
Journal of Industrial and Production Engineering

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tjci21

# Optimization of train formation plan based on technical station under railcar demand fluctuation

Bing Li, Shangtao Jiang, Yanjie Zhou & Hua Xuan

JOURNAL OF INDUSTRIAL AND PRODUCTION ENGINEERING
https://doi.org/10.1080/21681015.2023.2221699



# Optimization of train formation plan based on technical station under railcar demand fluctuation

Bing Li, Shangtao Jiang, Yanjie Zhou and Hua Xuan

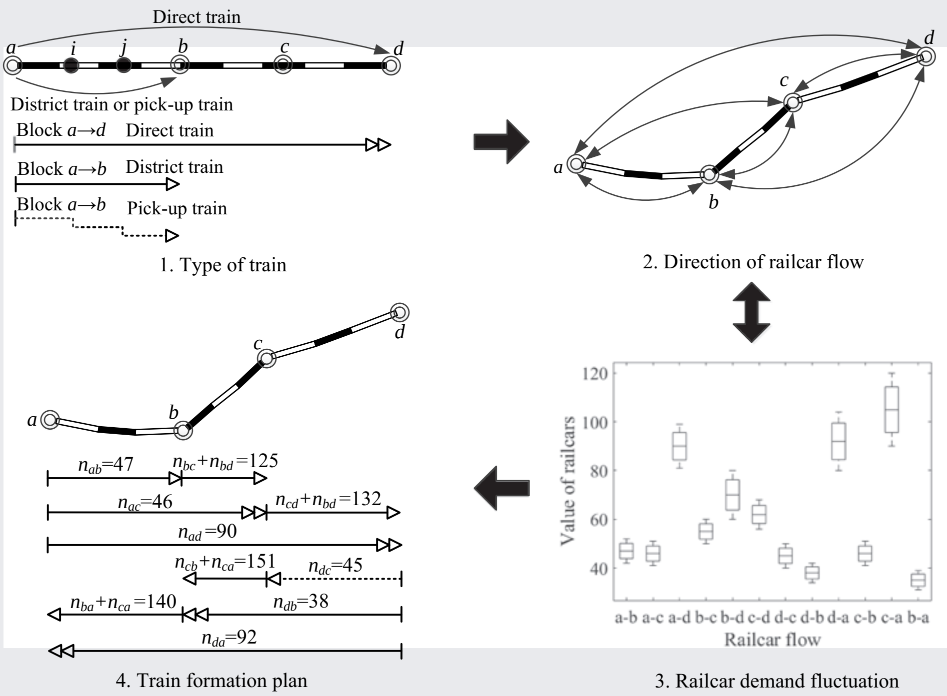
School of Management, Zhengzhou University, Zhengzhou, China

**ABSTRACT**

Train formation planning (TFP) is essential for rail freight logistics services. The fluctuation of railcar flows dramatically compared with before the outbreak of COVID-19. This paper studies train formation planning, considering three types of train services provided for railcar flow between pairs of technical stations (TS), including direct trains, district trains, and pickup trains. This paper introduces an optimization model with average railcars flow data (OMAD) and an optimization model with dynamic railcars flow data (OMDD) for the train formation planning based on TS under railcar demand fluctuation while minimizing railcar-hour consumption. The OMAD is a deterministic model, and the OMDD is a probability constraint model. To solve the OMDD, an approach for transforming probability constraints into deterministic constraints is presented. Various groups of scenarios are given to verify the effectiveness of the proposed models.

**ARTICLE HISTORY**
Received 26 December 2022
Revised 26 May 2023
Accepted 29 May 2023

**KEYWORDS**
technical station; formation plan; direct train; district train; pick-up train; railcar flow classification




## 1. Introduction

In recent years, the outbreak of COVID-19 has caused many enterprises to switch between production shutdown and resumption. This results in sharp fluctuations in demand for raw materials and finished products. Railway transport mainly delivers bulk commodities and undertakes main-line transport tasks. So the cargo of railcars is directly affected. To minimize the influence of frequently working out railcar flow organization schemes on railway transport management, the method fitting railcar demand fluctuation needs to be developed.

The railcar flow organization is the basis of railway transport management. The key to railcar flow organization is working out the train formation plan, which consolidates the railcars into the most reasonable freight train service sequence. Generally, the trains providing freight services between pairs of classification stations mainly include direct trains, district trains, and pickup trains. The classification station is the specific railroad station where the expensive marshaling

**CONTACT** Yanjie Zhou ieyjzhou@zzu.edu.cn School of Management, Zhengzhou University, Zhengzhou 450001, China

infrastructure is composed of the arrival yard, the classification bowl with a sophisticated braking system, and the departure yard is equipped. The railcar classification and train formation work can be operated in the classification station. So the classification station is also called the technical station in the China railway system. To maintain consistency of the expression, we will uniformly term technical station in this paper. The long-distance direct train runs between a pair of non-adjacent technical stations. But, the district train and pickup train must be operated between pair of adjacent technical stations.

This study aims to achieve the train formation plan based on technical station (TFP&TS) to minimize the total railcar-hour consumption induced by consolidating railcars to form trains in the origin technical station and classifying trains at the middle technical station. Its core task is making block-train assignment decisions, including pairs of technical stations providing train service (block plan) and the train type providing for the block.

The remainder of this study is organized as follows. The previous studies are summarized in section 2. Section 3 introduces a criterion for deciding the service type of train and sets up the optimization model with average data and the model with dynamic data. Methodologies for transforming the probabilistic constraints into deterministic expressions are introduced in Section 4. Section 5 presents a case study. Finally, the conclusions are given in Section 6.

## 2. Literature review

The general train formation plan (also called the blocking plan in the North American railway system) consists of the blocking problem, the train routing problem and the train makeup problem.

The objective of the blocking problem is to determine the blocking policy at each technical station and the shipment-to-block assignment. There is a lot of literature about this. Bodin et al. [1] formulated the blocking problem using an arc-based mixed integer programming model. Its some capacity constraints indicate each technical station in terms of the maximum number of blocks and the maximum number of railcar. Chen et al. [2] investigated the single-block train formation problem in railway freight transportation, the aim is to service all demands with minimal costs, and a novel solution methodology was proposed. Newton et al. [3] transformed the problem of formation plan optimization into the problem of network design, took a formation direction as an arc and a technical station as a node to form a formation direction network, and constructed a model with minimum cost as the goal. Xiao et al. [4] presented a solution for the freight train formation plan problem in China using both the single-block trains and the two-block trains. The solution is tested on an existing railway network in China. Crevier et al. [5] proposed an optimization model which encompasses pricing decisions and network planning policies for the railcar blocking. Xiao and Lin [6] proposed a mathematical programming model to solve the comprehensive problem of the one-block train and two-block train. The model aimed at minimizing the railcar-hour consumption of transporting the commodities. A heuristic optimization approach based on the ant colony system was proposed to solve the model.

Many researchers use heuristic algorithms to investigate the problem of train formation plan optimization from different problem perspectives. The reason is that manual approaches are highly time-consuming and highly rely on the experience of the designer. Fast and high-quality solutions can be provided by modern optimization models and heuristic algorithms based on computers. Keaton [7] proposed a mixed integer programming model for the combined problem, and a heuristic approach based on Lagrangian relaxation was presented to solve the model. Keaton [8] proposed a model for the problem by using the service network design method, and the Lagrangian relaxation was implemented by the dual-adjustment method. Lin et al. [9] proposed a linear integer programming model to solve the train service network problem of the China railway system. A simulated annealing algorithm is applied to the real-life Chinese railway network, and a significant improvement in the total railcar-hour consumption is achieved. Yaghini et al. [10] presented a meta-heuristic algorithm based on ant colony optimization to solve the railway blocking problem, and it aimed at minimizing the railcar-hour consumption of transporting goods. There are other researchers proposing heuristic optimization algorithms to solve the problem, including genetic algorithm [11,12], Lagrangian relaxation heuristic algorithm [13], and large-scale neighborhood search algorithm [14].

The train routing problem concerns the operating policies for the freight transportation and the railcar fleet management. In America and Europe, the routing problem is an important part of the train formation plan. The shipment-to-block and block-to-train problems are considered in a daily or weekly periodic cycle. After working out the train formation plan, the operating route of the train is also determined. However, in China, the routing problem is not a part of the train formation plan. The dynamic change of the traffic flow is not considered in solving the train formation plan problems. The routing problem is typically made as a separate part together with the train formation plan. Fuegenschuh et al. [15] proposed a mixed integer programming model for the freight railcar routing problem of Deutsche Bahn. It aims at finding the railcar routes with the lowest railcar-hour consumption, the train travel distance, and the amount of used sorting tracks were considered. Alikhani-Kooshkak et al. [16] proposed a multi-objective model to solve the optimization problem of train routing and train formation plan. It aimed at optimizing the total profit, satisfaction level of customers, and underutilized train capacity. Sama et al. [17] formulated the real-time train routing selection problem as an integer linear programming model and proposed a meta-heuristic algorithm based on ant colony optimization to solve the problem. Fang et al. [18] developed a mixed-integer programming model and a heuristic solution strategy. The scheduling and routing of rail shipments of hazardous materials have been investigated. Lin [19] constructed a comprehensive optimization model of formation plan and route based on direct trains and district trains with the method of combinatorial optimization. Zhu et al. [20] proposed a model for the network design problem for freight rail transportation. Car classification, train formation, service selection, and scheduling were considered, and time-dependent customer delivery routes were built. Lin et al. [21] proposed the 0–1 programming model for the comprehensive optimization of the train formation plan and the train route, and the large-scale case was solved.

There is some literature in regard to railway transport with demand fluctuation, but not for freight trains, they are for passenger trains. For example, Cao and Feng [22] considered passenger demand uncertainty and constructed a two-stage stochastic integer programming model to get the capacity allocation solutions in order to meet random fluctuations of passenger demands in the daily operations. There is also some literature on the problem of demand fluctuations in highway transport. For example, Tarhini and Bish [23] investigated network routing and traffic controls of highway transport under demand uncertainty. The literature on these issues is extensive and will not be covered here. The research contents and methods of the main references in this paper are listed in Table 1. The differences between the research contents of this paper and them can be clearly seen in Table 1.

It can be known from Table 1 and the actual situation, at present, there is little or no research on the fluctuation of railway freight transport. In addition, many studies of train formation plan assume that district trains must be operated between two adjacent technical stations. The assumption is likely to cause lower practicability of train formation plan in a realistic railway transportation work.

In this paper, we will study the train formation plan based on the technical station under railcar demand fluctuation. The contributions of this paper are as follows: first, we systematically analyze and summarize the train service types between pairs of technical stations. Second, we study the organizational conditions of pickup trains and district trains without assuming that district trains must be operated between two adjacent stations. The critical railcar flow condition deciding whether the district train between two adjacent stations is operated or not is designed, and then the criterion of deciding train service type is obtained. Third, we establish two optimization models based on average data and dynamic data aiming at the minimum railcar-hour consumption at all the stations in the railway network. Fourth, we develop the approach for transforming probability constraints into deterministic constraints. Finally, for different railcar flow data fluctuation ranges, we test and solve the proposed model on a railway network composed of eight technical stations. The result shows that the formation plan obtained by solving the optimization model based on dynamic data is highly adaptive within the range of railcar flow fluctuation. It avoids frequently remaking the train formation plan and brings great convenience to railway transportation management.

## 3. Problem statement

In this section, we first analyze the core task of the train formation plan based on the technical station. And then, we introduce the notations used in this paper and propose two detailed mathematical model formulations for the TFP&TS problem.

**Table 1.** Comparison of previous studies and this paper on train formation plan of technical station.

| | Type of train | | | Model | | Block train | | |
|---|---|---|---|---|---|---|---|---|
| Studies | Direct train | District train | Pick-up train | Linear model | Non-linear model | One-block train | Two-block train | Demand fluctuation |
| Barnhart et al [13] | ✓ | ✓ | × | × | ✓ | ✓ | × | ✓ |
| Crevier et al. [5] | ✓ | ✓ | × | × | ✓ | ✓ | × | ✓ |
| Tarhini and Bish (2016) [23] | × | × | × | × | ✓ | × | × | × |
| Fugenschuh et al. [15] | ✓ | ✓ | × | ✓ | × | ✓ | × | ✓ |
| Xiao and Lin [6] | ✓ | ✓ | × | × | ✓ | ✓ | ✓ | ✓ |
| Chen et al. (2019)[2] | ✓ | ✓ | × | ✓ | × | ✓ | × | ✓ |
| [4] | ✓ | ✓ | × | × | ✓ | ✓ | ✓ | ✓ |
| Alikhani-Kooshkak et al. [16] | ✓ | ✓ | × | ✓ | × | ✓ | × | ✓ |
| Cao and Feng [22] | ✓ | × | × | × | ✓ | × | × | ✓ |
| Lin et al. [21] | ✓ | ✓ | ✓ | × | ✓ | ✓ | × | ✓ |
| This paper | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | × | ✓ |

### 3.1. Analysis of train formation plan

The railcar flow between pairs of technical stations is usually carried by three types of trains, i.e. direct train, district train, and pickup train. When the railway section between pairs of technical stations provides the train service, the railway section is named as one block. The main task of the train formation plan studied in this paper is to make block-train assignment decisions, i.e. aims at specifying the following elements of the problem:

(i) Blocking plan: which pairs of technical stations should provide train service?
(ii) Train type: which train should the block be carried?

The direct train is the primary mode among the three category trains. The remote railcar flows, which have the same destination, should be consolidated into the direct train. They run along railway sections between two nonadjacent technical stations. There is at least one technical station between the origin and destination on its itinerary. The direct train does not require it to be reclassified at any middle technical station that it passes through on its itinerary. As a result, the block between pairs of nonadjacent technical stations is carried by a direct train.

If the amount of railcar flow with adjacent origin-destination satisfies the sufficient condition, they will be consolidated into the district trains. They run along railway sections between two adjacent technical stations. There are many intermediate stations between two adjacent technical stations. Unlike technical stations, the intermediate stations are mainly used to pick and deliver railcars from the loading-unloading points. However, the district train does not require to perform railcar handling at any intermediate station that it passes through on its itinerary. As a result, the block between pairs of adjacent technical stations is carried by a district train.

When the amount of railcar flow with adjacent origin-destination does not satisfy the sufficient condition, they will be shipped by pickup trains. The pickup train will perform railcar handling at an intermediate station located on its itinerary. The operations of railcar handling include detaching railcars from the pickup train and putting them into the loading-unloading points within the intermediate station, retrieving railcars from the loading-unloading points within an intermediate station and attaching them to the pickup train, and so on. So, the block between pairs of adjacent technical stations is carried by a pickup train.

The train service can be illustrated in Figure 1.

In this study, three of the most practical and crucial train formation patterns based on the technical station are elaborated on and analyzed. The optimization of the TFP&TS problem aims at determining the optimal train services between pairs of origin-destination technical stations. The objective function intends to minimize the total railcar-hour consumption induced by consolidating railcars to form a train in the origin technical station, classifying trains at the middle technical station, and additional operation of pickup trains. Some practical conditions and logical constraints among decision variables are considered in the model.

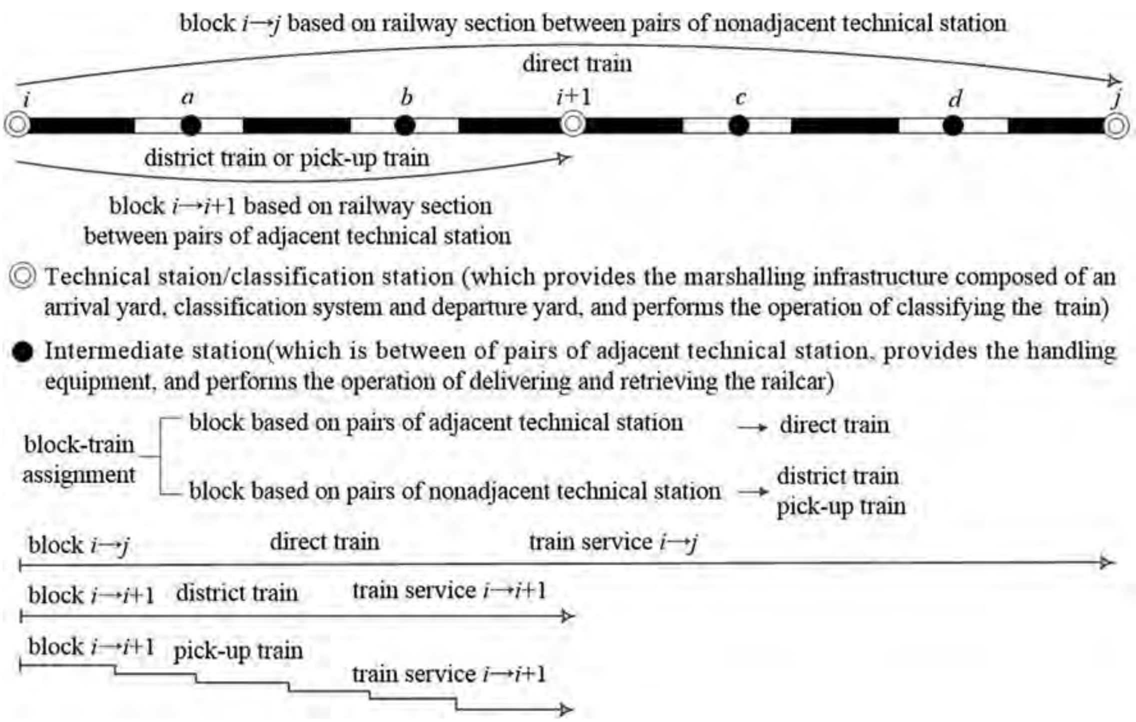



Figure 1. The train service in rail.

Train formation of loading stations (TFLS) is essential for rail freight logistics services for bulk goods. The main mode of it is the through train that originates from a loading point to an unloading station. After loading station organizes this train, the remaining railcar flows are sent to the adjacent technical station, they are transported by trains organized by the technical station. The optimization of the TFLS problem aims at determining the optimal train services among loading and unloading stations for bulk materials minimizing the total car-hour consumption induced by the loading, unloading, and reclassification operations and satisfying practical conditions and logical constraints among decision variables.

### 3.2. Problem definition and notation

Before introducing the mathematical formulations of the studied problem, the notations and the decision variables used in this paper are listed in Tables 2 and 3.

### 3.3. Criterion of deciding train service type

Deciding optimal train service type between pairs of the origin-destination technical station is the core of solving the problem. Remote railcar flows with the same origin and destination should be consolidated into the direct train. Short-distance railcar flows between adjacent origin-destinations can be optionally assembled to form a district train or pickup train. Compared with the operating conditions of direct trains, the selection of district and pickup trains is more complex and challenging. This section focuses on the train service types between pairs of adjacent technical stations.

(1) Analysis of railcar flows between pairs of adjacent technical stations

There are two adjacent technical stations $i$ and $j$, and two intermediate stations $a$ and $b$ in the railway section in Figure 2

The network includes one solid line with an arrow, representing the district railcar flow from technical station $i$ to $j$. Its weight $p_{ij}$ represents the value of district railcar flow. The network includes two dot lines with arrows, representing the pickup railcar flow delivered to intermediate stations $a$ and $b$. Their weights $f_{ia}$ and $f_{ib}$ represent the value of pickup railcar flow. The network includes two dashed lines with an arrow, meaning the pickup railcar flow retrieved from intermediate stations $a$ and $b$. Their weights $f_{aj}$ and $f_{bj}$ represent the value of pickup railcar flow.

(2) Analysis of independent operation mode of pickup train

If we independently operate the pickup trains between adjacent technical stations $i$ and $j$, the railcar flow consolidated by the pickup train consists of the following three parts.

**Table 2.** Notations.

| Set | |
|---|---|
| $M$ | Set of technical stations and $\widetilde{m}$ is the cardinality of $M$ |
| $M(i,j)$ | Set of technical stations located at the physical paths from station $i$ to $j$ excluding technical station $i$ and $j$ |
| $M(i)$ | Set of technical stations adjacent to station $i$ |
| $M^{'}(i)$ | Set of technical stations nonadjacent to station i |
| $H$ | Set of statistical time periods through planning horizon and $\widetilde{h}$ is the cardinality of $H$ |
| **Index** | |
| $i$ | Technical station |
| $j$ | Technical station |
| $k$ | Technical station |
| $h$ | Statistical time period |
| **Parameters** | |
| $a_i$ | Maximum number of railcars that can be handled by technical station $i$ |
| $b_i$ | The utilization rate of classification capacity of technical station $i$, which means the proportion of the remaining classification capacity after deducting the reserved capacity at technical station $i$ |
| $c_i$ | Assembling parameter at technical station $i$, which reflects the random arrival of railcars to form a train at technical station $i$ |
| $d_i$ | Number of available classification tracks at technical station $i$ |
| $e_i$ | Number of railcars that can be accommodated in each classification track at technical station $i$ |
| $g$ | Size of a train, which means the number of railcars forming the train $i$ |
| $t_i$ | Relative delay at technical station $i$. It consists of delays in the arriving, inspecting, classification, assembling, and departing processes for a railcar at technical station $i$ |
| $t_{ij}^{d}$ | Travel time of the district trains from technical station $i$ to technical station $j$ |
| $t_{ij}^{l}$ | Travel time of the pick-up trains from technical station $i$ to technical station $j$ |
| $\theta_{ij}$ | Critical value of railcar flow used to decide whether district trains can be operated between pairs of adjacent technical stations $i$ and $j$ or not |
| $q_{ij}$ | Addition railcar-hour consumption when the railcars are carried by pick-up train from technical station $i$ to $j$ |
| **Random variables** | |
| $n_{ij}^{h}$ | Original O-D railcar flow volume from $i$ to $j$, i.e. number of railcar flow that originates at technical station $i$ and destines to technical station $j$ in the statistical time period $h$ |
| $f_{ij}^{h}$ | Actual railcar flow volume from $i$ to $j$, i.e. number of railcar flow that originates or reclassifies at technical station $i$ and destines to technical $j$ in the statistical time period $h$ |
| $p_{ij}^{h}$ | Service railcar flow volume from $i$ to $j$, i.e. number of railcar flow that originates or reclassifies at technical station $i$, reclassifies or destines at technical station $j$ in the statistical time period $h$ |

**Table 3.** Decision variables.

| Decision variables | |
|---|---|
| $x_{ij}^k$ | $x_{ij}^k \in \{0, 1\}$. Railcar flow reclassification variable. It is a binary variable. Its value is 1 if the railcar flow originating technical station *i* and destining technical station *j* is reclassified at middle technical station *k*. Otherwise, it is 0 |
| $y_{ij}$ | $y_{ij} \in \{0, 1\}$. Train service variable. It is a binary variable. Its value is 1 if train service types (including direct train and district train) is provided to block $i \to j$ |
| $z_{ij}$ | $z_{ij} \in \{0, 1\}$. Pick-up train variable. It is a binary variable. Its value is 1 if any pick-up train is provided from technical station *i* to adjacent station *j*. Otherwise, it is 0 |

The first one is the pickup railcar flow delivered to intermediate stations between pairs of the adjacent technical stations, whose value is written by $f_{ia}$ and $f_{ib}$.

The second one is the pickup railcar flow retrieved from intermediate stations between pairs of the adjacent technical stations, whose value is written by $f_{aj}$ and $f_{bj}$.

The third one is the district railcar flow that originates at the rear stations of the original technical station and travels to the front stations' destination technical station, whose value is written by $p_{ij}$. Here, the station to which the train is heading while running is the front station. The station opposite the train heading direction is the rear station.

Let $f_{ij}^T$ denote the total pickup railcar flow value. It is composed of the first one and the second one. We can calculate its value using the following equation:

$$f_{ij}^T = f_{ia} + f_{ib} + f_{aj} + f_{bj} \tag{1}$$

Because the district railcar flow and pickup railcar flow are carried by pickup trains. Let $\sigma_1$ denote the total railcar-hour consumption induced by these two kinds of railcar flows in this railway section. So it can be expressed by the following formula:

$$\sigma_1 = \left(f_{ij}^T + p_{ij}\right)t_{ij}^l + \left(f_{ij}^T + p_{ij}\right) \times T_i\left(f_{ij}^T + p_{ij}\right) \tag{2}$$

where $t_{ij}^l$ represents the travel time of the pickup trains from technical station *i* to technical station *j*, and $T_i(\cdot)$ represents the waiting time of the pickup trains at the origin technical station *i*.

(3) Analysis of the hybrid operation mode of the district and pickup train

If we synchronously operate the pickup train and district train between adjacent technical stations *i* and *j*, the pickup railcar flow composed of the first one and the second one will be carried by the pickup train, and the remaining district railcar flow will be carried by the district train. Let $\sigma_2$ represent the total railcar-hour consumption under synchronous operation conditions of the district and pickup train. This can be expressed by the following formula:

$$\sigma_2 = c_i g + p_{ij} t_{ij}^d + f_{ij}^T t_{ij}^l + f_{ij}^T T_i\left(f_{ij}^T\right) \tag{3}$$

Where $c_i$ represents the assembling parameter at technical station *i*.

(4) Criterion of deciding train service types

If $\sigma_2 < \sigma_1$, we will operate the district train between adjacent stations *i* and *j*. So operation condition of the district train can be formulated as follows:

$$c_i g + p_{ij} t_{ij}^d + f_{ij}^T T_i\left(f_{ij}^T\right) < p_{ij} t_{ij}^l + \left(f_{ij}^T + p_{ij}\right) \times T_i\left(f_{ij}^T + p_{ij}\right) \tag{4}$$

When the volume of district railcar flow is low, it will not increase the number of pickup trains that both district and pickup railcar flows are carried by pickup train. That is to say, the waiting time of the original technical station *i* will stay the same when the

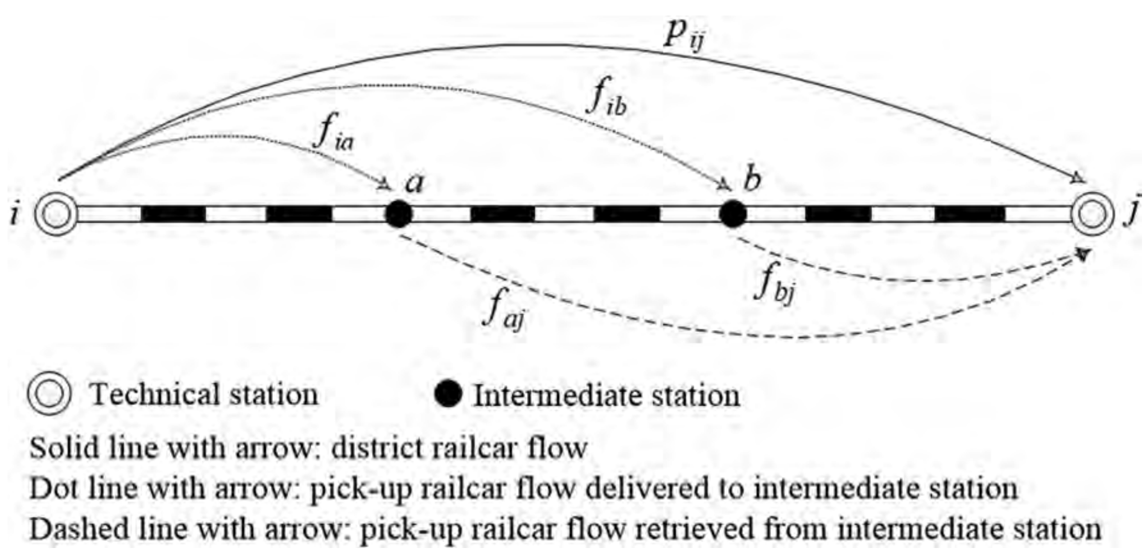



**Figure 2.** Railcar flows between pairs of adjacent origin-destination technical stations.

additional district railcar flow is also carried by the pickup train apart from the pickup railcar flow. So, we have:

$$T_i\left(f_{ij}^T + p_{ij}\right) \approx T_i\left(f_{ij}^T\right) \tag{5}$$

And then we put Equation (5) into Equation (4), and a simple equation can be obtained:

$$c_i g < p_{ij}\left(t_{ij}^l - t_{ij}^d\right) + p_{ij} \times T_i\left(f_{ij}^T\right) \tag{6}$$

We convert the formula into the following form:

$$p_{ij} > \frac{c_{ij} g}{t_{ij}^l - t_{ij}^d + T_i\left(f_{ij}^T\right)} \tag{7}$$

Equation (7) is the criterion for deciding the train service type between two adjacent stations. Suppose Equation (7) is tenable. In that case, we will additionally operate the district train apart from the pickup train, i.e. a hybrid operation mode of the district and pickup train. Otherwise, we perform an independent operation mode of the pickup train.

The right-hand side of Equation (7) means the lower bound of railcar flow volume used to decide whether to operate the district train additionally or not. We call it the critical railcar flow volume of district trains. In this way, we have:

$$\theta_{ij} = \frac{c_i g}{t_{ij}^l - t_{ij}^d + T_i\left(f_{ij}^T\right)} \tag{8}$$

From Equation (8), we can obviously conclude that the critical railcar flow volume of district train is higher under the shorter travel time $t_{ij}^l$ of the pickup trains.

### 3.4. Optimization model with average data

To reduce the difficulty of solving the model, we set up the model using the average daily railcar flow data in the range of the planning horizon. That is to say, the average value of railcar flow is used, i.e. $\bar{f}_{ij}$, $\bar{n}_{ij}$, and $\bar{p}_{ij}$. So the optimization model with average demand data (OMAD) for the TFP&TS problem is formulated as a mathematical programming model whose objective function and constraints are expressed as follows:

$$(\text{OMAD})\,Min \quad Z = \sum_{i\in M}\sum_{j\in M} c_i g y_{ij} + \sum_{i\in M}\sum_{j\in M'(i)}\sum_{k\in M(i,j)} \bar{f}_{ij} x_{ij}^k t_k + \sum_{i\in M}\sum_{j\in M(i)} z_{ij} q_{ij} \bar{p}_{ij} \tag{9}$$

$$\text{s.t.} \sum_{j\in M} \bar{p}_{kj} \le d_k e_k, k \in M \tag{10}$$

$$\sum_{i\in M}\sum_{j\in M'(i)} \bar{f}_{ij} x_{ij}^k \le a_k b_k, k \in M(i,j) \tag{11}$$

$$y_{ij} + \sum_{k\in M(i,j)} x_{ij}^k = 1, i \in M, j \in M'(i) \tag{12}$$

$$y_{ij} + z_{ij} = 1, i \in M, j \in M(i) \tag{13}$$

$$x_{ij}^k \le y_{ik}, i \in M, k, j \in M'(i) \tag{14}$$

$$z_{ij} = \begin{cases} 1 & \theta_{ij} - \bar{p}_{ij} > 0 \\ 0 & \theta_{ij} - \bar{p}_{ij} \le 0 \end{cases}, i \in M, j \in M(i) \tag{15}$$

$$y_{ij}, x_{ij}^k \in \{1, 0\}, i, j \in M, k \in M(i,j) \tag{16}$$

Equation (9) is the objective function, which minimizes the sum of the railcar collection time of the outgoing train at the technical departure station, train reclassification time at the middle technical station, and additional operation time for the pickup train.

The auxiliary variable $\bar{f}_{ij}$ is the average value of the actual railcar flow volume $f_{ij}^h$. It is composed of two parts. (i) The original railcar flow that originates at technical station $i$ and destines to technical station $j$. (ii) The reclassification railcar flow that originates at the rear technical station of technical station $i$ reclassifies at the technical station $i$ and is destined to technical station $j$. They are written by:

$$f_{ij}^h = n_{ij}^h + \sum_{m\in M} f_{mj}^h x_{mj}^i, i, j \in M \tag{17}$$

$$\bar{f}_{ij} = \bar{n}_{ij} + \sum_{m\in M} \bar{f}_{mj} x_{mj}^i, i, j \in M \tag{18}$$

The auxiliary variable $\bar{n}_{ij}$ is the average value of the original O-D railcar flow volume $n_{ij}^h$ in the statistical time period. It is written by:

$$\bar{n}_{ij} = \sum_{h\in H} n_{ij}^h / \tilde{h}, i, j \in M \tag{19}$$

The auxiliary variable $\bar{p}_{ij}$ is the average value of the service railcar flow $p_{ij}^h$ in the statistical period. It is composed of three parts: (i) The original railcar flow that originates at technical station $i$ and destines to technical station $j$; (ii) The reclassification railcar flow that originates at the rear technical station of technical station $i$ reclassifies at the technical station $i$ and destined to technical station $j$; (iii) The reclassification railcar flow that originates at the technical station $i$, reclassifies at the technical station $j$, and destines to the front technical station of technical station $j$. We can calculate by using the following equation: When the technical stations $i$ and $j$ are nonadjacent, we have:

$$p_{ij}^h = f_{ij}^h y_{ij} + \sum_{k\in M} f_{ik}^h x_{ik}^j, i \in M, j \in M'(i) \tag{20}$$

$$\bar{p}_{ij} = \bar{f}_{ij} y_{ij} + \sum_{k\in M} \bar{f}_{ik} x_{ik}^j, i \in M, j \in M'(i) \tag{21}$$

When the technical stations $i$ and $j$ are adjacent, we have:

$$p_{ij}^h = f_{ij}^h + \sum_{k\in M} f_{ik}^h x_{ik}^j, i, j \in M \tag{22}$$

$$\bar{p}_{ij}=\bar{f}_{ij}+\sum_{k\in M}\bar{f}_{ik}x_{ik}^{j},i,j\in M \tag{23}$$

The auxiliary variable $q_{ij}$ is the additional railcar-hour consumption when the railcars are carried by the pickup train from technical station $i$ to $j$. It consists of two parts. (i) The assembling time of outgoing pickup train made-up at departure station $i$. (ii) The additional traveling time due to the pickup train delivering and retrieving the railcar flow at the intermediate station when the railcar is carried by pickup train instead of district train between pairs of the adjacent technical station. We can calculate $q_{ij}$ by using the following equation:

$$q_{ij}=\left(T_i\left(\bar{p}_{ij}\right)+t_{ij}^{l}-t_{ij}^{d}\right)*\bar{p}_{ij},i\in M,j\in M(i) \tag{24}$$

Constraint (10) ensures that the number of occupied tracks is less than the number of available classification tracks. Constraint (11) provides that the number of classified railcars is lower than the available classification capacity of the technical station. Parameter $b_k$ means the proportion of the remaining classification capacity after deducting the reserved capacity at the technical station. Constraint (12) ensures that the railcar flow can either be directly delivered to the destination technical station or destined to the destination technical station after being classified at more than one technical station on its itinerary. Constraint (13) guarantees that when the average value of railcar flow between adjacent technical stations $i$ to $j$ is less than the critical railcar flow volume, the district trains cannot be operated. Constraint (14) ensures that the train can select $k$ as the first reclassification technical station only if the train service $i\to k$ (block $i\to k$) is provided. Constraint (15) ensures that the auxiliary variable $z_{ij}$ is the unit step function of the critical railcar flow volume of the district train. Finally, constraint (16) is the restriction on decision variables.

### 3.5. Optimization model with dynamic data

To improve the performance of the solution, we directly adopt the daily statistical railcars flow data to set up the optimization model with dynamic demand data (OMDD) for the TFP&TS. Compared with the optimization model with average data (OMAA) belonging to the deterministic model, the optimization model with dynamic data (OMDD) is a stochastic model. Here, we update the OMAD model by designing three probability constraints so that the new OMAD model can be developed.

(1) Probability constraint of train formation plan with high railcar-hour consumption

In order to avoid the solution of the problem falling into the local optimal trap and then improve the performance of the solution, the strategy of accepting inferior solutions within the maximum allowable range is given. And then, a new probability constraint which represents accepting the inferior solution under a certain probability is added to the model. The probability constraint of accepting the inferior solution is formulated as follows:

$$\text{Pro}\left\{\sum_{i\in M}\sum_{j\in M}c_i g y_{ij}+\sum_{i\in M}\sum_{j\in M'(i)}\sum_{k\in M(i,j)}\bar{f}_{ij}x_{ij}^{k}t_k+\sum_{i\in M}\sum_{j\in M(i)}z_{ij}q_{ij}p_{ij}^{h}\le G\right\}\ge\alpha,h\in H \tag{25}$$

Where parameter $\alpha$ represents the acceptance rate of the inferior solution and is within the range of [0.9, 1]. Here, the inferior solution denotes the train formation plan with high railcar-hour consumption. Next, let us define the proposal term "train formation plan with high railcar-hour consumption."

We put the average daily railcars flow data directly to the OMDD model. A train formation plan can be obtained by solving the OMDD model. Since the planning horizon includes the statistical time periods, we put dynamic railcar flows into the train formation plan, and the train formation plans, whose number is the same as the time periods, can be obtained. We rank the train formation plans in ascending order of railcar-hour consumption and obtain the train formation plan sequence. Define the train formation plans whose precedence is the position behind $\alpha$ in the sequence as train formation plans with high railcar-hour consumption. Thus let $G$ be the set of railcar-hour consumption of the train formation plan with high railcar-hour consumption. Constraint (25) implies that all train formation plans whose railcar-hour consumptions are set $G$ can all be accepted. The number of acceptable solutions is $(1-\alpha)\cdot\widetilde{h}$.

(2) Probability constraint of classification tracks number restriction

We allow the solution with not satisfying classification track number restriction to be accepted in a particular proportion. The constraint (10) representing classification tracks number restriction in the OMAD model is updated as the following probabilistic constraint form:

$$\text{Pro}\left\{\sum_{j\in M}p_{kj}^{h}\le d_k e_k\right\}\ge\beta,k\in M,h\in H \tag{26}$$

where the parameter $\beta$ is the satisfying rate of classification track number restriction, it means the proportion of the solution meeting classification track number restriction and is within the range of [0.9, 1].

Constraint (26) implies that the proportion of the train formation plan satisfying classification track number restriction is no less than $\beta$ in the $\widetilde{h}$ train formation plan obtained by solving the OMDD model. That is, the number of train formation plans satisfying classification tracks number restriction is no less than $\beta\cdot\widetilde{h}$.

(3) Probability constraint of classification capacity restriction

We allow the solution with not satisfy technical station classification capacity restriction to be accepted in a particular proportion. The constraint (11) representing technical station classification capacity restriction in the OMAD model is updated as the following probabilistic constraint form:

$$\text{Pro}\left\{\sum_{i\in M}\sum_{j\in M'(i)} f_{ij}^{h}x_{ij}^{k} \le a_k b_k\right\} \ge \gamma, k\in M(i,j), h\in H \tag{27}$$

where the parameter $\gamma$ is the satisfying technical station classification capacity restriction rate. It means that the proportion of the solution that meets technical station classification capacity restriction is within the range of [0.9, 1].

Constraint (27) implies that the proportion of train formation plan satisfying classification capacity restriction is no less than $\gamma$ in the $\widetilde{h}$ train formation plan obtained by solving the OMDD model. That is, the number of train formation plans satisfying classification capacity restriction is no less than $\gamma\cdot\tilde{h}$.

In conclusion, the optimization model with daily dynamic demand data (OMDD) for the TFP&TS problem is formulated as a mathematical programming model whose objective function and constraints are expressed as follows:

$$(\text{OMAD})Min \quad G \tag{28}$$

Subject to *Constraints (12) - (16), (25) - (27).*

Equation (28) means that the train formation plan with the lowest railcar-hour consumption in set $G$ is the optimal solution of the OMDD.

## 4. Methodology

In the OMDD model, the constraints (25), (26), and (27) are formulated as the probability expression equation. It is an unclear mathematical, analytical expression. If these probabilistic constraints are not treated, it will lead to difficulty in solving the model. Here, we expand the statistical railcar flow data and then transform the probabilistic constraints into deterministic expressions. Then, the model can be solved using conventional methods.

(1) Expanding statistical railcars flow data based on discrete uniform distribution

Considering that the statistical railcar flow data in the statistical period is insufficient, the following numerical simulation method is used to expand it.

Step 1: we find the minimum value $u_1$ and maximum value $u_2$ from railcar flow $n_{ij}^{h}$ between pairs of technical station $i$ to $j$ in $\widetilde{h}$ time period. Since the discrete uniform distribution has the excellent characteristics that every discrete value within the value range of a random variable has the same probability of occurrence, we use this distribution to describe the railcar flow random variable. We denote it as $n_{ij}^{h}\sim U(u_1,u_2)$.

Step 2: we generate $\widetilde{r}$ random numbers between [0,1] and denoted as $R=\{r|0\le r\le 1\}$. Let $\widetilde{r}\gg\widetilde{h}$.

Step 3: we use the equation $n_{ij}^{r}=ru_1+(1-r)u_2$ to generate $\widetilde{r}$ new railcar flows data $n_{ij}^{r}$. As a result, the expanding data set of fluctuant railcar flows can be obtained and denoted as $N^{R}$.

(2) Finding a train formation plan with the lowest high railcar-hour consumption

The acceptance rate of the inferior solution is used to find the train formation plan with the lowest railcar-hour consumption from set $G$ by the expanding railcar flow data set.

Step 1: because there is $\widetilde{r}$ expanding railcars flow data in set $N^{R}$. So the acceptance amount of the train formation plan with high railcar-hour consumption is $\psi=\lceil\widetilde{r}-\widetilde{r}\cdot\alpha$, which denotes minimum integer greater than $\tilde{r}-\widetilde{r}\cdot\alpha$.

Step 2: the $\widetilde{r}$ expanding railcars flow data in set $N^{R}$ replaces the average railcar flow data of the OMAD model in turn. The OMAD model based on expanding railcars flow data is solved $\widetilde{r}$ times, and $\widetilde{r}$ solutions can be obtained.

Step 3: the $\widetilde{r}$ solutions are sorted by the ascending order of their objective function value. Their objective functions are denoted as $\{Z_1',\cdots,Z_\theta',\cdots,Z_{\widetilde{r}}'\}$, where $\theta=\widetilde{r}-\psi+1$. Obviously, the objective function value set of train formation plan with high railcar-hour consumption is $G=\{Z_\theta',\cdots,Z_{\widetilde{r}}'\}$. The objective function value of train formation with the lowest railcar-hour consumption in set $G$ is $Z_\theta'$.

(3) Counting satisfying rate of railway network capacity of train formation plan

The following procedure is used to test the satisfying rate of railway network capacity of a train formation plan by the expanding railcar flow data.

Step 1: all train formation plan is placed into the constraints (10) and (11) of the OMAD model.

Step 2: furthermore, all $\widetilde{r}$ expanding railcar flow data in set $N^{R}$ is placed into the constraints (10) and (11) of the OMAD model for judging whether the two constraints are both satisfied. The amount of expanding railcars flow data satisfying the two constraints are recorded and denoted as $\widetilde{r}'$.

Step 3: the satisfying rate of railway network capacity of the train formation plan can be counted by the equation $\widetilde{r}'/\tilde{r}$.

## 5. A case study

In this section, we test the model and the solution approach in a railroad network composed of eight technical stations.

### 5.1. Scenario setting of the experiments

An experimental railway network composed of eight technical stations in China is shown in Figure 3. The pickup railcar flow between the adjacent stations is fixed relatively. When the district railcar flow is small, it will not increase the number of pickup trains that both district and pickup railcar flows are carried by pickup train. So the assembling time $T_i(\cdot)$ of the pickup train can be simplified as a constant. The basic parameters of the technical stations, including assembling parameters, assembling time, reclassification time, classification capacity, and track, are shown in Table 4. The technical stations are numbered from 1 to 8 in Table 1 for the convenience of subsequent expression. The value of the average daily railcars flow between pairs of O-D is shown in Table 5. The trip time between pairs of adjacent technical stations and critical railcar flow volume are shown in Table 6. The size of the train, which means the number of railcars forming the train, is 55. The number of railcars that can be accommodated in each classification track is 200.

### 5.2. Expanding statistical railcars flow data

Since the transportation of bulk goods (such as coal) in China is mainly undertaken by railway transportation, in China, Shanxi province has the largest coal reserve and annual output. The coal output before and after the epidemic in Shanxi province is taken as an example to illustrate and estimate the fluctuation of railway railcar flow. The coal output of Shanxi province in 2019 and 2020 is shown in Tables 7 and 8.

As can be seen from Tables 7 and 8, the fluctuation range is $459.1/8016.3 = 5.73\%$ before the outbreak of COVID-19, i.e. in 2019, and the fluctuation range is $872.5/8761 = 9.96\%$ after the outbreak of COVID-19, i.e. in 2020. Obviously, the fluctuation range after the outbreak is larger, so we used the fluctuation range of $9.96\% \approx 10\%$ to carry out the study. We use the discrete uniform distribution to describe the railcar flow as a random variable. Its mean is the value of the average daily railcar flow shown in Table 9. Its fluctuation range is arranged within the range from −10% to 10%. The railcar flow data with the uniform distribution function is displayed in Table 9.

Since the railcar flow data obtained throughout the planning horizon are insufficient for solving the OMDD model, we use the proposed uniform distribution function to expand 56 railcar flows listed in Table 4. Here, the expanding size of railcar flow data is set as $\tilde{r} = 100$ for each railcar flow.

### 5.3. Train formation plan based on the OMAD model

The daily average railcar flow value in Table 2 is put into the OMAD model. Because optimization software LINGO 18.0 is fast in calculation and convenient in programming, we use it to solve the model

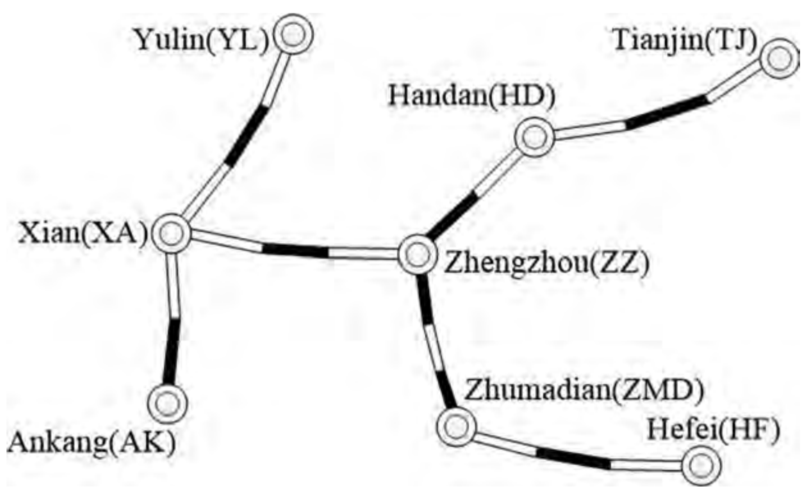



**Figure 3.** Railway network composed of eight technical stations in China.

**Table 4.** Parameters of technical stations.

| No. | Station | Assembling parameter | Assembling time | Reclassification time | Classification capacity | Track number |
|---|---|---|---|---|---|---|
| 1 | AK | 11.3 | 2.5 | 3.4 | 580 | 7 |
| 2 | XA | 11.5 | 2.3 | 3.9 | 750 | 9 |
| 3 | ZZ | 11.2 | 2.8 | 4.2 | 900 | 10 |
| 4 | HD | 11.6 | 2.4 | 4.5 | 836 | 9 |
| 5 | TJ | 11.8 | 2.6 | 3.6 | 695 | 6 |
| 6 | YL | 11.7 | 2.9 | 3.8 | 660 | 8 |
| 7 | ZMD | 11.4 | 2.5 | 3.7 | 625 | 10 |
| 8 | HF | 11.6 | 2.7 | 3.7 | 740 | 7 |

Note: 1 * The number in column 1 denotes the serial number of technical station.

**Table 5.** Value of daily average railcars flow from origin to destination.

| No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 132 | 116 | 178 | 154 | 186 | 142 | 189 |
| 2 | 164 | 0 | 246 | 173 | 155 | 79 | 132 | 118 |
| 3 | 146 | 182 | 0 | 166 | 144 | 239 | 134 | 186 |
| 4 | 153 | 149 | 86 | 0 | 86 | 234 | 175 | 158 |
| 5 | 147 | 158 | 68 | 156 | 0 | 127 | 108 | 122 |
| 6 | 178 | 163 | 114 | 164 | 185 | 0 | 56 | 137 |
| 7 | 164 | 152 | 84 | 136 | 175 | 126 | 0 | 146 |
| 8 | 154 | 184 | 156 | 179 | 178 | 165 | 149 | 0 |

Note: 1 * The number in column 1 denotes the serial number of technical station.

**Table 6.** Trip time between pairs of adjacent technical stations and critical railcar flow volume.

| Rail section | Trip time of district trains | Trip time of pick-up trains | volume of upward rail section | volume of downward rail section |
|---|---|---|---|---|
| 1–2 | 2.8 | 4.9 | 147.50 | 137.50 |
| 2–3 | 5.6 | 8.2 | 124.02 | 114.07 |
| 3–4 | 3.5 | 5.4 | 131.06 | 149.65 |
| 4–5 | 4.5 | 6.8 | 136.91 | 130.20 |
| 2–6 | 6.7 | 9.1 | 129.08 | 118.30 |
| 3–7 | 3.2 | 5.1 | 131.06 | 141.25 |
| 7–8 | 3.8 | 5.7 | 141.25 | 138.70 |

Note: 1 * The upward rail section in column 4 denotes the rail section from a small serial number technical station to a large serial number technical station. And downward rail section in column 5 represents the rail section from a large serial number technical station to a small serial number technical station.

**Table 7.** Monthly coal output of Shanxi province in 2019 (10,000 tons).

| Month | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | Mean |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coal output | 6732.9 | 6732.9 | 8282.8 | 8019.5 | 8477.1 | 8782 | 8420.2 | 8254.6 | 8182 | 8174.2 | 8308.4 | 7828.4 | 8016.3 |
| The difference from the mean | 1283.3 | 1283.3 | 266.5 | 3.2 | 460.8 | 765.7 | 403.9 | 238.3 | 165.7 | 157.9 | 292.1 | 187.9 | 459.1 |

Note: 1 * The data comes from China Statistical Yearbook.

**Table 8.** Monthly coal output of Shanxi province in 2020 (10,000 tons).

| Month | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | Mean |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Coal output | 6340 | 6340 | 9479.4 | 8595.4 | 8533.4 | 9298.3 | 8946.7 | 9477.7 | 9490.7 | 9543.5 | 9494.6 | 9592.3 | 8761 |
| The difference from the mean | 2421 | 2421 | 718.4 | 165.6 | 227.6 | 537.3 | 185.7 | 716.7 | 729.7 | 782.5 | 733.6 | 831.3 | 872.5 |

Note: 1 * The data comes from China Statistical Yearbook.

**Table 9.** Railcar flow data with uniform distribution function.

| Railcar flow | Uniform distribution function | Railcar flow | Uniform distribution function | Railcar flow | Uniform distribution function |
|---|---|---|---|---|---|
| $n_{12}$ | $U$(118, 146) | $n_{46}$ | $U$(210, 258) | $n_{73}$ | $U$(75, 93) |
| $n_{13}$ | $U$(104, 128) | $n_{47}$ | $U$(157, 193) | $n_{72}$ | $U$(136,168) |
| $n_{14}$ | $U$(160, 196) | $n_{48}$ | $U$(142, 174) | $n_{71}$ | $U$(147, 181) |
| $n_{15}$ | $U$(138, 170) | $n_{56}$ | $U$(114, 140) | $n_{65}$ | $U$(166, 204) |
| $n_{16}$ | $U$(167, 205) | $n_{57}$ | $U$(97, 119) | $n_{64}$ | $U$(147, 181) |
| $n_{17}$ | $U$(127, 157) | $n_{58}$ | $U$(109, 135) | $n_{63}$ | $U$(102, 126) |
| $n_{18}$ | $U$(170, 208) | $n_{67}$ | $U$(50, 62) | $n_{62}$ | $U$(146, 180) |
| $n_{23}$ | $U$(221, 271) | $n_{68}$ | $U$(123, 151) | $n_{61}$ | $U$(160, 196) |
| $n_{24}$ | $U$(155, 191) | $n_{78}$ | $U$(131, 161) | $n_{54}$ | $U$(140, 172) |
| $n_{25}$ | $U$(139, 171) | $n_{87}$ | $U$(134, 164) | $n_{53}$ | $U$(61, 75) |
| $n_{26}$ | $U$(71, 87) | $n_{86}$ | $U$(148, 182) | $n_{52}$ | $U$(142, 174) |
| $n_{27}$ | $U$(118, 146) | $n_{85}$ | $U$(160, 196) | $n_{51}$ | $U$(132, 162) |
| $n_{28}$ | $U$(106, 130) | $n_{84}$ | $U$(161, 197) | $n_{43}$ | $U$(77, 95) |
| $n_{34}$ | $U$(149, 183) | $n_{83}$ | $U$(140, 172) | $n_{42}$ | $U$(134, 164) |
| $n_{35}$ | $U$(129, 159) | $n_{82}$ | $U$(165, 203) | $n_{41}$ | $U$(137, 169) |
| $n_{36}$ | $U$(215, 263) | $n_{81}$ | $U$(138, 170) | $n_{32}$ | $U$(163, 201) |
| $n_{37}$ | $U$(120, 148) | $n_{76}$ | $U$(113, 139) | $n_{31}$ | $U$(131, 161) |
| $n_{38}$ | $U$(167, 205) | $n_{75}$ | $U$(157 ,193) | $n_{21}$ | $U$(147, 181) |
| $n_{45}$ | $U$(77, 95) | $n_{74}$ | $U$(122 ,150) | | |

Note: 1 * The symbol $U$ in columns 2, 4, and 6 denotes the uniform distribution function, and the number expresses the upper and lower bounds of uniform distribution.

and obtain the optimal global solution. The solving work has been executed using a computer with Intel(R) Core(TM) i5-8400 h CPU @ 2.80 GHz 2.30 GHz. The optimal train formation plan is shown in Table 10 and Figure 4. Table 10 lists the rail section, consolidated rail flow, railcar volume, and train frequency. Figure 4 shows block-train assignment decisions, including blocking plan and train type. The solid line with a double arrow means that it provides direct train service for the block between pairs of nonadjacent technical stations. The solid line with a single arrow means that it gives district

**Table 10.** The train formation plan based on the OMAD model.

| No. | Rail section | Consolidated railcar flow | Railcar volume | Train frequency | No. | Rail section | Consolidated railcar flow | Railcar volume | Train frequency |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1–2 | $n_{12}+n_{13}+n_{17}$ | 390 | 7.09 | 23 | 8–4 | $n_{84}$ | 179 | 3.25 |
| 2 | 1–4 | $n_{14}$ | 178 | 3.24 | 24 | 8–3 | $n_{83}$ | 156 | 2.84 |
| 3 | 1–5 | $n_{15}$ | 154 | 2.80 | 25 | 8–2 | $n_{82}$ | 184 | 3.35 |
| 4 | 1–6 | $n_{16}$ | 186 | 3.38 | 26 | 7–6 | $n_{76}+n_{86}$ | 291 | 5.29 |
| 5 | 1–8 | $n_{18}$ | 189 | 3.44 | 27 | 7–5 | $n_{75}$ | 175 | 3.18 |
| 6 | 2–3 | $n_{23}+n_{13}+n_{63}$ | 476 | 8.65 | 28 | 7–4 | $n_{74}$ | 136 | 2.47 |
| 7 | 2–4 | $n_{24}$ | 173 | 3.15 | 29 | 7–3 | $n_{73}$ | 84 | - |
| 8 | 2–5 | $n_{25}$ | 155 | 2.82 | 30 | 7–2 | $n_{72}$ | 152 | 2.76 |
| 9 | 2–6 | $n_{26}$ | 79 | - | 31 | 7–1 | $n_{71}+n_{81}$ | 318 | 5.78 |
| 10 | 2–7 | $n_{27}+n_{17}+n_{28}+n_{67}$ | 448 | 8.15 | 32 | 6–5 | $n_{65}$ | 185 | 3.36 |
| 11 | 3–4 | $n_{34}$ | 166 | 3.02 | 33 | 6–4 | $n_{64}$ | 164 | 2.98 |
| 12 | 3–5 | $n_{35}$ | 144 | 2.62 | 34 | 6–2 | $n_{62}+n_{63}+n_{67}$ | 333 | 6.05 |
| 13 | 3–6 | $n_{36}+n_{56}$ | 366 | 6.65 | 35 | 6–1 | $n_{61}$ | 178 | 3.24 |
| 14 | 3–7 | $n_{37}+n_{57}$ | 242 | 4.40 | 36 | 5–4 | $n_{54}$ | 156 | 2.84 |
| 15 | 3–8 | $n_{38}+n_{58}$ | 308 | 5.60 | 37 | 5–3 | $n_{53}+n_{56}+n_{57}+n_{58}$ | 425 | 7.73 |
| 16 | 4–5 | $n_{45}$ | 86 | - | 38 | 5–2 | $n_{52}+n_{51}$ | 305 | 5.55 |
| 17 | 4–6 | $n_{46}$ | 234 | 4.25 | 39 | 4–3 | $n_{43}$ | 86 | - |
| 18 | 4–7 | $n_{47}+n_{48}$ | 333 | 6.05 | 40 | 4–2 | $n_{42}$ | 149 | 2.71 |
| 19 | 6–8 | $n_{68}$ | 137 | 2.49 | 41 | 4–1 | $n_{41}$ | 153 | 2.78 |
| 20 | 7–8 | $n_{78}+n_{28}+n_{48}$ | 422 | 7.67 | 42 | 3–2 | $n_{32}+n_{31}$ | 328 | 5.96 |
| 21 | 8–7 | $n_{87}+n_{86}+n_{81}$ | 468 | 8.51 | 43 | 2–1 | $n_{21}+n_{31}+n_{51}$ | 457 | 8.31 |
| 22 | 8–5 | $n_{85}$ | 178 | 3.24 | | | | | |

Note: 1 * Railcar volume in column 4 denotes the number of railcars carried by train per day in the rail section.
2* Train frequency in column 5 denotes the number of trains operating per day in the rail section.
3* The information in rows 9, 16, 29, and 39 denotes the railcar flow carried by the pick-up train.

train service for the block between pairs of adjacent technical stations. The dotted line with a single arrow means that it provides the pickup train service for the block between pairs of adjacent technical stations.

We use two measures to evaluate the performance of the OMAD model, which are railcar-hour consumption and the satisfying rate of railway network capacity. The railcar-hour consumption measure is the objective function value of the model. The satisfying rate of railway network capacity represents the proportion of feasible solutions that satisfy the model's constraints. Railcar-hour consumption of the optimal train formation plan obtained by solving the OMAD model is 32,662.2. The satisfying rate of railway network capacity is 71%.

### 5.4. *Train formation plan based on the OMDD model*

The 100 groups expanding data of 56 railcar flows generated by the approach given in section 4.2 are all put into the OMDD model. And then, we use the procedure developed in section 3 to transform the probability constraints of the OMDD model into deterministic expression. Let the parameter $\alpha$ be 0.96, which means the acceptance rate of the

**Table 11.** The train formation plan based on the OMDD model.

| No. | Rail section | Consolidated railcar flow | Railcar volume | Train frequency | No. | Rail section | Consolidated railcar flow | Railcar volume | Train frequency |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1–2 | $n_{12}+n_{13}+n_{17}$ | 390 | 7.09 | 23 | 8–4 | $n_{84}$ | 179 | 3.25 |
| 2 | 1–4 | $n_{14}$ | 178 | 3.24 | 24 | 8–3 | $n_{83}$ | 156 | 2.84 |
| 3 | 1–5 | $n_{15}$ | 154 | 2.80 | 25 | 8–2 | $n_{82}$ | 184 | 3.35 |
| 4 | 1–6 | $n_{16}$ | 186 | 3.38 | 26 | 7–6 | $n_{76}+n_{86}$ | 291 | 5.29 |
| 5 | 1–8 | $n_{18}$ | 189 | 3.44 | 27 | 7–5 | $n_{75}$ | 175 | 3.18 |
| 6 | 2–3 | $n_{23}+n_{13}+n_{63}$ | 476 | 8.65 | 28 | 7–4 | $n_{74}$ | 136 | 2.47 |
| 7 | 2–4 | $n_{24}$ | 173 | 3.15 | 29 | 7–3 | $n_{73}$ | 84 | - |
| 8 | 2–5 | $n_{25}$ | 155 | 2.82 | 30 | 7–2 | $n_{72}$ | 152 | 2.76 |
| 9 | 2–6 | $n_{26}$ | 79 | - | 31 | 7–1 | $n_{71}+n_{81}$ | 318 | 5.78 |
| 10 | 2–7 | $n_{27}+n_{17}+n_{28}+n_{67}$ | 448 | 8.15 | 32 | 6–5 | $n_{65}$ | 185 | 3.36 |
| 11 | 3–4 | $n_{34}$ | 166 | 3.02 | 33 | 6–4 | $n_{64}$ | 164 | 2.98 |
| 12 | 3–5 | $n_{35}$ | 144 | 2.62 | 34 | 6–2 | $n_{62}+n_{63}+n_{67}$ | 333 | 6.05 |
| 13 | 3–6 | $n_{36}+n_{56}$ | 366 | 6.65 | 35 | 6–1 | $n_{61}$ | 178 | 3.24 |
| 14 | 3–7 | $n_{37}+n_{57}$ | 242 | 4.40 | 36 | 5–4 | $n_{54}$ | 156 | 2.84 |
| 15 | 3–8 | $n_{38}+n_{58}$ | 308 | 5.60 | 37 | 5–3 | $n_{53}+n_{56}+n_{57}+n_{58}$ | 425 | 7.73 |
| 16 | 4–5 | $n_{45}$ | 86 | - | 38 | 5–2 | $n_{52}+n_{51}$ | 305 | 5.55 |
| 17 | 4–6 | $n_{46}$ | 234 | 4.25 | 39 | 4–3 | $n_{43}$ | 86 | - |
| 18 | 4–7 | $n_{47}+n_{48}$ | 333 | 6.05 | 40 | 4–2 | $n_{42}$ | 149 | 2.71 |
| 19 | 6–8 | $n_{68}$ | 137 | 2.49 | 41 | 4–1 | $n_{41}$ | 153 | 2.78 |
| 20 | 7–8 | $n_{78}+n_{28}+n_{48}$ | 422 | 7.67 | 42 | 3–2 | $n_{32}$ | 182 | 3.31 |
| 21 | 8–7 | $n_{87}+n_{86}+n_{81}$ | 468 | 8.51 | 43 | 3–1 | $n_{31}$ | 146 | 2.65 |
| 22 | 8–5 | $n_{85}$ | 178 | 3.24 | 44 | 2–1 | $n_{21}+n_{51}$ | 311 | 5.65 |

Note: 1 * Railcar volume in column 4 denotes the number of railcars carried by train per day in the rail section.
2* Train frequency in column 5 denotes the number of trains operating per day in the rail section.
3* The information in rows 9, 16, 29, and 39 denotes the railcar flow carried by the pick-up train.

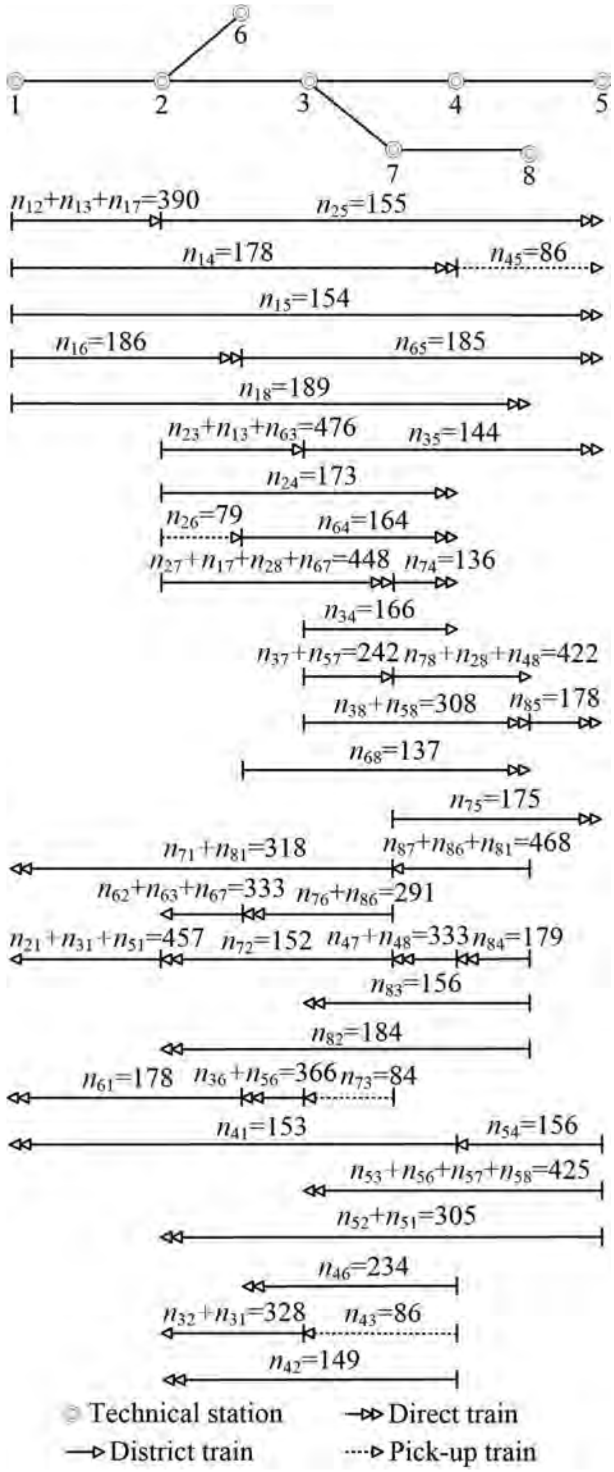


Figure 4. Optimal block-train assignment decisions obtained by solving the OMAD model.

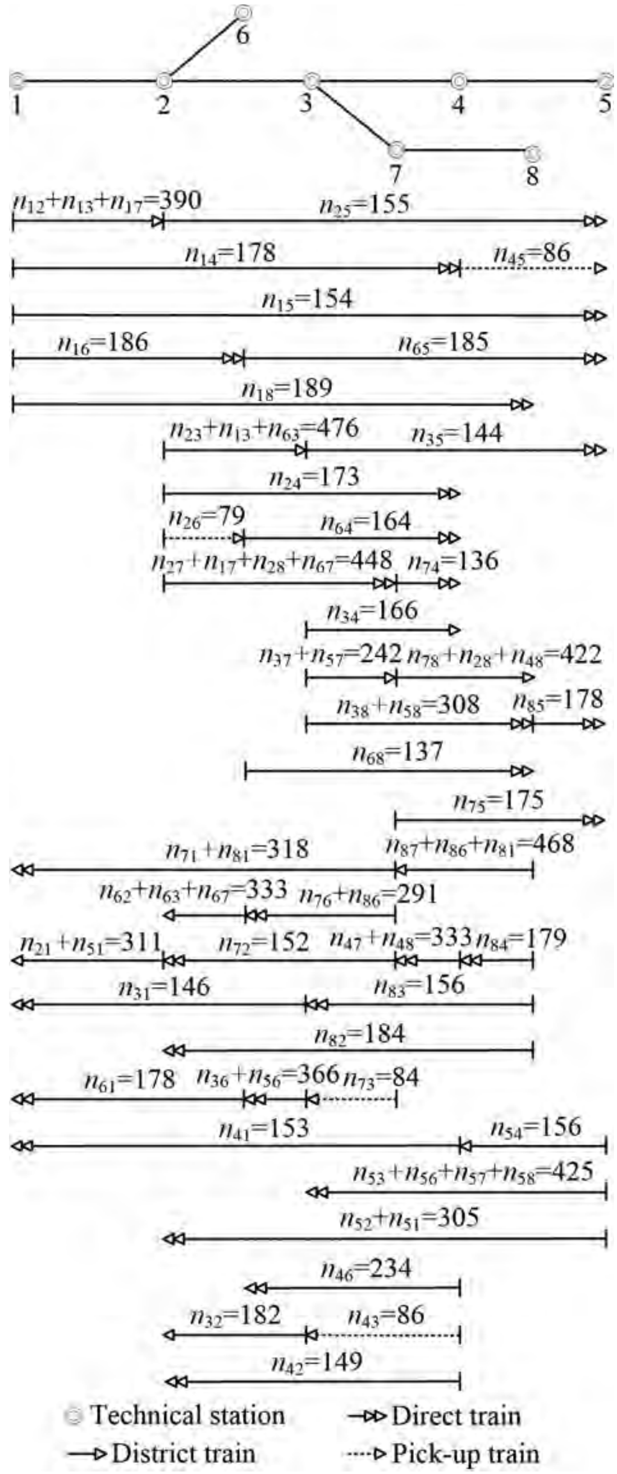


Figure 5. Optimal block-train assignment decisions obtained by solving the OMDD model.

train formation plan with high railcar-hour consumption is 96%. Let the parameters $\beta$ and $\gamma$ be 0.95, which means that the satisfying rate of the classification track number and technical station classification capacity is 95%. The OMDD model is solved by the same method and software. The optimal train formation plan is shown in Table 11 and Figure 5. Table 11 lists the rail section, consolidated rail flow, railcar volume, and train frequency. Figure 5 shows block-train assignment decisions, including blocking plan and train type.

We use the same measures to evaluate the performance of the OMDD model. Railcar-hour consumption of the optimal train formation plan obtained by the OMDD model is 32,708.8. The satisfying rate of railway network capacity is 98%.

**Table 12.** Comparison of the results of the OMAD model and the OMDD model.

| Model | Railcar-hour consumption | Satisfying rate |
|---|---|---|
| OMAD | 32662.2 | 71% |
| OMDD | 32708.8 | 98% |

Note: * Railcar-hour consumptions in column 2 imply the objective function value of two models.
*Satisfying rate in column 3 imply the ratio of feasible solutions obtained by solving the two models.

**Table 13.** Comparison of train formation plan between the OMAD model and the OMDD model under different fluctuation ranges of railcar flow data.

| Fluctuation range | Train formation plan of the OMAD model | | Train formation plan of the OMDD model | |
|---|---|---|---|---|
| | Railcar-hour consumption | Satisfying rate | Railcar-hour consumption | Satisfying rate |
| ±10% | 32662.2 | 71% | 32708.8 | 98% |
| ±20% | 32662.2 | 49% | 32736.3 | 99% |
| ±30% | 32662.2 | 44% | 32736.3 | 98% |
| ±50% | 32662.2 | 26% | 32804.0 | 95% |

### *5.5. Comparison of train formation plans*

We compare the solutions obtained from OMAD and OMDD. We can see the difference between the two measures, i.e. railcar-hour consumption and satisfying rate. The results are displayed in Table 12.

From the comparison results, it can be known that the difference between the total railcar-hour consumption obtained by the two models is less than 0.14%, while the satisfying rate of the OMDD model is 38% more than that of the OMAD model because the satisfying rate measure represents the proportion of feasible solutions which satisfy all constraints of the model. That is to say. This measure implies model stability. The OMDD model dramatically improves the solution's stability by slightly increasing railcar-hour consumption.

Because the fluctuation range of ± 10% is based on the coal output of Shanxi province every month, the region is large, and the time span is long. Considering the possibility of greater fluctuations when the region is smaller or the time span is shorter, sensitivity analysis was carried out on the fluctuation range of railcar flow. Under the condition that all parameters and calculation methods are the same, experimental cases with the fluctuation range of railcar flow of ± 20%, ±30%, and ± 50% were added to further illustrate the effectiveness of the proposed method. The results are shown in Table 13.

It can be seen from Table 13, for the OMAD model, the greater the fluctuation range of railcar flow data, the lower the satisfying rate of train formation plan, and the two are inversely proportional. When the fluctuation range of railcar flow data is within a range of ± 50%, the satisfying rate of train formation plan of the OMAD model is 26%, and it is almost impossible to use normally. For the OMDD model, with the increase of the fluctuation range of railcar flow data, the satisfying rate of train formation plan remains basically unchanged and can be maintained at a high level all the time, and the railcar-hour consumption increases slightly, but the increase is small. When the fluctuation range is ± 50%, the formation plan of the OMDD model has the highest railcar-hour consumption, which is 32,804, and the difference in railcar-hour consumption between the OMDD model and the OMAD model is 0.43%, the difference in satisfying rate was 265%. Obviously, a very small increase in railcar-hour consumption in exchange for a large increase in satisfying rate is very advantageous. It further validates the effectiveness of the proposed method.

## 6. Conclusions

The continuous spread and frequent disturbance of the COVID-19 epidemic have significantly impacted society's well production. As a result, sharp fluctuations in raw material and production consumption for the whole of society happen. Because railway transport mainly delivers bulk commodities and undertakes main-line transport task, the demand for the railcars used by technical stations is affected. To minimize the influence of frequently remaking train formation plans on railway transport management, the train formation plan fitting railcar demand fluctuation needs to be created as soon as possible.

In this paper, three train service types provided for railcar flow between pairs of technical stations, including direct train, district train, and pickup train, are analyzed. And then, the criterion for deciding train service type is designed. Furthermore, we proposed two optimization models for the TFP&TS problem. The objective function intends to minimize the sum of the railcar collection time of the outgoing train at the technical departure station, train reclassification time at the middle technical station, and additional operation time for the pickup train. Some constraints indicating a unique formation plan, the classification tracks number restriction, and technical station classification capacity restriction are considered. First, we use the average daily railcars flow data in the statistical period to develop the OMAD model. And then, we use directly the daily dynamic railcars flow data to build the OMDD model.

The OMAD model is a deterministic model, and the OMDD model is a stochastic model. To solve the OMDD model that includes three probability constraints, we develop the procedure of transforming probability constraints into deterministic expressions. Finally, we test the two models on a railway network composed of eight technical stations. The block-train assignment decisions, including the blocking plan and the service type of train, can be obtained using the optimization software LINGO 18.0 to solve the two models. When the fluctuation range of railcar flow data is within a range of ± 10%, compared to the train formation plan obtained using the OMAD model, the train formation plan obtained using the OMDD model increases the railway network capacity satisfy ratio by 38%. It increases the total railcar-hour consumption by 0.14%. When the fluctuation range is within a range of ± 20%, ±30%, and ± 50%, the results are similar. So we conclude that the OMDD model dramatically improves the solution's stability by slightly increasing railcar-hour consumption.

## Acknowledgments

The authors would like to thank the editors and anonymous referees for their careful and fruitful comments to improve the quality of this paper. This study is supported by Philosophy and Social Science Research Major Project of Henan Province Education Department of China (Grant no. 2022-YYZD-24), Henan Province Philosophy and Social Science Planning Project (Grant no. 2021BJJ087), and National Natural Science Foundation of China (Grant no. U1604150, U1804151, and 72201252).

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Notes on contributors

*Bing Li* is a Professor at the School of Management, Zhengzhou University, China. He currently holds the administrative post of vice Dean of School of Management of Zhengzhou University. He is a doctoral supervisor. He received the Ph.D. degree in transportation planning and management from Southweast Jiaotong University, Chengdu, China. His research interest is transportation planning and management.

*Shangtao Jiang* is a Master student at the School of Management, Zhengzhou University, China. His research interest is transportation planning and management.

*Yanjie Zhou* received Ph.D. degree from the Department of Industrial Engineering at Pusan National University in 2020 and received B.S. Degree and M.S. Degree in Computer Science and Computer Applied Technology from Zhengzhou University in 2012 and 2015, respectively. He is currently an associate professor with the school of management at Zhengzhou University. His research areas include optimization problems in industrial engineering, game theory, maritime logistics, and intelligence healthcare.

*Hua Xuan* is a Professor at the School of Management, Zhengzhou University, China. She received the Ph.D. degree in control theory and application from Northeastern University, Shenyang, China. Her current research interests include modeling and optimization in flow shop, scheduling of manufacturing system.

## ORCID

Bing Li http://orcid.org/0000-0001-8495-8336
Shangtao Jiang http://orcid.org/0000-0002-8136-5516
Yanjie Zhou http://orcid.org/0000-0003-2222-9140
Hua Xuan http://orcid.org/0000-0003-0962-949X

## References

[1] Bodin LD, Golden BL, Schuster AD, et al. A model for the blocking of trains. Transp Res E Method. 1980;14 (1–2):115–120.
[2] Chen CS, Dollevoet T, Zhao J. One-block train formation in large-scale railway networks: an exact model and a tree-based decomposition algorithm. Transportation Research Part B-Methodological. 2018;118:1–30.
[3] Newton H, Barnhart C, Vance P. Constructing railroad blocking plans to minimize handling costs. Transp Sci. 1998;32(4):330–345.
[4] Xiao J, Lin BL, Wang JX. Solving the train formation plan network problem of the single-block train and two-block train using a hybrid algorithm of genetic algorithm and tabu search. Transp Res Part C. 2018;86:124–146.
[5] Crevier B, Cordeau J-F, Savard G. Integrated operations planning and revenue management for rail freight transportation. Transportation Research Part B-Methodological. 2012;46(1):100–119.
[6] Xiao J, Lin BL. Comprehensive optimization of the one-block and two-block train formation plan. J Rail Transp Plann Management. 2017;218–236. DOI:10.1016/j.jrtpm.2016.09.002
[7] Keaton M. Designing optimal railroad operating plans: Lagrangian relaxation and heuristic approaches. Transportation Research Part B-Methodological. 1989;23(6):415–431.
[8] Keaton M. Designing railroad operating plans: a dual adjustment method for implementing Lagrangian relaxation. Transp Sci. 1992;26(4):263–279.
[9] Lin BL, Wang ZM, Ji LJ, et al. Optimizing the freight train connection service network of a large-scale rail system. Transportation Research Part B-Methodological. 2012;46(5):649–667.
[10] Yaghini M, Foroughi A, Nadjari B. Solving railroad blocking problem using ant colony optimization algorithm. Appl Math Modell. 2011;35(12):5579–5591.
[11] Gorman M. An application of genetic and tabu searches to the freight railroad operating plan problem. Ann Oper Res. 1998;78:51–69.
[12] Yaghini M, Momeni M, Sarmadi M, et al. A fuzzy railroad blocking model with genetic algorithm solution approach for Iranian railways. Appl Math Model. 2015;39(20):6114–6125.
[13] Barnhart C, Jin H, Vance P. Railroad blocking: a network design application. Journal Of Transportation Engineering Part A-Systems. 2000;48(4):603–614.

[14] Ahuja RK, Jha KC, Liu J. Solving real-life railroad blocking problems. Appl Math Model. 2007;37 (5):404–419.
[15] Fuegenschuh A, Homfeld H, Schuelldorf H. Single-car routing in rail freight transport. Transp Sci. 2015;49 (1):130–148.
[16] Alikhani-Kooshkak R, Tavakkoli-Moghaddam R, Jamili A, et al. Multi-objective mathematical modeling of an integrated train makeup and routing problem in an Iranian railway company. Sci Iran. 2019;26 (6):3765–3779.
[17] Sama M, D'Ariano A, Corman F, et al. A variable neighbourhood search for fast train scheduling and routing during disturbed railway traffic situations. Comput Oper Res. 2017;480–499.
[18] Fang K, Ke GY, Verma M. A routing and scheduling approach to rail transportation of hazardous materials with demand due dates. Eur J Oper Res. 2017;261 (1):154–168.
[19] Lin BL. An overall optimizing routing makeup model combined local and through freight trains. J China Railway Soc. 1996;18(5):1–7.
[20] Zhu E, Crainic TG, Gendreau M. Scheduled service network design for freight rail transportation. Oper Res. 2014;62(2):383–400.
[21] Lin BL, Zhao YN, Lin RX, et al. Integrating traffic routing optimization and train formation plan using simulated annealing algorithm. Appl Math Modell. 2021;93:811–830.
[22] Cao C, Feng Z. Optimal capacity allocation under random passenger demands in the high-speed rail network. Eng Appl Artif Intell. 2020;88:103363.
[23] Tarhini H, Bish DR. Routing strategies under demand uncertainty. Netw Spat Econ. 2016;16(2):665–685.