarai station (MRI), Cikampek station (CKP), Cirebon station (CN), Semarang Tawang station (SMT), and Surabaya Pasar Turi station (SBI). For this example, we only consider stations MRI, CN, SMT, and SBI as stopping stations in our model due to data availability and with only one-way direction. Table 4 summarises the most important topological characteristics.

We assume that the travel time of a passenger is defined as the time between arriving at the origin station and arriving at the destination station. If the stopping time at each station for boarding and alighting is designed to be around 5 min, then for example, nonstop arc (1,3) or MRI to SMT will have the travel time as (5 + 96 + 104) min = 205 min. The stopping time at station 1 is added to the travel time.

As the Jakarta–Surabaya semi high-speed train is a new service planned by the Indonesian government, there is no existing data available for *OD* demand at its stations. Therefore, we rely on *OD* demand information from the conventional train, the Jakarta–Surabaya executive train in 2017 (Table 5), which is planned to be replaced by the Jakarta–Surabaya semi high-speed train. However, we assume a 50% increase in *OD* demand for the year 2025, when this train is predicted to be operational. This assumption is based on rough estimates of modest passenger growth and the displacement of passengers from other modes, such as airplanes (please refer to Appendix B for the rationale behind this increase). Consequently, the *OD* demand details in Table 5 will be

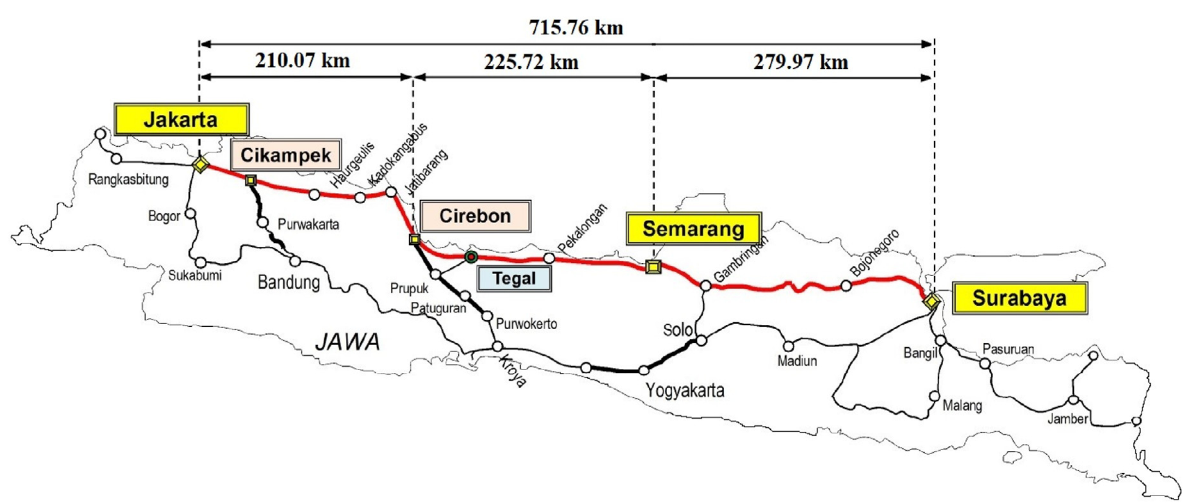


Fig. 5. Map of the Jakarta–Surabaya semi high-speed network.

Table 4
Length of each track segment of Jakarta–Surabaya high-speed railway network with its nonstop travel time.

Segment number	From	Station number	To	Station number	Length (km)	Nonstop travel time	Average speed (km/h)
1	MRI	1	CN	2	210	1 h 36 min	131.25
2	CN	2	SMT	3	225	1 h 44 min	129.81
3	SMT	3	SBI	4	280	2 h 10 min	129.23
Total distance		—	—	—	715	—	—

Table 5
Passenger demand data information (daily demand) for Jakarta–Surabaya executive train in the 2017 (Utomo et al., 2020).

O/D	Jakarta	Cirebon	Semarang	Surabaya
Jakarta	0	1,306	888	1,781
Cirebon	1,206	0	80	93
Semarang	820	73	0	574
Surabaya	1,644	86	530	0

Table 6
Passenger demand data information (daily demand) for Jakarta–Surabaya semi high-speed train (predicted 50% increase in demand by 2025).

O/D	Jakarta	Cirebon	Semarang	Surabaya
Jakarta	0	1,959	1,332	2,672
Cirebon	1,809	0	120	140
Semarang	1,230	110	0	861
Surabaya	2,466	129	795	0

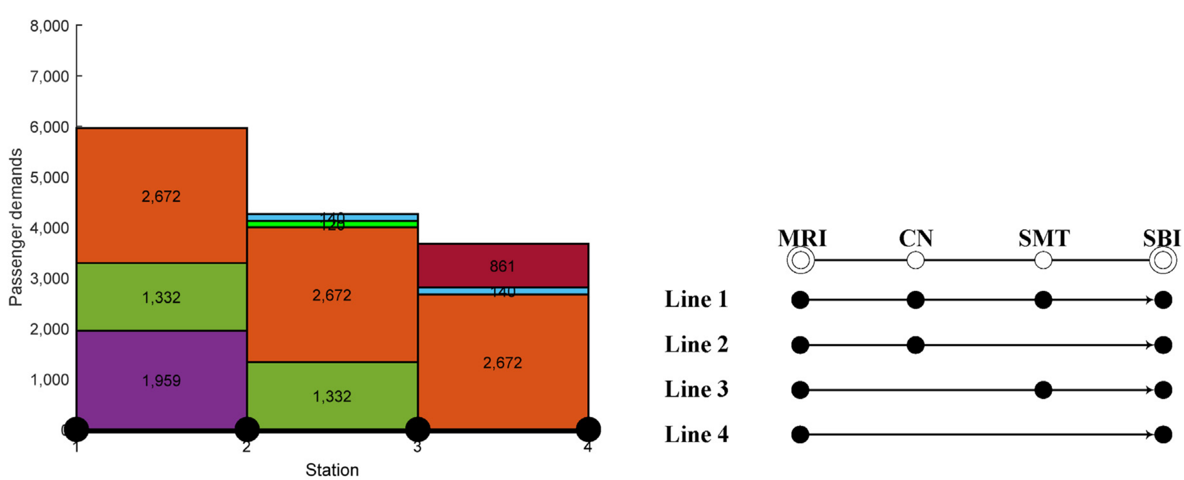
adjusted to reflect this increase, as shown in Table 6.

The minimum number of necessary trains is the same as the maximum number of required trains across links or segments. This number is calculated over links for an uni-directional double-track railway because different links represent different tracks. For bi-directional single-track, this number is determined over segments because the two directions are using the same track. The first task is to analyse the number of passengers passing through each link (or segment), as shown in Fig. 6a. Passengers starting at station 1 and reaching station 3 will naturally traverse through link (2,3), as shown in the moss green area in Fig. 6a. From the calculation of leg traffic load, the passenger flows through link (1,2), link (2,3), and link (3,4) for the upstream direction (Jakarta to Surabaya) are 5,963, 4,264, and 3,673; respectively. If we consider that the train's maximum capacity is designated as 500, thus for each link it needs at least 12, 9, and 8 trains, respectively. Therefore, the minimum number of necessary trains equals the maximum number needed across links, max {12, 9, 8} = 12.

The possible stopping patterns used in the first scenario can be observed in Fig. 6b. The operational cost for each train on line in the Jakarta–Surabaya route (c_k), pertains only to running a train in the network, and the values are expected to vary with specified lines. However, in this paper, we assume that the operational cost of running a train incorporates capital cost, operating cost (fuel cost), and maintenance cost, with the value being uniform across all lines. In our example, we define the ticket fares in Table 7 for each *OD* pair on the Jakarta–Surabaya route in the network. The value of time (*VoT*) from passengers using the Jakarta–Surabaya executive train, who move to the semi high-speed train is assumed to be 48,295 IDR/h. In this experimental case, we convert this value to minute units, resulting in 805 IDR/min. With this data and these assumptions, we can populate all the input parameters used in our model in both cases with and without stopping patterns given in advance, as provided in Appendix C.

Table 7
Operation cost data and ticket fare design for each *OD* pair (in Indonesian currency, IDR).

	Line 1	Line 2	Line 3	Line 4
Operation cost, c_k	110,000,000	110,000,000	110,000,000	110,000,000
Ticket fare, P_k^{OD}				
OD 1,2	162,000	162,000	—	—
OD 1,3	310,000	—	335,000	—
OD 2,3	173,000	—	—	—
OD 1,4	500,000	525,000	525,000	550,000
OD 2,4	363,000	388,000	—	—
OD 3,4	215,000	—	215,000	—



(a) (b)

Fig. 6. (a) Number of passengers traversing each link for upstream direction (Jakarta to Surabaya); (b) set of possible stop patterns for Jakarta–Surabaya semi high-speed train network.

Table 8
Summary of results with maximum trains run on each line is 3 ($N_k = 3, \forall k \in L$) and 4 stopping patterns (4 lines).

	Objective 1-P	Objective 2-P	Objective 3-P	Objective 4-P	Objective 5-P	Objective 6-P
No. of trains (unit)	12	12	12	12	12	12
No. of line 1 (1–2–3–4)	3	3	3	3	3	3
No. of line 2 (1–2–4)	3	3	3	3	3	3
No. of line 3 (1–3–4)	3	3	3	3	3	3
No. of line 4 (1–4)	3	3	3	3	3	3
Operation cost (IDR)	**1,320,000,000**	**1,320,000,000**	**1,320,000,000**	**1,320,000,000**	**1,320,000,000**	**1,320,000,000**
Total travel time (min)	**1,534,674**	**1,534,674**	1,539,694	**1,534,674**	1,540,764	1,540,064
Total travel time cost (IDR)	**1,235,412,570**	**1,235,412,570**	1,239,453,670	**1,235,412,570**	1,240,315,020	1,239,751,520
Total profit (IDR)	**1,144,073,000**	**1,144,073,000**	1,118,973,000	**1,144,073,000**	1,113,623,000	1,117,123,000
No. of empty seats	3,744	3,744	2,740	3,744	**2,526**	2,666
Empty seat–kilometres (empty seat–km)	**1,049,930**	**1,049,930**	**1,049,930**	**1,049,930**	**1,049,930**	**1,049,930**
Average computational time (s)	2.114	1.999	2.080	2.078	1.965	1.977

Table 9
Summary of results with maximum trains run on each line is 10 ($N_k = 10, \forall k \in L$) and 4 stopping patterns (4 lines).

	Objective 1-P	Objective 2-P	Objective 3-P	Objective 4-P	Objective 5-P	Objective 6-P
No. of trains (unit)	12	14	12	12	12	12
No. of line 1 (1–2–3–4)	1	1	4	1	1	1
No. of line 2 (1–2–4)	3	4	0	3	4	4
No. of line 3 (1–3–4)	3	3	8	3	3	3
No. of line 4 (1–4)	5	6	0	5	4	4
Operation cost (IDR)	**1,320,000,000**	1,540,000,000	**1,320,000,000**	**1,320,000,000**	**1,320,000,000**	**1,320,000,000**
Total travel time (min)	1,529,674	**1,528,814**	1,542,894	1,529,674	1,534,259	1,534,074
Total travel time cost (IDR)	1,231,387,570	**1,230,695,270**	1,242,029,670	1,231,387,570	1,235,078,495	1,234,929,570
Total profit (IDR)	**1,169,073,000**	953,373,000	1,102,973,000	**1,169,073,000**	1,146,148,000	1,147,073,000
No. of empty seats	2,744	4,416	4,100	2,744	**2,327**	2,364
Empty seat–kilometres (empty seat–km)	**1,049,930**	1,764,930	**1,049,930**	**1,049,930**	**1,049,930**	**1,049,930**
Average computational time (s)	2.296	1.917	2.095	2.089	1.966	1.978

3.2. Results of case with predetermined stopping patterns

With various objective functions from our proposed model, we compare the results of each line planning optimisation for a case with predetermined stopping patterns. In our experiment using the Jakarta–Surabaya semi high-speed railway network, the model efficiently finds an optimal solution within a brief period using the MATLAB solver. Tables 8 and 9 show the results of line planning where the maximum number of trains running on each line are set to 3 and 10, respectively. The maximum number of trains running on each line is changed so that we can see the influence of this parameter on the results. From Tables 8 and 9, setting a larger number for the maximum number of trains running on each line to 10 achieves better results. It is possible to change further parameters to achieve a better line plan design. For example, setting $N_k = 10$, can decrease empty seats for scenario objectives 1-P, 4-P, 5-P, and 6-P but has worse values for scenario objectives 2-P and 3-P. For scenario objective 2-P, the limitation of the maximum number of trains run on each line is set to 3, resulting in 3 trains on each line with 12 trains in total for the network. There are no more options to accommodate all *OD* passenger demand. When the limit is set to 10, the model has the flexibility to adjust the optimal number of trains running on each line as long as the total passenger travel time is minimised. Its total passenger travel time is lower than $N_k = 3$. For scenario objective 3-P, the impact of setting different parameters for N_k can be better explained if we include variable costs in our proposed model, which is not considered yet.

The total operating costs and empty seat–kilometres exhibit a similar pattern of values in these tables due to their correlation with the total number of trains in the network and the identical start and stop stations for all lines. For scenario objective 2-P, the total number of trains is 14 for $N_k = 10$,. Since the fixed cost is applied for running a train in the network, running 14 trains will cost more than running only 12 trains. Moreover, for empty seat–kilometres, except for scenario using objective 2-P with $N_k = 10$, the results are exactly the same because we have fixed demand for all *OD* pairs. As long as the total frequency of trains running in the network remains the same across all scenarios, empty seat–kilometres will always be the same.

From the passenger assignment results illustrated in Fig. 7, we can improve the line plan designs by changing the stopping patterns, which have been influenced by previous results. For example, we change the stopping patterns for line 1 (Fig. 6b) to have only three stops (1–2–3) – this line then belongs to a different route – and re-run the optimisation for all scenarios, including for objectives 5-P and 6-P which already have good results for empty seats and empty seat–kilometres, to see whether their results can be improved. The outcomes in Table 10 show improvements in all scenarios, especially in terms of empty seats and empty seat–kilometres. For objectives 5-P and 6-P, total travel time decreased, and total profit increased.

3.3. Results of case without predetermined stopping patterns

We now use the same data as before for the operational cost of running a train in the network, ticket fares for each *OD* pair and value of travel time but consider the case where stopping patterns are not predetermined. Even so, the desired number of different stopping patterns (or lines) must still be given as an upper bound. For this case, the input parameters can be seen in Appendix C. First, we set the maximum number of desired stopping patterns (or lines) on the route to 4 ($K = 4$) and second, we decrease it to 2 ($K = 2$). Tables 11 and 12 show the comparison of the two simulations. For the first simulation with $N_k = 10$, $\forall k \in L$ and $K = 4$, the results for all optimal objective values (highlighted in bold in Table 11) are identical compared to a case where stopping patterns are given in advance, where $N_k = 10, \forall k \in L$ with 4 stopping patterns (Table 9). The only exception is for objective 5-NP, where there is a very slight variation in the optimal empty seats. This discrepancy is attributed to different solving methods in MATLAB, influenced by the nature of the problems, which are non-linear IP. Although some terms related to other objective functions (those not in bold) may not be as favourable as in the previous case, the optimisation results for stopping patterns not given in advance are still quite satisfactory. This is because the optimal objective values for each scenario, from objective 1-NP to 6-NP, are still achieved. The difference between these solutions lies in the passenger flow assignment results, which can be observed in Fig. 8. This indicates that the problem has alternative optimal solutions. It is common for most problems, particularly linear programming, to have multiple optimal solutions sourced from practical scenarios. These alternative optimal represent situations wherein a different set of reactions can be used by the system to reach the exact same quantitative objective value.

For two scenarios in Table 11 and i.e., objectives 3-NP and 6-NP, the line planning results in only two stopping patterns. If we set the operational cost for running a train in the network (c_k) as seen in Table 7 to non-identical values for all lines, the results could potentially improve. The total operation costs for all scenarios would vary because running two lines and four lines might incur different expenses, even if the total number of trains running in the network remains the same. For example, running a train with more stop schemes could lead to higher energy costs, resulting in higher expenses compared to running a direct train without stops. Consequently, the total operation costs for running 12 trains in the network with two lines (stopping patterns) and four lines would differ. We acknowledge this as our limitation, and will be improved in further research. Similar to prior simulations, we discovered that operating costs and empty seat–kilometres are closely related to the total number of trains in the network as well as the route. The value of empty seat–kilometres will always be the same across all scenarios if there are the same number of trains running in the network and the same start and stop stations for all lines (the same route), since the *OD* demands are fixed.

For the second simulation, when we set the desired number of stopping patterns to 2 ($K = 2$), the overall results for all objectives are not quite satisfactory compared to $K = 4$. In terms of empty seats, for example, having 4 stopping patterns (lines) results in less capacity loss, although for empty seat–kilometres, all values remain the same due to the total of 12 trains running in the network (and 15 trains for objective 2). This indicates that allowing more stop schemes improves the performance of line plan designs.

All experiments for both cases with and without predetermined stopping patterns are carried out in a MATLAB R2022a environment on a personal computer equipped with an Intel Core i5 processor running at 2.50 GHz and 16 GB of RAM. The computer operates on Windows 11 Education 64-Bit.

4. Discussion

4.1. Choices of objectives

Following the results for both cases with and without predetermined stopping patterns, we found that our proposed model allows changing the parameters expected to improve the quality of the solutions resulting from model optimisation. From the comparative analysis, railway planners can examine which scenario suits their conditions and make parameter adjustments to assess potential improvements.

In cases like new service development, such as the Jakarta–Surabaya semi high-speed train in our experimental case, planners might be primarily interested in maximising efficiency (travel time and capacity) and minimising costs for the initial periods of operation. Considering expected passenger demand, planners might limit the maximum number of trains passing through each line, starting with a small number first and adjusting to the maximum number they can afford to provide. Setting a higher number of trains that can run on each line naturally allows the model to be flexible, adjusting the optimal number of trains running on each line as long as the efficiency parameters are achieved.

During the initial phase of operation, it is advisable for railway planners to prioritise objectives 1, 2, 3, 5, or 6, focusing on maximising technical efficiency (such as minimising travel time and maximising capacity) and minimising costs. However, in the long run, as the market demand for the Jakarta–Surabaya semi high-speed train system grows, planners may shift their focus to more commercial concerns through profit maximisation. This could involve implementing a dynamic pricing policy to maximise revenue and minimise costs, making objective 4 the primary focus of line plan designs. It is worth noting that the profit shown in our numerical experiment is based on fixed demand at one time period. Therefore, the results using objectives 1 and 4 are similar, indicating that objective 4 in long-term usage could potentially result in the same quality as objective 1, which considers passenger satisfaction. With sustained passenger engagement, a profit-oriented line plan has the potential to yield higher profits with increasing passenger demand compared to the initial plan.

In real-world applications, railway planners will conduct comprehensive reviews routinely for their planning strategies, focusing on the objective as necessary. The frequency of this review depends on some conditions, such as changes in demand, technological advancement, economic conditions, shifting priorities or changes in policy regulation. It could be an annual assessment or an even more frequent review. Our model offers an adaptable optimisation approach to address this challenge, accommodating various objectives based on the contextual requirements during adjustment. Our approach provides a more comprehensive perspective by generating a wide range of solutions rather than a single solution. We then use comparative analysis to evaluate their performance to obtain valuable insights for policy recommendations.

In this study, adjusting objectives in the line planning can be useful for optimising different time horizons since the model allows for the proposal of short-term, medium-term, and long-term line plans. In the short term, the emphasis may be on efficiently meeting current demand. In the medium term, the focus may be on increasing passenger engagement. In the long term, the focus could shift to achieving financial viability and accommodating future growth.

In addition, this model allows the consideration of various scenarios in line planning design, such as different optimal passenger flow assignments. Between cases with and without predetermined stopping patterns, even when the frequency of each line remains the same for the same objective, the passenger flow could differ, representing alternative optimal solutions. In practice, alternative optimal are advantageous as they provide decision-makers with a range of choices without compromising the objective function (Tsai et al., 2008). This flexibility enables the planning scheme to strategically allocate passengers based on an analysis of alternative assignments.

Furthermore, the optimal passenger flow assignment results, obtained through various objective schemes, illustrate how the distribution of passenger flow varies with changes in line plan objectives. This valuable information can be effectively utilised to enhance ticket-selling management. For instance, the insights gained from passenger flow assignments shed light on the popularity of different routes and lines. This knowledge is instrumental in implementing optimal pricing strategies, as previously mentioned, such as dynamic pricing, where ticket prices are adjusted based on demand patterns. Although the benefits of this approach may be more evident in the long-term operational phase, it remains a crucial consideration.

Additionally, this information helps the railway planners optimise capacity planning by adjusting the number of available tickets based on expected demand. This ensures capacity aligns with the number of passengers assigned to each route and service. Another aspect that could possibly derived from passenger flow assignment is the potential for enhancing service provision. For example, for routes and lines experiencing high demand, operators may provide additional services to upgrade the passenger experience, attracting more customers for future reference. Conversely, promotions and marketing techniques can improve routes and lines with lower assignments to raise the demand. Such strategies can be valuable for informing policy recommendations in future planning endeavours.

4.2. Choices of stopping patterns

As demonstrated, we conclude that determining stopping patterns in advance offers several advantages, including the ability to rerun the model for improved results based on initial optimisation. This proves

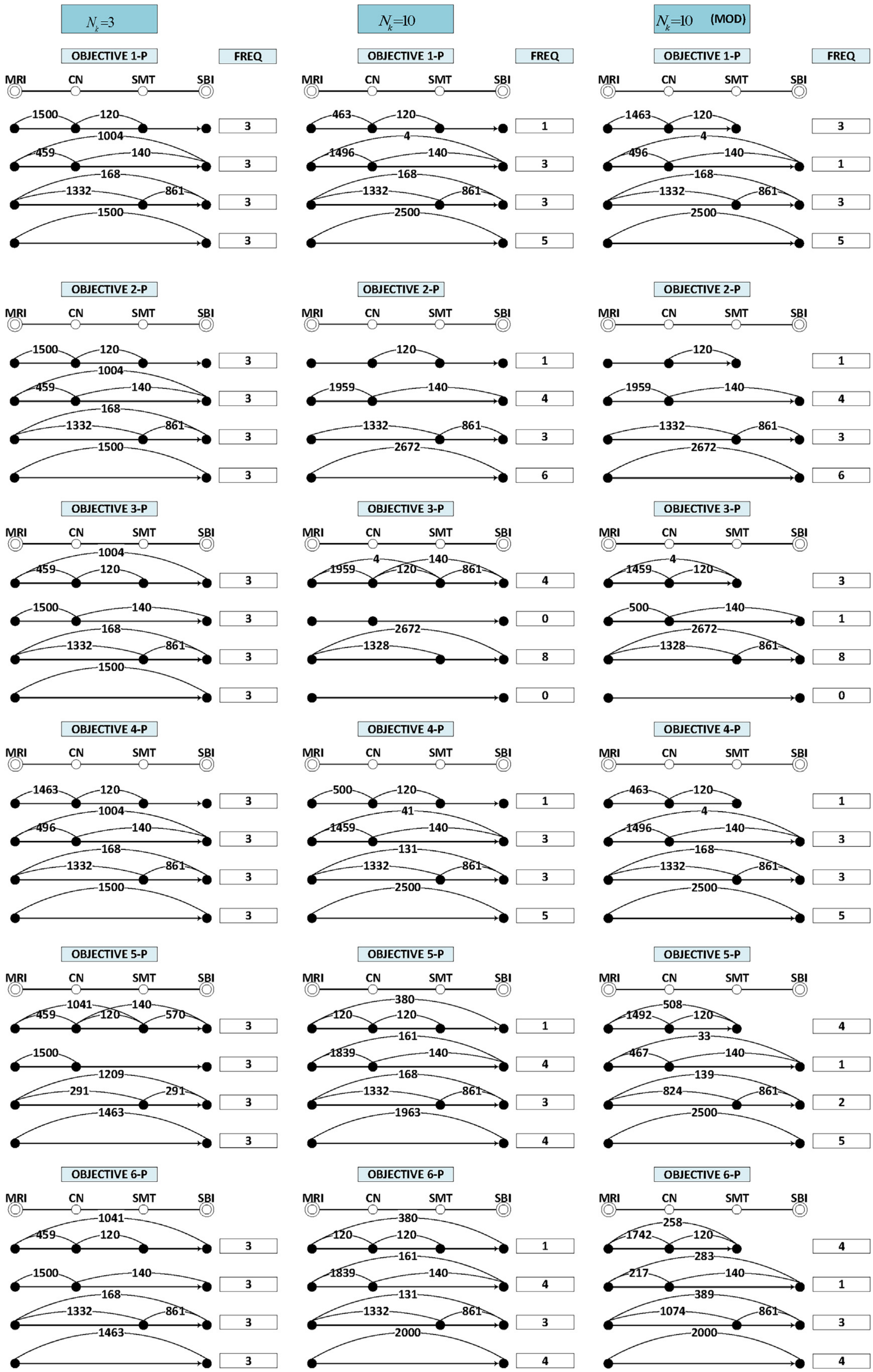


Fig. 7. Passenger flow assignments for all scenarios in a case with predetermined stopping patterns and changing parameters.

Table 10
Summary of results with maximum trains run on each line is 10 ($N_k = 10, \forall k \in L$) and 4 modified stopping patterns (4 modified lines).

	Objective 1-P	Objective 2-P	Objective 3-P	Objective 4-P	Objective 5-P	Objective 6-P
No. of trains (unit)	12	14	12	12	12	12
No. of line 1 (1–2–3)	3	1	3	1	4	4
No. of line 2 (1–2–4)	1	4	1	3	1	1
No. of line 3 (1–3–4)	3	3	8	3	2	3
No. of line 4 (1–4)	5	6	0	5	5	4
Operation cost (IDR)	**1,320,000,000**	1,540,000,000	**1,320,000,000**	**1,320,000,000**	**1,320,000,000**	**1,320,000,000**
Total travel time (min)	1,529,674	**1,528,814**	1,542,194	1,529,674	1,532,214	1,533,464
Total travel time cost (IDR)	1,231,387,570	**1,230,695,270**	1,241,466,170	1,231,387,570	1,233,432,270	1,234,438,520
Total profit (IDR)	**1,169,073,000**	953,373,000	1,106,473,000	**1,169,073,000**	1,156,373,000	1,150,123,000
No. of empty seats	2,244	3,916	2,240	2,244	**1,736**	1,986
Empty seat–kilometres (empty seat–km)	629,930	1,624,930	629,930	909,930	**489,930**	**489,930**
Average computational time (s)	2.125	1.962	2.089	2.080	1.957	1.985

Table 11
Summary of results with maximum trains run on each line is 10 ($N_k = 10, \forall k \in L$) and $K = 4$.

	Objective 1-NP	Objective 2-NP	Objective 3-NP	Objective 4-NP	Objective 5-NP	Objective 6-NP
No. of trains (unit)	12	15	12	12	12	12
Stops & No. of line 1	1–2–3–4 & 3	1–2–4 & 2	1–3–4 & 3	1–4 & 5	1–3–4 & 2	1–2–3–4 & 10
Stops & No. of line 2	1–3–4 & 3	1–3–4 & 3	– & 0	1–2–3–4 & 1	1–2–4 & 7	1–2–4 & 2
Stops & No. of line 3	1–4 & 5	1–2–3–4 & 2	1–2–3–4 & 9	1–2–4 & 3	1–3–4 & 1	– & 0
Stops & No. of line 4	1–2–4 & 1	1–4 & 8	– & 0	1–3–4 & 3	1–2–3–4 & 2	– & 0
Operation cost (IDR)	**1,320,000,000**	1,650,000,000	**1,320,000,000**	**1,320,000,000**	**1,320,000,000**	**1,320,000,000**
Total travel time (min)	1,529,674	**1,528,814**	1,555,579	1,529,674	1,546,574	1,562,894
Total travel time cost (IDR)	1,231,387,570	**1,230,695,270**	1,252,241,095	1,231,387,570	1,244,992,070	1,258,129,670
Total profit (IDR)	**1,169,073,000**	843,373,000	1,039,548,000	**1,169,073,000**	1,084,573,000	1,002,973,000
No. of empty seats	3,744	4,916	4,063	2,744	**2,364**	3,100
Empty seat–kilometres (empty seat–km)	**1,049,930**	2,122,430	**1,049,930**	**1,049,930**	**1,049,930**	**1,049,930**
Average running time (s)	2.770	2.513	2.528	2.725	146.556	2.504

Table 12
Summary of results with maximum trains run on each line is 10 ($N_k = 10, \forall k \in L$) and $K = 2$.

	Objective 1-NP	Objective 2-NP	Objective 3-NP	Objective 4-NP	Objective 5-NP	Objective 6-NP
No. of trains (unit)	12	15	12	12	12	12
Stops & No. of line 1	1–2–3–4 & 7	1–4 & 8	1–2–3–4 & 10	1–4 & 5	1–2–3–4 & 3	1–2–3–4 & 6
Stops & No. of line 2	1–4 & 5	1–2–3–4 & 7	1–4 & 2	1–2–3–4 & 7	1–2–4 & 9	1–2–3–4 & 6
Operation cost (IDR)	**1,320,000,000**	1,650,000,000	**1,320,000,000**	**1,320,000,000**	**1,320,000,000**	**1,320,000,000**
Total travel time (min)	1,537,894	**1,536,174**	1,553,264	1,537,894	1,549,074	1,562,894
Total travel time cost (IDR)	1,238,004,670	**1,236,620,070**	1,250,377,520	1,238,004,670	1,247,004,570	1,258,129,670
Total profit (IDR)	**1,127,973,000**	806,573,000	1,051,123,000	**1,127,973,000**	1,072,073,000	1,002,973,000
No. of empty seats	4,100	5,944	4,026	4,100	**2,364**	4,100
Empty seat–kilometres (empty seat–km)	**1,049,930**	2,122,430	**1,049,930**	**1,049,930**	**1,049,930**	**1,049,930**
Average running time (s)	2.466	2.251	2.236	2.365	89.051	2.173

beneficial for smaller networks. However, for larger networks, it may be impractical to predefine stopping patterns due to the vast number of possible station combinations. In such cases, applying the model without predetermined stopping patterns for line planning problems becomes advantageous, eliminating the need for pre-processing to determine all potential stop schemes in the network. Giving an appropriate prediction of the number of stopping patterns (lines) number is also important, since low numbers of stopping patterns yield lower performance due to a lack of service options offered for passengers.

Furthermore, optimised stopping patterns provide flexibility. For instance, operators can swiftly adjust routes and lines to accommodate changes in demand, infrastructure issues, or other unforeseen circumstances, without experiencing any deterioration in the objective function. This approach demonstrates higher adaptability to changing conditions, particularly in larger networks. However, it is important to note that we have not considered a larger network in this study. We acknowledge that addressing larger and more realistic network problems requires a sophisticated algorithm and warrants exploration in future research.

5. Conclusions and further research

5.1. Conclusions

We presented an integer linear programming optimisation model with various objective functions (and one non-linear objective for a case without predetermined stopping patterns) to identify all viable line plan designs for the railway network. The model considers both passengers' and operators' perspectives. Our proposed model generates diverse line plan designs, providing railway operators with options for short-term, medium-term, and long-term plans. The results exhibit varying train frequencies and passenger assignments for different objectives. In the context of introducing a new mode, such as a high-speed train, this modelling approach assesses multiple proposals, enabling comparative analysis to evaluate efficiency parameters and offering insights for policy recommendations. Additionally, the passenger assignment results offer valuable information for enhancing ticket-selling management such as implementing optimal pricing strategies, adjusting ticket availability based on expected demand, and enhancing service customisation.

Two scenarios of stopping patterns are considered. With predetermined stopping patterns, the modelling approach has advantages,

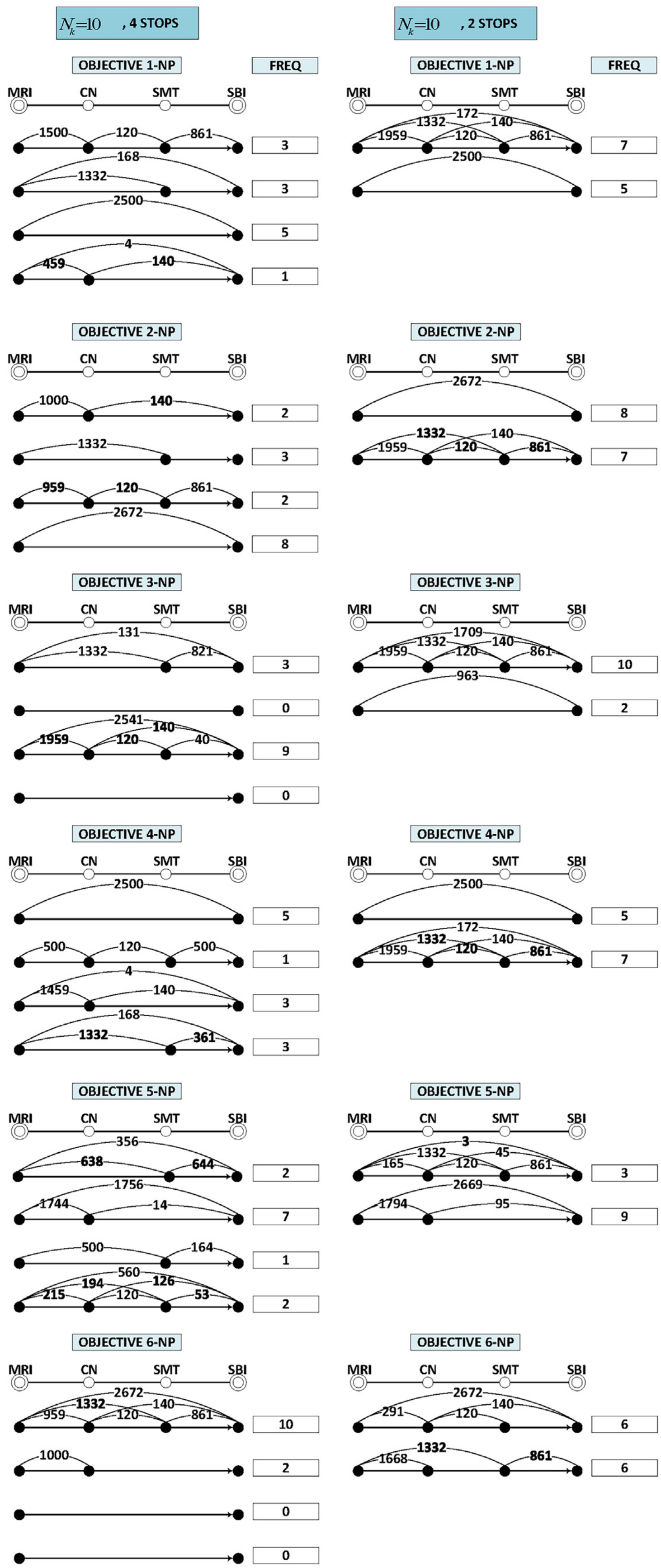


Fig. 8. Passenger flow assignments for all scenarios in a case without predetermined stopping patterns.

such as the ability to rerun the model for better results based on initial optimisation, making it useful for exploring improved line plan designs in small networks. However, it becomes challenging for larger networks due to the exhaustive nature of giving the design of stops in advance, considering the massive number of possible combinations.

In conclusion, both of our proposed models allow for setting different objectives to obtain various scenarios of optimal line plans. This flexibility aids in determining better planning strategies for new railway service development, assisting railway planners in shaping the future of these services with high-quality line plan designs for short-, medium-, and long-term goals. Exploring numerous options to maximise benefits is feasible with these models.

5.2. Further research

Firstly, a larger network with more stations and segments is needed to have a broader and more complex outlook for new railway services development. Secondly, we can improve the cost of operating a train by considering factors such as the time needed for completing a journey, energy costs, and other potential variable costs. For example, running a train with more stops will incur higher energy costs compared to running a direct train or one with fewer stops. Thirdly, we will further develop more efficient algorithms for solving the line planning problem with various objective functions, particularly for larger networks. Fourthly, as line plan results currently do not integrate trains' arrival and departure times, accurately determining passengers' train choices based on actual journey time becomes almost impossible during the line planning phase. To address this, we intend to integrate line planning and timetable optimisation with various objective functions, introducing a new modelling approach. For this, we can also consider time-varying demand in the model. Additionally, we will explore the potential of a multi-objective optimisation approach for the integrated problem since, in the timetabling stage, tactical plans might observe broader interests. This comprehensive approach is anticipated to serve as a decision support tool for policymakers in formulating more effective planning strategies for new railway service development.

Replication and data sharing

The authors confirm that the data supporting the findings of this study are available within the article, Appendix A, Appendix B, and Appendix C. The complete code may only be provided with restrictions from the corresponding author upon reasonable request.

CRediT authorship contribution statement

Prasetyaning Diah Rizky Lestari: Conceptualization, Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Ronghui Liu:** Methodology, Supervision, Validation, Writing – review & editing. **Richard Batley:** Methodology, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This study was funded by the Indonesian Endowment Fund for Education (LPDP). We acknowledge the Agency for the Assessment and Application of Technology (BPPT), the former name of the National Research and Innovation Agency of Indonesia (BRIN), for providing research data. Special appreciation goes to Prof. Xiaobo Qu, the Editor-in-Chief, and the two anonymous reviewers for their contributions and insightful comments. Their feedback has significantly enhanced the overall quality of the paper.

Appendix A

The effectiveness of non-linear solver (heuristic method) embedded in MATLAB

We aim to demonstrate the effectiveness of MATLAB's non-linear solver, typically a heuristic method, compared to its linear solver (exact method) in solving our case study involving predetermined stopping patterns. MATLAB offers a non-linear solver based on the genetic algorithm to address non-linear integer programming problems. Table A1 compares the objective function values obtained from the linear solver ('intlinprog' solver) with those from the non-linear solver ('ga' solver). 'intlinprog' solver employs an exact method (Branch-and-Bound) to solve the integer linear programming (ILP) problem. The iterations of both solvers are provided in the chart until the optimal solution is reached. Additionally, the computation times for both solvers are included.

Table A1
Performance comparison between 'intlinprog' solver and 'ga'solver embedded in MATLAB for our case study.

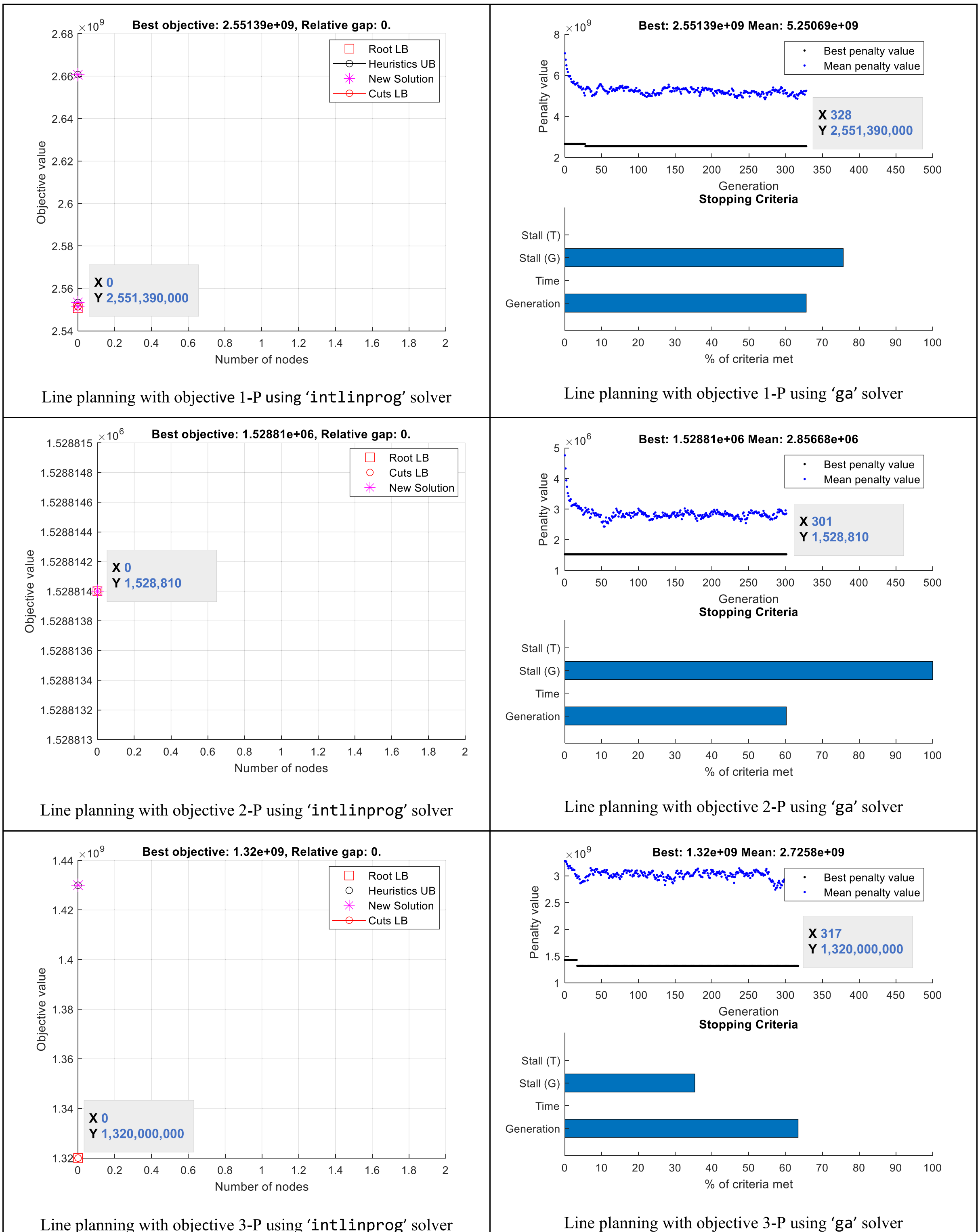
Objective	'intlinprog' solver		'ga' solver	
	Objective value	Computational time (s)	Objective value	Computational time (s)
Objective 1-P	2,551,387,570	2.296	2,551,387,570	71.493
Objective 2-P	1,528,814	1.917	1,528,814	70.313
Objective 3-P	1,320,000,000	2.095	1,320,000,000	71.658
Objective 4-P	1,169,073,000	2.089	1,169,073,000	83.679
Objective 5-P	2,327	1.966	2,327	87.769
Objective 6-P	1,049,930	1.978	1,049,930	52.024

The comparison reveals no difference in the objective function values obtained from both solvers, indicating that the 'ga' solver is also effective for solving our line planning problem, even in scenarios without predetermined stopping patterns. However, the computation time increases for the 'ga' solver when a higher number of generations is set. A high number of generations is required to increase the probability of the 'ga' solver finding the optimal objective values, leading to the optimal solution. In our experiment, we set 'MaxGeneration' to 500 and 'MaxStallGeneration' to 300. 'MaxGeneration' determines the overall maximum number of generations, while 'MaxStallGeneration' controls termination based on the lack of improvement over a specific number of generations. 'Stall (T)' and 'Stall (G)' act as stopping criterions for 'ga' solver finding the optimal objective values. 'Stall (T)' refers to the stopping criterion based on the change in the penalty value or fitness function over a certain number of generations. If the best penalty value does not improve significantly over a specified number of generations, the algorithm is considered to have 'stalled' and stops. 'Stall

(G)' refers to a similar stopping criterion but is based on the average penalty value or fitness function across the population. If the average penalty value does not improve significantly over a specified number of generations, the algorithm stops.

The difference in objective values between Table A1 and Fig. A1 is solely due to rounding in MATLAB; the actual values are identical to those presented in Table A1. Additionally, in our case featuring predetermined stopping patterns, where the optimisation problem is relatively simple, the number of nodes in a Branch-and-Bound algorithm is reported as 0 ('numnodes' = 0). This means 'intlinprog' solved the problem before branching. This is one indication that the result is reliable. Moreover, the presence of the heuristic upper bound in the optimisation process using the 'intlinprog' solver (please refer to the Legend in each chart in Fig. A1) only influences the Branch-and-Bound algorithm's decision-making process. 'intlinprog' remains an exact method for solving ILP problems, ensuring that the obtained solution is provably optimal within the specified tolerance limits.

In summary, the proposed framework in this paper is transferrable and effective for real-world applications.



Line planning with objective 1-P using 'intlinprog' solver

Line planning with objective 1-P using 'ga' solver

Line planning with objective 2-P using 'intlinprog' solver

Line planning with objective 2-P using 'ga' solver

Line planning with objective 3-P using 'intlinprog' solver

Line planning with objective 3-P using 'ga' solver

Fig. A1. Iterations of 'intlinprog' solver and 'ga' solver in solving the case study with predetermined stopping patterns.

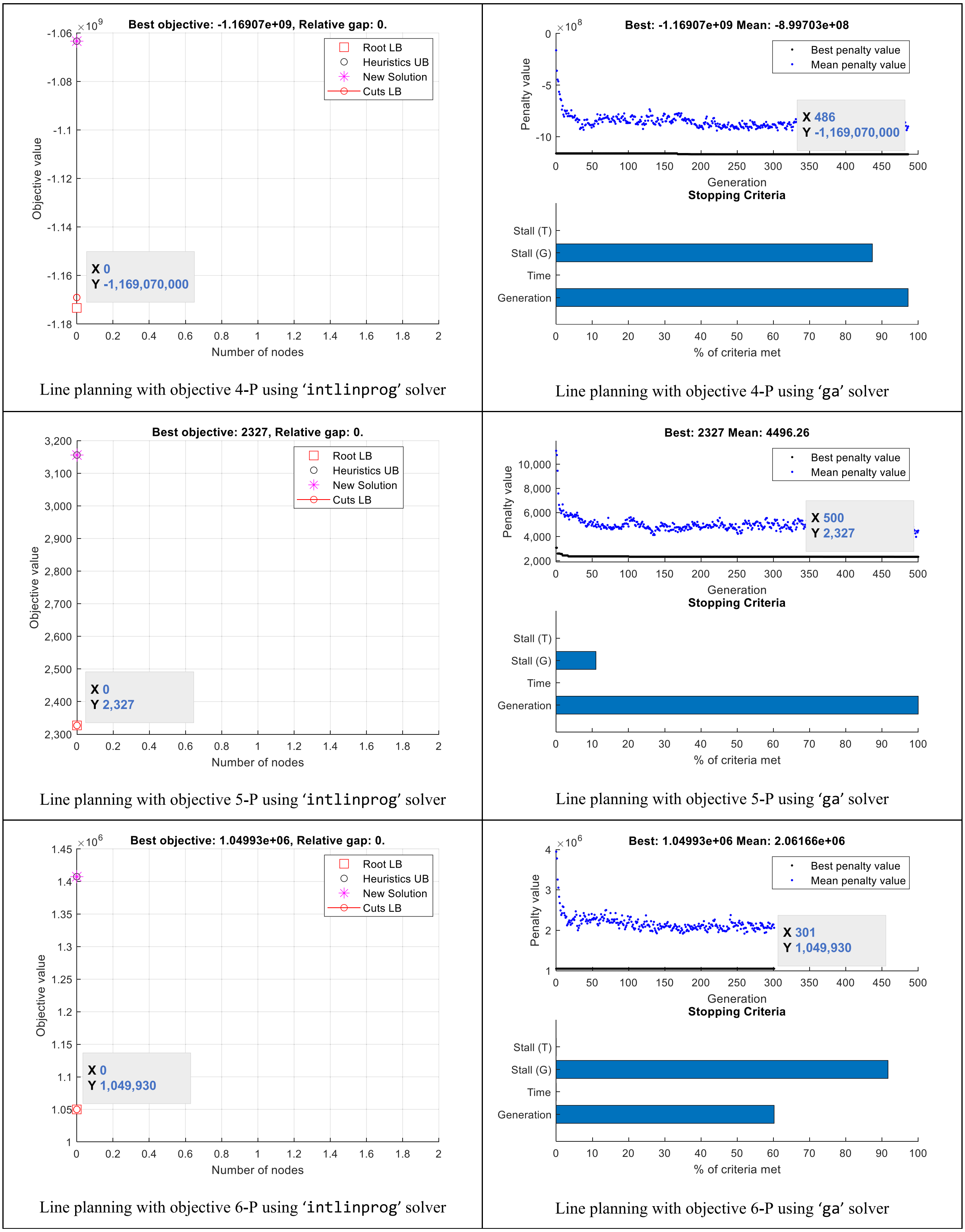


Fig. A1. (*continued*).

Appendix B

Estimating passenger demand for the Jakarta–Surabaya semi high-speed train

Fig. B1 illustrates the overall number of passengers for rail transport and air transport in Indonesia from 2017 to 2023, with projections for 2024 to

2025, as analysed by authors using data from BPS-Statistics Indonesia (2024). The passenger projections do not account for the existence of the Jakarta–Surabaya semi high-speed train and solely rely on historical data. Drawing from the findings of Lubis et al. (2019), we adopt an optimistic scenario which predicts that approximately 23% of airplane passengers will shift to high-speed trains by 2025. Given that the Jakarta–Surabaya route still dominates train passenger distribution in Indonesia with exceptionally high demand for both rail and air transport, we assume a linear relationship between passengers travelling on the Jakarta–Surabaya semi high-speed train and the total number of train passengers. Hence, the overall number of passengers using the Jakarta–Surabaya semi high-speed train would be equivalent to the total passengers from trains (as the executive service will be replaced with HSR) plus 23% of airplane passengers. Consequently, this would result in an estimated 54% increase compared to 2017.

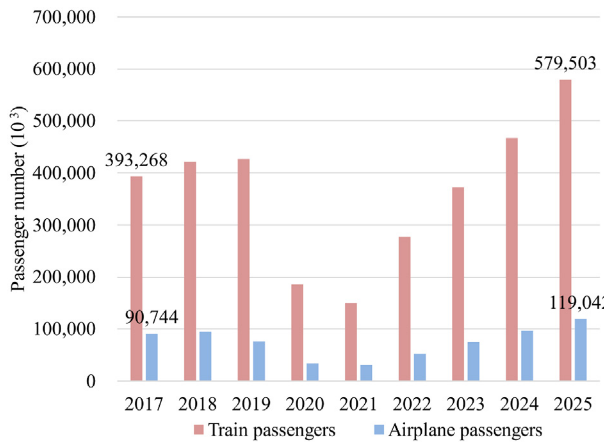


Fig. B1. Number of passengers for rail transport and air transport in Indonesia (2017–2025) (Analysed by authors by data from BPS-Statistics Indonesia (2024)).

Given the absence of data for future Jakarta–Surabaya semi high-speed train passengers, we rely on real *OD* demand information from the conventional train, the Jakarta–Surabaya executive train in 2017 (Table 7), which is planned to be replaced by the Jakarta–Surabaya semi high-speed train. Factoring in the earlier prediction, we opt for a 50% increase from the *OD* demand data in 2017, as reflected in Table 8.

Appendix C

Input parameters for example illustrations

Table C1
Parameters used in the example illustrations for cases with and without stopping patterns.

Parameter	Value
Set of stations, N	{1,2,3,4}
Set of links, Λ	{(1,2), (2,3), (3,4)}
Start station, s & end station, t	{1} & {4}
Set of *OD* pairs, S	{(1,2), (1,3), (2,3), (1,4), (2,4), (3,4)}
Set of line indices, L^1	{1,2,3,4}
Minimum & maximum number of trains which can run on the link, $M_{ij}\&T_{ij}, \forall(i,j) \in A$	12 & 15 trains/segment
Case with predetermined stopping patterns	
Set of nonstop arcs:	
A_1	{(1,2), (2,3), (3,4)}
A_2	{(1,2), (2,4)}
A_3	{(1,3), (3,4)}
A_4	{(1,4)}
Set of possible *OD* pairs (passengers) who can travel in nonstop arc $(i,j) \in A_1$	
$S_{(1,2)}$	{(1,2), (1,3), (1,4)}
$S_{(2,3)}$	{(1,3), (1,4), (2,3), (2,4)}
$S_{(3,4)}$	{(1,4), (2,4), (3,4)}
Set of possible *OD* pairs (passengers) who can travel in nonstop arc $(i,j) \in A_2$	
$S_{(1,2)}$	{(1,2), (1,4)}
$S_{(2,4)}$	{(2,4), (1,4)}
Set of possible *OD* pairs (passengers) who can travel in nonstop arc $(i,j) \in A_3$	
$S_{(1,3)}$	{(1,3), (1,4)}
$S_{(3,4)}$	{(3,4), (1,4)}
Set of possible *OD* pairs (passengers) who can travel in nonstop arc $(i,j) \in A_4$	
$S_{(1,4)}$	{1,4}
Set of line indices for $OD \in S$	
$L^{(1,2)}$	{1,2}
$L^{(1,3)}$	{1,3}
$L^{(2,3)}$	{1}

(*continued on next column*)

Table C1 (*continued*)

Parameter	Value
$L^{(1,4)}$	{1,2,3,4}
$L^{(2,4)}$	{1,2}
$L^{(3,4)}$	{1,3}
Case without predetermined stopping patterns	
Set of nonstop arcs: A	{(1,2), (1,3), (2,3), (1,4), (2,4), (3,4)}
Set of possible *OD* pairs (passengers) who can travel in nonstop arc $(i,j) \in A$	
$S_{(1,2)}$	{(1,2), (1,3), (1,4)}
$S_{(1,3)}$	{(1,3), (1,4)}
$S_{(2,3)}$	{(1,3), (1,4), (2,3), (2,4)}
$S_{(1,4)}$	{(1,4)}
$S_{(2,4)}$	{(2,4), (1,4)}
$S_{(3,4)}$	{(1,4), (2,4), (3,4)}

References

Assad, A.A., 1980. Modelling of rail networks: toward a routing/makeup model. Transp. Res. Part B Methodol. 14, 101–114.

Borndörfer, R., Grötschel, M., Pfetsch, M.E., 2007. A column-generation approach to line planning in public transport. Transport. Sci. 41, 123–132.

Borndörfer, R., Klug, T., Lamorgese, L., Mannino, C., Reuther, M., Schlechte, T., 2017. Recent success stories on integrated optimization of railway systems. Transport. Res. C Emerg. Technol. 74, 196–211.

BPS-Statistics Indonesia, 2024. Jumlah penumpang kereta api (Ribu Orang), 2024. https://www.bps.go.id/id/statistics-table/2/NzIjMg==/jumlah-penumpang-kereta-api.html.

Bracaglia, V., D'Alfonso, T., Nastasi, A., Sheng, D., Wan, Y., Zhang, A., 2020. High-speed rail networks, capacity investments and social welfare. Transport. Res. Part A Policy Pract. 132, 308–323.

Bussieck, M., 1998. Optimal Lines in Public Rail Transport. Technische Universität Braunschweig, Braunschweig, Germany. Ph.D. Dissertation.

Bussieck, M.R., Winter, T., Zimmermann, U.T., 1997a. Discrete optimization in public rail transport. Math. Program. 79, 415–444.

Bussieck, M.R., Kreuzer, P., Zimmermann, U.T., 1997b. Optimal lines for railway systems. Eur. J. Oper. Res. 96, 54–63.

Canca, D., De-Los-Santos, A., Laporte, G., Mesa, J.A., 2016. A general rapid network design, line planning and fleet investment integrated model. Ann. Oper. Res. 246, 127–144.

Chang, Y.H., Yeh, C.H., Shen, C.C., 2000. A multiobjective model for passenger train services planning: application to Taiwan's high–speed rail line. Transp. Res. Part B Methodol. 34, 91–106.

Claessens, M.T., 1994. A Mathematical Programming Model to Determine a Set of Operation Lines at Minimal Costs. WIT Press, Southampton, UK.

Claessens, M.T., van Dijk, N.M., Zwaneveld, P.J., 1998. Cost optimal allocation of rail passenger lines. Eur. J. Oper. Res. 110, 474–489.

Eliasson, J., 2021. Efficient transport pricing–why, what, and when? Commun. Transport. Res. 1, 100006.

Fu, H., Nie, L., Meng, L., Sperry, B.R., He, Z., 2015. A hierarchical line planning approach for a large-scale high speed rail network: the China case. Transport. Res. Part A Policy Pract. 75, 61–83.

Ghoseiri, K., Szidarovszky, F., Asgharpour, M.J., 2004. A multi-objective train scheduling model and solution. Transp. Res. Part B Methodol. 38, 927–952.

Goossens, J.W., van Hoesel, S., Kroon, L., 2004. A branch-and-cut approach for solving railway line–planning problems. Transport. Sci. 38, 379–393.

Goossens, J.W., van Hoesel, S., Kroon, L., 2006. On solving multi-type railway line planning problems. Eur. J. Oper. Res. 168, 403–424.

Huang, W., Shuai, B., 2019. Approach and application on high-speed train stop plan for better passenger transfer efficiency: the China case. Int. J. Rail. Transp. 7, 55–78.

Li, Z.C., Lam, W.H.K., Wong, S.C., Sumalee, A., 2011. Design of a rail transit line for profit maximization in a linear transportation corridor. Procedia Soc. Behav. Sci. 17, 82–112.

Liu, D., Durán Micco, J., Lu, G., Peng, Q., Ning, J., Vansteenwegen, P., 2020. A matheuristic iterative approach for profit–oriented line planning applied to the Chinese high-speed railway network. J. Adv. Transport. 2020, 4294195.

Lubis, H.A.R., Pantas, V.B., Farda, M., 2019. Demand forecast of jakarta–surabaya high speed rail based on stated preference method. IJTech 10, 405.

Lusby, R.M., Larsen, J., Ehrgott, M., Ryan, D., 2011. Railway track allocation: models and methods. Spectrum 33, 843–883.

Official Journal of the European Union, 2011. The technical specification for interoperability relating to the subsystem 'telematics applications for passenger services' of the trans-European rail system Text with EEA relevance. http://data.europa.eu/eli/reg/2011/454/oj.

Official Journal of the European Union, 2014. The technical specification for interoperability (TSI) relating to the 'telematics applications for freight' subsystem of the rail system in the European Union. http://data.europa.eu/eli/reg/2014/1305/oj.

Park, B.H., Seo, Y.I., Hong, S.P., Rho, H.L., 2013. Column generation approach to line planning with various halting patterns—application to the Korean high-speed railway. Asia Pac. J. Oper. Res. 30, 1350006.

Pu, S., Zhan, S., 2021. Two-stage robust railway line-planning approach with passenger demand uncertainty. Transport. Res. Part E Logist. Transp. Rev. 152, 102372.

Schlechte, T., 2014. Railway track allocation. In: Operations Research Proceedings 2012, pp. 15–20.

Schöbel, A., 2012. Line planning in public transportation: models and methods. Spectrum 34, 491–510.

Schöbel, A., Scholl, S., 2006. Line planning with minimal traveling time. 5th Workshop on Algorithmic Methods and Models for Optimization of Railways (ATMOS'05) 2, 1–16.

Scholl, S., 2006. Customer-oriented Line Planning. Technischen Universitat Kaiserslautern, Kaiserslautern, Germany. Ph.D. Dissertation.

Song, Y.J., Shoji, K., 2016. Effects of diversification strategies on investment in railway business: the case of private railway companies in Japan. Res. Transport. Econ. 59, 388–396.

Torres, L.M., Torres, R., Borndorfer, R., Pfetsch, M.E., 2008. Line planning on paths and tree networks with applications to the quito trolebus system. In: 8th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS'08), vol. 9, pp. 1–13.

Tsai, J.F., Lin, M.H., Hu, Y.C., 2008. Finding multiple solutions to general integer linear programs. Eur. J. Oper. Res. 184, 802–809.

Utomo, D.P., Lestari, P.D.R., Sucipto, S., Nurhidayat, A.Y., 2020. Impact of the Jakarta-Surabaya high-speed train on passengers of the Jakarta–Surabaya executive train corridor. Majalah Ilmiah Pengkajian Industri. 14, 1–10.

Wang, L., Jia, L.M., Qin, Y., Xu, J., Mo, W.T., 2011. A two-layer optimization model for high-speed railway line planning. J. Zhejiang Univ. Sci. A. 12, 902–912.

Wang, Q., Deng, L., Chen, Y., Zhou, W., Niu, F., 2020. Optimizing a high-speed railway operation plan based on train capacity and service frequency. J. Transport. Eng. Part A. Syst. 146, 04020109.

Wang, Y., Li, X., 2022. Optimality analysis of train platforming and routing with different interlocking modes. Commun. Transport. Res. 2, 100061.

Yan, F., Goverde, R.M.P., 2019. Combined line planning and train timetabling for strongly heterogeneous railway lines with direct connections. Transp. Res. Part B Methodol. 127, 20–46.

Yang, L., Qi, J., Li, S., Gao, Y., 2016. Collaborative optimization for train scheduling and train stop planning on high-speed railways. Omega 64, 57–76.

Yue, Y., Wang, S., Zhou, L., Tong, L., Saat, M.R., 2016. Optimizing train stopping patterns and schedules for high-speed passenger rail corridors. Transport. Res. C Emerg. Technol. 63, 126–146.

Zhang, W., Zhao, H., Xu, M., 2021. Optimal operating strategy of short turning lines for the battery electric bus system. Commun. Transport. Res. 1, 100023.

Zhao, S., Wu, R., Shi, F., 2021. A line planning approach for high-speed railway network with time-varying demand. Comput. Ind. Eng. 160, 107547.

Zhou, Y., Yang, H., Wang, Y., Yan, X., 2021. Integrated line configuration and frequency determination with passenger path assignment in urban rail transit networks. Transp. Res. Part B Methodol. 145, 134–151.

Prasetyaning Diah Rizky Lestari is a Ph.D. student at the Institute for Transport Studies (ITS), University of Leeds, UK. She is also a researcher at the National Research and Innovation Agency of Indonesia (BRIN). She received her B.Sc. degree in Applied Mathematics from IPB University, Indonesia, and her M.Sc. degree in Computational Science from Bandung Institute of Technology, Indonesia. Her research interests include mathematical and computational modelling to develop and evaluate innovative solutions for various transport and energy problems. She is particularly interested in simulation, optimisation, and modelling of traffic and railway networks, energy systems, and renewable resources, delivering high-quality outputs and insights for policy and decision-making.

Ronghui Liu is a Professor of Networks and Transport Operations at the Institute for Transport Studies (ITS), University of Leeds, UK. She received her B.Sc. degree from Peking University and her Ph.D. degree from Cambridge University. Her research areas span several themes in the field of transport studies: vehicle dynamics and traffic microsimulation model development, traffic control theory and algorithms, travel behaviour and intelligent mobility, stochastic models and reliability analysis, public transport operations and controls, timetabling and schedule coordination, and train control and railway traffic management systems. She is particularly interested in the interrelationships between these themes and methodologies for the estimation, design, and evaluation of their network-wide effects.

Richard Batley is a Professor of Transport Demand and Valuation at the Institute for Transport Studies (ITS), University of Leeds, UK. He received his B.Sc. degree from the Newcastle University, and both his M.A. and Ph.D. degrees from the University of Leeds. With a disciplinary background in transport economics, Richard's specialist expertise covers three related areas: first, non-market valuation of travel; second, forecasting the impacts of changes in fares, journey time and journey quality on demand and revenue; and third, the theory and practice of economic appraisal. He has mainly operated at the interface between academia and public policy, and can demonstrate lasting impacts from his research, especially in the form of official UK policy and practitioner guidance issued to transport operators and transport scheme promoters.