



# *Article* **Urban Infrastructure Construction Planning: Urban Public Transport Line Formulation**

**Silin Zhang <sup>1</sup> [,](https://orcid.org/0000-0001-7363-4045) Buhao Zhang <sup>2</sup> , Yi Zhao <sup>3</sup> , Shun Zhang <sup>4</sup> and Zhichao Cao 1,5,[\\*](https://orcid.org/0000-0002-7815-3889)**

- <sup>1</sup> School of Transportation and Civil Engineering, Nantong University, Nantong 226019, China; 12114213@bjtu.edu.cn
- <sup>2</sup> Zhejiang Intelligent Transportation Engineering Technology Research Center, Hangzhou 310023, China; bhzhang@enjoyor.net
- <sup>3</sup> Standards and Metrology Research Institute, China Academy of Railway Sciences Corporation Limited, Beijing 100081, China; yizhao2013@bjtu.edu.cn
- 4 Jiangsu CRRC Mechanical & Electric Technology Co., Ltd., Nantong 226019, China; zhangshun@crrcitet.com
- <sup>5</sup> Hubei Key Laboratory of Transportation Internet of Things, Wuhan University of Technology, Wuhan 430070, China
- **\*** Correspondence: caozhichao@bjtu.edu.cn; Tel.: +86-130-23562889

**Abstract:** Urban public transport line formulation has its appeal in promoting public convenience and developing environmentally friendly cities. During the bus line planning stage, the line frequency and stop location determination is a key issue for decision makers. Our study focuses on the integrated formulation problem between line frequency and stop planning featuring multi-type vehicles. The multi-type vehicles are able to accommodate the various passenger demands at either peak hours or off-peak hours. The a priori magnitudes of user demands are investigated by drone-based technique methods in the tactical-level plan. The collected geospatial data can assist the public transport user forecast. A mixed-integer linear programming (MILP) model is proposed. The objective is to minimize the walking cost of passengers, the building cost of stops, and the operation cost of service frequency. The effectiveness of the model is validated by a real case in Nantong, China. CPLEX is used to resolve the MILP model. Yielding to the budget constraint, in high-price, medium-price, and low-price scenarios, the optimal high-quantity stop scheme can save 3.04%, 3.11%, and 3.38% in overall cost compared with the medium-quantity stop scheme, respectively; their cost savings are 8.53%, 8.70%, and 9.09% more than the costs of the low-quantity stop scheme.

**Keywords:** public transport; line formulation; construction planning; multi-type vehicle; stop plan

# **1. Introduction**

China is enthusiastic about reducing carbon emissions by actively developing public transport (PT) buildings. Pure electric PT featuring its public convenience and environmentally friendly low-pollution merits is credited for its green development. A majority of megacities, such as Beijing, Hong Kong, Singapore, and Tokyo, prioritize supporting policy that supports facilitating the public in favoring and becoming accustomed to riding buses/subways. The PT mode-sharing ratio accounts for about 70–90% [\[1\]](#page-17-0). For this purpose, most cities in China are positively developing PT infrastructure construction, the basis of which is formulating the optimal stop location and line frequency. The nature of the line planning work should match the user-oriented passenger demand distributions well. Thus, the potential passenger demand points play a core role in the line formulation stage. Specifically, the origin and destination of the passenger demand points are defined as the origin-destination (OD) pairs, such as home, workplace, schools, leisure, and shopping. In other words, the stop locations mean different walking distances between stops and these centers, which lead to heterogeneous travel costs for passengers. Furthermore, the reasonable bus line frequency and stop planning have a considerable contribution to aggrandizing PT's attractiveness by a valid alternative to private cars.



**Citation:** Zhang, S.; Zhang, B.; Zhao, Y.; Zhang, S.; Cao, Z. Urban Infrastructure Construction Planning: Urban Public Transport Line Formulation. *Buildings* **2024**, *14*, 2031. [https://doi.org/10.3390/](https://doi.org/10.3390/buildings14072031) [buildings14072031](https://doi.org/10.3390/buildings14072031)

Academic Editor: Pramen P. Shrestha

Received: 9 June 2024 Revised: 26 June 2024 Accepted: 29 June 2024 Published: 3 July 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license [\(https://](https://creativecommons.org/licenses/by/4.0/) [creativecommons.org/licenses/by/](https://creativecommons.org/licenses/by/4.0/)  $4.0/$ ).

Typically, PTs are confronting temporal variations of passenger demand between peak hours and off-peak hours. This variability entails us to employ multi-type vehicles to deal with the demand distribution fluctuation challenges. On the other hand, the increase in service frequency is indeed able to reduce users' waiting times but also means more operation cost and carbon emission. Thus, it is necessary to implement an integrated optimization between line frequency and stop plan featuring multi-type vehicles in facilitating the geospatial data application.

Our study aims to determine the reasonable number and locations of stops along a fixed bus route. To be clear, PT routing does not belong in this work. In other words, the routing trend is determined and the incumbent formulating task is to examine the locations/number of stop candidates. For the sake of convenience, the set of candidate stops is not groundless, but rather depends on the land use property. In this study, we define the different land modules as the demand points (or the so-called interest points). A tactical drone-based survey that needs to be fulfilled before the formulation provides a data estimation basis for the PT passenger demand. In particular, 'the PT passenger demand' refers to the number of users who choose the PT mode depending on the urban PT mode-sharing calculation. Thus, private transport demands are not accounted for in the paper.

The aim of a building plan is to seek for the trade-off in net public interests between users and operators. These interests are defined in a quantized manner as threefold: (i) users' walking costs from centers to stops, (ii) new building costs for entire stops, and (iii) operation costs of service frequencies. This integrated programming model will help the planners to examine the adaptive schemes based on the real investigation information. Since the PT stop plan featuring multi-type vehicle outcomes is limited in the literature, we hereinafter focus on this topic as per the programming model.

## *1.1. Literature Review*

This section introduces three directions of urban PT infrastructure construction planning: stop determination, line frequency, and transport accessibility. They provide a solid research basis for our work.

## <span id="page-1-0"></span>1.1.1. Stop Determination Problem

The methodology of stop determination in the railway study field provides an essential base for bus stop formulation. Some state-of-the-art modelling approaches have inspired us to develop our study concept. Yang et al. [\[2\]](#page-17-1) propose a collaborative optimization method for both stop planning and train schedules with different speeds, specifically that refer to the maximum velocities of 250 km/h and 300 km/h, respectively. To handle this integration problem, an MILP model is built and resolved by a CPLEX solver availably. Repolho et al. [\[3\]](#page-17-2) construct a strategic planning model for a high-speed rail corridor. A mixed-integer optimization model with two decision variables, i.e., station location and train type, is proposed. Throughout, a real case of #8 line involving a high-speed rail project in Portugal, the model is solved through the commercial software to validate its effectiveness. On the operation level, Yuan et al. [\[4\]](#page-17-3) propose a skip-stop strategy schedule optimization model to cope with the congestion demand in rail transit. Multi-criteria on passengers' waiting times and station crowding are taken into consideration in the model. A decomposition and approximate dynamic programming approach is devised to resolve the model. Chang et al. [\[5\]](#page-17-4) propose a multi-objective programming model for the high-speed rail station schemes. The ternary minimum objectives are to reduce the rolling stocks, the running distances of trains, and total travel times of passengers. The model is handled by a fuzzy mathematical programming approach and clarified by a Taiwan's high-speed rail case.

Tang et al. [\[6\]](#page-17-5) propose a skip-stop strategic electrical bus schedule optimization model. A heuristic algorithm invoking a left-shifting trip arrival time pretreatment is designed to handle the model yielding to nonlinear objectives and nonlinear constraints. A Dandong bus case in China demonstrates that the overall operation cost savings is 15.09% more and

the energy use ratio is increased by 9.02%. Given OD passenger demand pairs along a highspeed railway line, Cacchiani et al. [\[7\]](#page-17-6) focus on the sequence of trains' departures/arrivals and simultaneously determine the stop locations. The robustness at the planning level is carried out to handle the uncertain passenger demand fluctuation problem. The fundamental model and the extended robust models are solved by CPLEX and testified by the real data of the Wuhan–Guangzhou high-speed railway corridor. Shao et al. [\[8\]](#page-17-7) focus on an equity opportunity for users who board from different stations. An integration model between stop planning and timetabling is constructed with a focus on exploring fairness considerations. An adaptive large-scale neighborhood search algorithm is used and verified based on the Chinese Shanghai–Jinshan suburban railway.

## 1.1.2. Line Frequency Problem

Canca et al. [\[9\]](#page-17-8) developed a line frequency and transport capacity optimization model for a dense railway network. In this model, the OD passenger demand is fixed as the input. The model assigns the reasonable line frequencies and capacities over the same open tracks. The multi-objectives are to minimize the average trip time of all trains, the costs of operation, maintenance, and fleet acquisition, synchronously. The model is solved by the extended cutting plane method. Shi et al. [\[10\]](#page-17-9) balance the non-equilibrium user aggregation as per a timetable and stop plan collaborative optimization model. An efficient iterated local search algorithm is tailored to deal with a large case on the Fangshan line of the Beijing metro. In order to control the passenger flow density in the COVID-19 pandemic period, Weert and Gkiotsalitis [\[11\]](#page-17-10) present a mixed-integer quadratic programing model to optimize bus line frequencies to conform public health security. Their approach deals with high-demand bus lines in the Netherlands. The results indicate that short turning can help in reducing the social distance. Liang et al. [\[12\]](#page-17-11) developed a self-equalizing bus headway-control and stop-skipping model to eliminate the bunching phenomenon. An optimal holding time calculation algorithm is proposed to guarantee high-frequency regulation.

Considering crowding discomfort experiences, time-dependent demand, left-behind users, and random travel times, Sadrani et al. [\[13\]](#page-17-12) built a service frequency and vehicle size integration optimization model for automated bus systems. An automated bus deployment case demonstrates the model's effectiveness in reducing in-vehicle passenger crowding ranging from 20% to 60%. Fei et al. [\[14\]](#page-17-13) focus on the profitability of electric bus development and power grid services. The 'Frequency Control Reserve' contract is examined, with a focus on how discharge power responds to immediate grid requests. The analysis delves into both summer and winter frequency optimization, drawing insights from data pertaining to a Japanese city.

## <span id="page-2-0"></span>1.1.3. Public Transport Accessibility

In their paper, Tong et al. [\[15\]](#page-18-0) present a quantitative estimation method to measure urbanization. Walking distance and bus travel time are taken into account. An integrated algorithm between Dijkstra and Monte Carlo simulation is proposed to resolve the model. Kim and Kim [\[16\]](#page-18-1) optimized two key operational tactics: distance-based fares and headway. With maximizing the interest of bus operators, the budget, demand, and maximum headway constraints were taken into consideration. The fixed PT line case testifies the feasibility of the proposed method in Springfield, USA. Yang and Liang [\[17\]](#page-18-2) focus on the connectivity metrical approach between rail and PT. An entropy weighting method is used to derive the connection indicators. A case study of Wuxi, China, rail and PT, was validated to find that 57.5% of rail stations in Wuxi possess low connectivity. Nadimi et al. [\[18\]](#page-18-3) strike the trade-off for school bus service between social equity and sustainable operations. Grounded theory and structural equation model quantitatively examined the influence factors, i.e., cost, walking times, waiting times, running times, and fixed/flexible route. Li et al. [\[19\]](#page-18-4) developed an on-demand PT service for a faraway community. A bi-level programming model is proposed where the upper level determines an optimized frequency and ticket price schemes while the lower level formulates the travelers' choosing probability. Risso et al. [\[20\]](#page-18-5) devise a multi-level PT network in Montevideo, Uruguay. CPLEX and the evolutionary algorithm are utilized to resolve the proposed multi-objective model. The objective is to minimize the overall cost of flow paths. Su et al. [\[21\]](#page-18-6) present an advanced PT accessibility estimation technique. The users' choice preferences for different PT modes are analyzed based on the data of Shenzhen.

Here, we distinguish the stop planning problem and line frequency problem as follows. The former decides the number and locations of stop along the bus corridor. The latter refers to the number of trips in a certain period, e.g., one hour or one day. The difference between them lies in the space assignment and time schedule. They are important aspects for public transport formulation. On the other hand, 'stop' directly impacts users' accessibility (that highly depends on the walking distance to stops) and potential demands. 'Frequency' means the amount of service supplies. Thus, the three modules—stop, frequency, and accessibility—jointly compose the study framework.

## *1.2. Contributions and Features*

Concentrating on the above three modules, our study integrates them together into a fundamental collaborative optimization model, in which the objective function is to minimize the costs of (i) building stops, (ii) users' walking, namely, accessibility, and (iii) operation for service frequency; explicitly, they are one-to-one correspondences for Sections [1.1.1](#page-1-0)[–1.1.3,](#page-2-0) respectively. To date, a limited amount of literature has studied this integrated optimization problem.

Drone-based investigation enlightens us to adapt a priori passenger flow estimation for facilitating PT line formulation projects. Another significant insight is that the programming model is expected to cope with the integrated line planning problem at the tactical level. The problem covers stop patterns, multi-type vehicle schedules, and service frequency subject to the budget constraints.

This work aims to make the following contributions for PT line formulation as per drone-acquired data. (1) Our work innovatively adapts the drone-based demand investigation method to obtain the order of magnitudes of potential users. During the planning stage, either a hundred-level or thousand-level of demands are feasible for precision to direct a brand-new PT line formulation/construction. We refer to this stage as the pre-feasibility estimation. (2) We integrate the stop plan with multi-type vehicle and service frequency jointly in a programming model. The objective is to minimize (i) the building cost of new stops, (ii) the walking time costs of all users, and (iii) the operation cost for service frequency, simultaneously. Yielding to the different budgets, high-quantity, medium-quantity, and low-quantity stop schemes are derived to analyze the trade-off between supplies and demands. Distinct from previous studies, we build an integrated optimization model between multi-type vehicle assignment and departure frequency. To the best of the authors' knowledge, it is the first time to provide a joint optimization method to facilitate the stop location formulation problem.

The rest of this paper is organized as follows. Section [2](#page-3-0) introduces the bus line formulation problem statement involving with the number of and location of stops, and service frequency. Section [3](#page-6-0) proposes a mixed-integer linear programming model to achieve the integrated optimization. Section [4](#page-8-0) testifies the effectiveness of the model and analyzes the applicability of multiple schemes. Section [5](#page-15-0) draws the study conclusions and extends the future directions.

## <span id="page-3-0"></span>**2. Problem Statement**

In 2023, Nantong, as a mid-size Chinese city, opened the first subway line No. 1. At the same time, Nantong public transport agency had been transforming diesel buses to electric bus vehicles. In order to facilitate an optimal geospatial data application for urban infrastructure construction, the new line formulation is addressed for accommodating the new passenger flow demand. In this regard, a formulation model is supposed to determine the reasonable number of and locations of stops for a new bus line building. Furthermore,

the service frequency scheme of the line is optimized subject to different budgets. Herein, the multi-type vehicles are taken into consideration in this study to accommodate different level of demands. The appearance and interior of multi-type vehicles are exhibited in<br>Figures of the multi-legale of the multi-legale of the latter concerns the multi-legale of the multi-legale of Figure [A1](#page-17-14) in Appendix [A.](#page-17-15) In detail, Types 1, 2, and 3 refer to the larger, medium, and right *III in Appendix 11.* In detail, 19 pes 1, 2, and 3 feet to the larger, meaning and small vehicle, respectively. It means the load capacity is different from its in-vehicle space. Simultaneously, their purchase costs are considerably distinguished. Thus, our study seeks for the optimal trade-off between service spaces and operation costs.

Allowing for the accessibility of public transport's serviceability, both the number Anowing for the accessionity of public transport's serviceability, both the humber<br>and locations of stops contribute to the attractiveness of choosing the PT mode. The former focuses on the coverage scale of the bus service while the latter concerns the multilegged walking distances. In detail, the number of stops are subject to the construction cost budget and land usage. The spacing distance between two adjacent stops needs to be considered. On the other hand, the stop locations are key elements on whether the Feasible considered. On the other hand, the stop foculous are key clements on whether the convenience of choosing the bus service is available for the surrounding demand users. If walking times/distances are too long, un-timely travel and large baggage would become the passengers' hurdles on riding buses. On the contrary, the short walking times mean a considerable cost and resource supply, i.e., a majority of stops. A questionnaire-based<br>proximately contributed with the walking the walking times of compute accessoring survey is a feasible technique to investigate the walking times of common passengers. But very is a reashere teening to investigate the wanting times or common passengers.<br>However, the absence of incumbent bus stops makes us unable to follow this means. Thus, the distances from the formulated stop candidates and walk speed of comment users can help to approximately compute the walking times.

Concentrating on the new line formulation problem, we seek for an interest trade-off Concentrating on the new line formulation problem, we seek for an interest trade-on<br>between the number of and locations of stops, shown in Figure [1.](#page-4-0) In the following case (Section 4), we define and compare the high-quantity, medium-quantity, and low-quantity stop schemes. The green points (demand points) refer to the residential districts and schools as well as hospitals that usually accommodate a high density of users. Without loss of generality, we need to provide and ensure that one bus stop (red circle) can serve two to five demand points. The blue circles are the start and end points, defined as depots. The five demand points. above works jointly make up the PT route formulation.

<span id="page-4-0"></span>

**Figure 1.** Stop planning problem formulation scenario to accommodate surrounding demands.

At the planning stage, the available line frequency is the other important service degrees, and the operation costs. Over-crowding and left-behind phenomena resulting role in crowing comfort and left-behind possibilities. Hence, our study further focuses on the integration optimization between line frequency and multi-type vehicle options. For a<br>first line ages, the travel degree de degreed from degree degrips (centers are investigated meas the case, the travel definition costs from and lefting from  $\epsilon$  criters are the exigated.<br>by drone-based technique methods. The OD pairs are derived based on the OD distribution rate in this programming problem. Indeed, different vehicle types play a key role in this programming problem. supply basis. Typically, it mirrors the average waiting times, the in-vehicle crowding from full-load capacity would frustrate the users. Indeed, different vehicle types play a key fixed-line case, the travel demands derived from demand points/centers are investigated

To summarize, there are three core concerns to deal with in this study so as to strike te definition optimization between line frequency and multi-type vehicle options. For a fixedthe demand–supply balance problem as follows:

- (a) How many stops are supposed to be built to cover the accessibility of the surrounding residents (as the potential users)?
- (b) Where are the stop locations which achieve a short walking distance?
- (c) What are the schemes of the vehicle types yielding to different budgets?

#### *2.1. Assumption*

In order to facilitate this real-world problem, the related assumptions, terminologies, and modelling are proposed.

- (I) The multiple stop patterns yielding the different budget scenarios obtained in the programing model are analyzed based on the trade-off between users and operators.
- (II) The planned stops do not warrant to serve all neighboring *demand points* as per oneto-one correspondence. Although the stop candidate locations are considered to accommodate these demand points a priori, in general, they are partially elected by the criteria-based estimation. Both the economy and operation efficiency have a real impact on the number of stops.
- (III) The fleet sizes are assumed to be sufficient. In other words, vehicle procurement is not considered in this work.

## *2.2. Formulation Scale*

The OD trajectory means the process of passengers from the origin to the destination. Although most of the travel is served by buses, a walking process is necessary. Typically, the origin and the destination refer to home, workplace, schools, leisure, shopping and so on. In this study, we do not consider the discrete passenger demands but rather use passenger flow derived from the potential demand points in a cluster manner. Individual travel users are distributed randomly and hard to track. The item of '*demand points*' helps us to focus on major demand sources either from the origins or at the destinations, namely, *qoi* and *qid*, respectively. Quantitatively, the order of magnitudes of passengers in the *demand points* are feasible for PT line planning level. The radius covered by the demand points refers to ranging from 500 m to 1000 m, which allows passengers to walk from 5 min to 10 min.

The formulation problem focused on in this study covers how many *demand points* are along the PT line and where they are as well as what is the departure frequency and which vehicle types to choose.

# *2.3. Terminologies*

2.3.1. Sets

*I*: Set of stop candidates,  $I = \{i: i = 1, \ldots, j, \ldots, A\}$ , where *A* is the upper bound of the number of stop candidates.

*V*: Set of vehicle types,  $V = \{v: v = 1, \ldots, B\}$ , where *B* is the upper bound of the number of vehicle types.

*F*: Set of service frequencies,  $F = \{f: f = 1, \ldots, N\}$ , where *N* is the upper bound of the number of service frequencies.

*O*: Set of origin points,  $O = \{o: o = 1, \ldots, E\}$ , where *E* is the upper bound of the number of demand points.

*D*: Set of destination points,  $D = \{d: d = 1, \ldots, E\}$ , where *E* is the upper bound of the number of demand points, which is similar to set *O*.

*R*: Set of stop plans,  $R = \{\gamma : \gamma = 1, \ldots, \Re\}$ , where  $\Re$  is the maximum of stop patterns.

## 2.3.2. Parameters

 $q_{od}$ : The number of passengers from origin *o* (*o* ∈ *O*) to destination *d* (*d* ∈ *D*).

*toi*: Walking time from origin *o* to boarding stop *i*.

*tjd*: Walking time from alighting stop *j* to destination *d*.

*Cv*: Passenger-load capacity of vehicle type *v*.

*nv*: Maximum service frequency of vehicle type *v*.

 $\Omega_1$ : Unit cost of building one stop.

 $\Omega_2$ : Unit cost of passengers' walking times.

Ω*<sup>v</sup>* 3 : Unit cost of service frequency for vehicle type *v*.

2.3.3. Decision Variables

- $x_{\gamma v f} = 1$  if stop plan  $\gamma$  with vehicle type *v* and service frequency *f* is adapted; =0 otherwise.
- $\alpha_i = 1$  if stop plan includes candidate *i*; =0 otherwise.
- $\gamma_{ij} = 1$  if stop plan  $\gamma$  includes candidate *i* and *j*; =0 otherwise.

## <span id="page-6-0"></span>**3. Method**

# *3.1. Accessibility Constraint*

Basically, a reasonable stop plan is supposed to ensure a desired walking time for the surrounding residents, such as 5 min or 8 min. In fact, a long walking distance is defined as beyond the scope of bus accessibility. Then, walking times *Twalk* between stops and O/D are defined as per Equation (1).

$$
T_{walk} = \sum_{\begin{array}{l} o \in O, \quad i \in I, \\ d \in D, \quad j \in I, \\ o < d \quad i < j \end{array}} q_{od} \cdot (t_{oi} + t_{jd}), \tag{1}
$$

where  $q_{od}$  is the number of passengers from origin  $o$  ( $o \in O$ ) to destination  $d$  ( $d \in D$ ).  $t_{oi}$  is the walking time from origin  $o$  to stop  $i$  to board.  $t_{id}$  is the walking time from alighting stop *j* to destination *d*.

# *3.2. Stop Plan Constraint*

A stop pattern *γij* includes a series of stops that accommodate the circumambient residents. The greater the ridership, the greater the number of stops that are needed to supply service accessibility. Yielding to the building cost, the judgment whether one particular stop *α<sup>i</sup>* is to choose or not mainly depends on the number of demands. The corresponding constraints are formulated as follows:

$$
\gamma_{ij} = \alpha_i \cdot \alpha_j, \forall i \in I, j \in J,
$$
\n<sup>(2)</sup>

$$
\gamma_{ij} \le \alpha_i, \forall i \in I, j \in J,
$$
\n
$$
(3)
$$

$$
\gamma_{ij} \le \alpha_j, \forall i \in I, j \in J,
$$
\n<sup>(4)</sup>

$$
\alpha_i \le 1, \forall i \in I,\tag{5}
$$

$$
\alpha_i \geq 0, \forall i \in I,\tag{6}
$$

$$
\alpha_j \le 1, \forall j \in J,\tag{7}
$$

$$
\alpha_j \ge 0, \forall j \in J,\tag{8}
$$

where Equation (2) ensures stop pattern *γij* is equal to 1 only when stop candidates *α<sup>i</sup>* and  $\alpha$ <sup>*j*</sup> are 1, simultaneously. Equations (3) and (4) specify the relationship between the stop pattern *γij* and stop candidates *α<sup>i</sup>* , *α<sup>j</sup>* . Equations (5)–(8) indicate decision variables *α<sup>i</sup>* and *α<sup>j</sup>* are binary variables.

# *3.3. Service Capability Constraint*

Multi-type vehicles allow us to precisely improve service supply to attain an optimal PT source plan. The line frequency not only represents the operation costs but also determines users' waiting times. Furthermore, the multi-type load capacities impact whether the left-behind demands exist and how many. That is,

$$
\sum_{\gamma \in R} \sum_{v \in V} \sum_{f \in F} x_{\gamma v f} \cdot C_v \ge \sum_{\begin{array}{c} o \in O, \\ d \in D, \\ o < d \end{array}} q_{od},\tag{9}
$$

$$
n_v - \sum_{\gamma \in R} \sum_{f \in F} x_{\gamma v f} \ge 0,
$$
\n(10)

$$
N = \sum_{v \in V} n_v,\tag{11}
$$

where Equation (9) indicates that all service supply is equal to or more than all the OD demand. Equation (10) validates the relationship between maximum frequency of vehicle type  $n_v$  and decision variable  $x_{\gamma v f}$ . Equation (11) refers to the overall service frequency *N*, which helps for determining the headway value for an initial even-headway timetable. In practice, a basic even-headway timetable is usually employed at the planning level.

# *3.4. Objective Function*

We propose an overall objective function that integrates the net interests of operators and users. In detail, three items are considered simultaneously: (i) building cost of new stops, (ii) walking time costs of all users, and (iii) operation cost for service frequency. All costs are transformed in the same unit, U.S. dollars. The objective function (12) is given as

$$
Z = min[\Omega_1 \cdot \sum_{i \in I} \alpha_i + \Omega_2 \cdot T_{walk} + \sum_{v \in V} \Omega_3^v \cdot n_v],
$$
\n(12)

where the three items in Equation (12) correspond to the above sub-objectives (i), (ii) and (iii). The monetary coefficients  $\Omega_1$ ,  $\Omega_2$ , and  $\Omega_3^v$  can help transfer the number of stops  $\sum_{i\in I} \alpha_i$ , overall walking times *Twalk*, and frequency *N* to the uniform cost unit.

## *3.5. Model Framework*

Based on the model structure, the proposed optimization problem is a mixed-integer linear programming (MILP) model. The MILP model framework for the formulation problem reads as follows:

*Minimize* (12),

# *Subject to* (2)–(11).

The above model is a basic mathematical framework for devising a new public transport line. It should be noted that this MILP model is able to apply for the planning level without timetabling the design. Thus, this fundamental model owns the potential expansion sphere for meeting more particular requirements, for example, (i) one particular stop is located on one certain location although its demand is not booming now; and perhaps, it mainly depends on the future development direction or policy support; (ii) one stop that serves for the transfer passengers from airports; rail stations should be taken into consideration as one candidate. In summary, the fundamental model proposed has a capable of handling the specific and demand-oriented stop planning problem.

## *3.6. Computation Complexity*

Concentrating on the proposed MILP model, two pre-specified decision variables lead to a complex computation, namely, stop location judgment *α<sup>i</sup>* and vehicle type *v*. Theoretically, in order to cover the reasonable number of demand points, the stop candidate set tends to consider considerable locations to testify. For example, a single line with 30 stop candidates and two vehicle types will produce  $2^{30}\times 2^2$  possibilities. They are derived from two aforesaid variable sets {*I*} and {*V*}. In what follows, Table [1](#page-8-1) exhibits the overall number of decision variables and key constraints involved in the proposed model. To be clear, this

incumbent problem turns out to be a large-scale binary-variable linear programming model with a total number of  $10^9$  variables.

<span id="page-8-1"></span>**Table 1.** Number of numeration involved with related-variables and constraints.



#### <span id="page-8-0"></span>**4. Results**

In order to validate the availability of the proposed model in Section [3,](#page-6-0) we put it into use to a real case that belongs to a Nantong airport shuttle project in Nantong, China. The shuttle bus line is planned to be about 30 km as the total length. The line alignment is determined a priori but the possible stop locations along the line need to be formulated. Moreover, the service frequency and multiple vehicle types are supposed to be optimally devised to accommodate the various demand distribution.

The implementation of the developed model has two categories of data as the input: (i) data on cost, i.e., stops' building cost, travel cost for users, operation cost for multivehicle types; and (ii) data on the passenger demand desired to take the bus. In particular, the fundamental passenger investigation allows us to use drones to handle the demand magnitudes in Section [4.1.](#page-8-2)

## <span id="page-8-2"></span>*4.1. Drones-Implemented Demand Investigation*

For a brand-new public transport line formulation, the passenger flow does not yet generate. Essentially, the current passenger flow demand survey is a kind of forecast. Drones' application helps for surveying the geospatial information, i.e., area, land status, plot ratio, and the number/property of buildings. Thereby, we can estimate the number of potential users based on the travel intensity. Indeed, drones are used to handle the basic demand magnitude as one auxiliary analysis technique. On the other hand, traditional survey approaches, i.e., questionnaire survey or household surveys, are supposed to be executed at the next stage. They are more precise than drones' primary investigation. However, they require more surveying times, more investigators and higher costs than that of drones-based estimation. Thus, drones can be carried out at the initial stage of formulation, while traditional surveys further enhance their precision.

Given one candidate stop as the center, we manipulate the drones to investigate its circle region with its radius equal to 1000 m. Commonly, the formulation fields, prior to the real building, are '*unlimited height*' areas. It means that the drones can fly free without the negative impacts and control constraints. On the contrary, the '*height-permitted*' areas spur us to further survey by artificial approaches. In particular, the drone-based geospatial data survey needs to execute twice in one-round investigation for one identical demand area, i.e., in the day and at night. With regard to the day survey, aerial photos derived from drones can clearly demonstrate the accommodation capacity of residential quarters, factories, schools and so on. On the other hand, the night aerial photos can make us know the occupancy situations as per the lights being on or off. It should be noted that drone-based geospatial data investigation is for figuring out the order of magnitudes of the user demands rather than one kind of precise survey. Indeed, at the planning level, this precision level in knowing the primary demand amount is feasible at an early stage. These approximate data can be used as auxiliary data for proofreading.

Clearly, the drone-implemented investigation can not only save the survey duration, but also considerably reduce the human cost that is a significant expenditure component.

Drones were used to help with the pre-estimate passenger demand, which is one way for demand surveys. A drone-based geospatial data survey was executed by normal processes rather than relying on the particular hardware (cameras, zoom capability, etc.). Fair weather conditions are feasible for drone-based surveys. In this empirical investigation, the drones were purchased by the authors' institution, i.e., Nantong University, shown in Figure [2.](#page-9-0) Figure 2. Drones were used to help with the pre-estimate passenger demand, which is one way for Drones were used to neip with the pre-estimate passenger demand, which is one way for

but also considerably reduce the human cost that is a significant expenditure component.

<span id="page-9-0"></span>

**Figure 2.** Employed drones applied into the demand investigation. **Figure 2.** Employed drones applied into the demand investigation.

Typically, the PT operation schedule is divided per 1 h as the period unit. PT are Typically, the PT operation schedule is divided per 1 h as the period unit. PT are expected to operate between 05:00 and 21:00, which corresponds to 16-h intervals. We surveyed the number of floors for each building. The number of households for each floor relies on the manual interview. In general, the number of residents for one household assumed to be three. The geospatial information is recorded by the means of pictures is assumed to be three. The geospatial information is recorded by the means of pictures stored in drones, preliminarily obtaining the order of magnitudes of potential users. In stored in drones, preliminarily obtaining the order of magnitudes of potential users. In the planning level, either a hundred-unit or thousand-unit of demands are feasible precision for assisting a brand-new line formulation. The basic demand magnitude data derived from drones are settled as the third row of Table [2.](#page-11-0)









23



<u>51 54 710 7210 722 732 733 734 735 736 737 738 739 730 731 732 733 734 735 736 737 738 739 730 731 732 733 </u>

<span id="page-11-0"></span>**Table 2.** *Cont.*

<span id="page-11-1"></span>

**Figure 3.** Public transport line formulation scenario with stop candidates. **Figure 3.** Public transport line formulation scenario with stop candidates.

# *4.2. Public Transport Line Formulation Scenario 4.2. Public Transport Line Formulation Scenario*

28 candidate stops as the available infrastructure positions to determine, which are depicted by red circles. There are some demand points, depicted as the second row of Table [2.](#page-11-0) In general, residential/office buildings, factories, or schools are defined as the demand points. These demand points are investigated by the use of drones. They are too miscellaneous to intuitively identify, so no points are depicted in Figure 3 but rather [th](#page-11-1)e data of the demand points are described in Table [2.](#page-11-0) Commonly, the average walking speed is 1 m/s. Indeed, Nantong buses always have a low service sharing rate. Cars and electric bicycles account for the mot that second harger proportion for transport modes. Anagloong rimport projectric bi-<br>tends to facilitate passenger demand growth based on a reasonable stop formulation. Thus, cycles account for the first and second larger proportion for transport modes. Xingdong the case scenario is available to validate the effectiveness of the proposed model. Figure [3](#page-11-1) exhibits a planned bus line with a 30 km length in Nantong. The line includes for the first and second larger proportion for transport modes. Xingdong Airport project

Generally speaking, the costs of different items fluctuate subject to local economics, tax rates, or other real-life uncertainties. In this case, owing to the NDA (non-disclosure agreement) of this project, some cost values are derived from hypothetical estimation. In order to cover different budget situations, we propose three price-level budget expectations in Table [3.](#page-12-0) We name them as the high-price scenario, the medium-price scenario, and

Origin o

the low-price scenario; the capacities of one vehicle type are 20 pax, 30 pax, and 40 pax, respectively. The cost of building/maintaining one stop  $\Omega_1$  is 7300 US dollars. The travel cost of passengers  $\Omega_2$  is USD 13 per hour [\[2\]](#page-17-1). Yielding to the pure electric motor or alternatively diesel traction, the operation cost  $\Omega_3^v$  for vehicle type  $v$  is illustrated in Table [3,](#page-12-0) which is derived from full-life circle estimations. The passengers' OD distribution is demonstrated in Figure [4.](#page-12-1) For the sake of convenience, the OD matrix is recorded by distribution rate, in which the unit is percentage (%). For example, we focus on *OD*<sub>12</sub> equal to 1.84%. It means that for the overall 125 users (derived from  $45 + 30 + 50$ , in Table [2\)](#page-11-0) board from the first stop candidate, 1.84% of them are about to alight at the second stop candidate. The stop candidate set is  $I = \{i: i = 1, 2, 3, \ldots, 28\}$ . Clearly, with regards to the data from Table [2](#page-11-0) and Figure [4,](#page-12-1) simultaneously, the OD demand can be precisely calculated and applied into Equation (1).



<span id="page-12-0"></span>**Table 3.** Operation cost  $\Omega_3^v$  for vehicle type  $v$  in different price-based scenarios.

<span id="page-12-1"></span>

	$\mathbf{0}$	1.84	3.46	1.63	7.25	2.36	2.26	0.09	2.62	3.01	2.95	2.48	2.23	4.91	1.77	2.28	4.73	3.79	2.37	2.11	5.2	4.96	3.92	2.24	7.69	7.76	5.25	8.84	100
2	2.88	$\mathbf{0}$	5.52	4.96	2.94	2.89	5.81	5.6	7.04	4.46	2.52	5.88	1.55	6.82	1.95	1.93	8.08	1.07	5.42	1.17	3.55	3.65	1.01	3.78	1.13	1.84	0.2	9.23	
3	2.23	4.81	0	5.07	2.55	3.5	3.45	7.94	0.15	1.56	1.78	4.16	1.87	2.26	5.42	5.73	1.39	1.85	5.98	3.51	1.64	4.31	5.03	7.58	2.37	2.78	14.08	4.04	90
4	5.79	7.51	1.42	0	8.17	8.27	7.9	1.42	0.02	3.97	4.67	0.63	6.81	8.53	5.1	4.02	2.7	3.15	2.67	4.8	2.77	5.51	3.3	1.76	1.38	0.79	3.63	8.03	
5	6.17	1.99	6.95	0.68	$\bf{0}$	7.15	6.34	2.08	4.48	1.34	4.83	1.99	3.61	3.97	2.51	2.57	4.86	6.17	4.49	1.28	3.13	5.26	1.46	8.6	1.1	4.37	8.76	9.65	
6	1.97	4.79	3.62	3.08	2.14	$\bf{0}$	3.69	1.87	0.27	5.85	6.79	5.32	8.78	3.26	6.44	3.58	6.84	7.19	4.62	3.16	1.52	1.12	3.33	3.6	4.76	7.89	3.6	6.52	80
7	1.61	7.78	8.64	4.96	2.92	4.35	$\bf{0}$	$\overline{4}$	4.67	2.85	1.23	1.12	9.53	3.84	3.73	7.04	5.63	1.01	1.45	4.07	4.96	8.01	1.71	5.61	4.1	8.24	7.63	9.57	
8	5.18	5.42	1.71	4.23	1.73	6.57	1.58	0	5.15	6.31	3.06	3.59	1.64	4.43	2.94	5.01	5.65	3.74	4.31	4.94	6.55	5.53	5.62	6.14	3.72	3.56	8.72	9.39	
9	1.89	2.45	4.62	1.88	1.26	3.23	4.24	12.13	$\mathbf{0}$	3.9	3.03	2.28	4.44	5.74	7.36	7.8	2.16	1.66	7.39	3.71	3.69	5.43	1.63	11.02	4.91	8.03	8.4	7.42	70
10	1.3	1.75	1.87	2.47	8.22	5.48	5.67	5.84	2.23	$\circ$	2.42	3.54	1.45	6.55	3.61	7.69	5.76	3.5	8.79	7.58	7.54	4.75	7.8	8.25	5.85	7.04	2.64	5.24	
11	4.62	1.34	0.23	2.93	5.4	3.6	5.66	3.88	6.3	3.03	$\mathbf{0}$	2.48	8.03	4.77	3.83	4.26	9.48	2.52	3.14	4.48	8.03	4.37	3.38	10.19	6.25	8.79	5.03	10.97	
12	7.93	3.78	1.66	8.91	8.5	4.4	2.96	5.58	11.3	10.19	9.77	$\mathbf{0}$	5.96	4.88	6.11	1.54	7.08	4.5	6.17	2.52	5.69	9.65	4.82	13.12	11.02	11.17	4.16	1.61	60
13	1.55	3.37	4.94	5.27	1.53	6.39	2.64	5.82	4.83	1.27	10.96	4.49	$\circ$	11.09	6.51	13.54	6.22	2.27	10.08	2.92	9.9	3.39	8.25	5.45	8.12	4.49	2.73	5.04	
14	4.32	4.53	2.1	1.34	5.74	3.91	3.39	4.83	2.22	7.85	3.15	11.88	11.19	$\mathbf{0}$	13.72	13.62	12.78	13.97	4.41	2.24	7.26	5.12	5.79	2.35	2.93	6.4	7.4	2.01	50
15	2.88	3.02	3.55	7.25	2.51	5.31	3.14	3.03	4.31	10.58	2.57	2.92	5.17	4.75	0	2.51	5.85	7.02	3.35	11.62	11.07	4.93	11.82	11.46	4.91	10.58	11.47	3.41	
16	6.88	2.52	3.09	8.2	5.96	1.38	5.42	5.51	5.55	3.37	3.86	7.11	2.29	3.54	5.98	$\mathbf{0}$	11.32	5.17	3.05	12.08	12.47	11.71	12.81	3.27	12.46	5.47	10 <sup>1</sup>	0.19	
17	8.99	4.98	1.55	1.29	8.34	5.75	2.2	4.46	5.44	2.27	3.28	4.88	11.88	4.59	11.35	1.88	0	9.1	12.03	10.04	12.13	14.26	7.06	10.35	10.29	3.3	6.44	5	40
18	4.77	5.39	6.15	2.54	6.88	5.58	1.07	8.83	2.7	3.6	3.47	2.44	2.44	2.78	12.72	10.22	12.17	$\circ$	2.13	7.63	$11.12$   11.19		15.05	16.86	13.84	11.12	10.04	1.02	
19	3.39	4.81	4.17	0.75	1.33	2.3	4.4	10.28	11.19	4.21	4.34	4.12	11.18	10.72	4.37	4.5	13.05	13.56	0	15.4	7.29	18.07	10.43	9.17	11.93	5.62	12.11	9.98	
20	4.65	2.68	7.16	2.77	1.32	2.8	1.55	4.19	7.59	6.91	11.68	3.79	10.61	10.14	5.94	5.94	11.38	12.39	12.91	$\mathbf{0}$	12.49	11.11	11.95	11.76 17.66		10.84	12.64	11.55	30
21	1.61	4.3	6.33	4.96	2.95	5.61	5.64	2.21	5.47	2.16	4.91	3.86	7.5	10.13	5.23	12.27	14.59	13.13	11.57	5.13	$\bf{0}$	21.9	21.36	17.95	15.36	12.11	10.53	0.79	
22	4.77	2.86	0.86	3.63	4.15	1.53	6.43	2.55	3.57	2.98	11.19	8.77	3.66	11.55	2.97	12.83	11.02	13.15	11.03	12.23	12.53	0	28.69	26.39	23.28	17.81	3.35	0.48	
23	3.35	6.46	4.48	3.92	5.44	8.04	1.38	1.89	8.6	2.76	6.83	3.49	7.21	4.68	2.69	6.03	5.66	14.28	17.75	12.72	13.95	20.19	$\mathbf{0}$	20.37	16.63	20.69	20.66	21.65	20
24	1.73	4.56	3.02	7.61	5.28	5.5	1.64	1.54	3.07	6.26	5.78	10.9	4.01	12.58	10.58	5.44	4.66	12.45	10.44	13.71	16.86	21.18	26.4	$\circ$	35.1	21.81	21.14	21.95	
25	4.48	2.42	8.81	7.13	7.7	6.11	2.23	3.28	2.64	9.55	4.56	5.73	3.21	3.39	5.64	10.39	4.33	6.32	7.39	17.54	22.23	21.2	25.28	25.56	$\circ$	39.92	30.03	30.05	
26	1.99	3.02	3.87	6.18	4.31	7.53	5.69	2.25	3.27	7.95	4.05	5.29	3.27	6.44	8.45	11.65	3.27	3.69	5.89	13.58	11.65	18.26	20.77	20.18	32.18	$\mathbf{0}$	60.14	39.86	10
27	2.08	3.07	2.67	4.9	2.19	2.39	1.73	7.26	4.17	8.36	3.24	11.78	5.19	10.06	12.63	7.43	11.23	8.53	12.78	13.66	11.64	12.17		$11.23$   24.47	35.71	49.17	$\bf{0}$	100	
28	0.99	0.39	6.53	3.1	4.2	2.24	1.89	4.64	5.55	6.7	6.36	8.55	11.19	4.65		$11.45$   11.42	8.64	2.5	10.24	11.43	11.14	$\overline{7}$	16.32	29.79	32.11	50.83	100	0	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	
															Destination d														

**Figure 4.** Passengers' OD distribution rate from origin  $o$  ( $o \in O$ ) to destination  $d$  ( $d \in D$ ).

Subject to the budget for building the stops, we provide three scenario-based schemes regarding financial uncertainties. The budget hypotheses are given to depict different scenarios, that is, an overall USD 100,000 budget for the high-price scenario, USD 97,000 for the medium-price scenario, and USD 94,000 for the low-price scenario. The different budgets have a considerable impact on the number of selectable stops. This inference is clarified in the calculation results in Table [4.](#page-13-0)

<span id="page-13-0"></span>**Table 4.** Configurations of different stop schemes.



Note: The abbreviations of small, medium, and large vehicle type are S, M, and L, respectively.

# **4.3. Computation Results**

Given the linear programming problem, we adapt the IBM ILOG CPLEX MIP to handle this large-scale searching problem to derive the optimal scheme. This section executes the necessary numerical experiments based on a personal computer. Based on the preconditioning excavation, the computation is supposed to deal with 67,108,864 binary judgment decisions. As per global traversal, an optimized scheme is archived in less than 2 h running times. The gap between computation results for more than 10 times repeatedly is less than 0.5%, which is an acceptable calculation deviation. The three stop schemes are achieved and demonstrated in Figure 5. The numerical comparison is deployed in Tables [4](#page-13-0) and [5.](#page-14-0) Low-price scenario

<span id="page-13-1"></span>

**Figure 5.** Three scenario-based stop schemes. **Figure 5.** Three scenario-based stop schemes.



<span id="page-14-0"></span>**Table 5.** Computation results for different stop schemes.

Note: The abbreviations of small, medium, and large vehicle type are S, M, and L, respectively.

The results derived for determining the number and locations of stops as well as the number and types of vehicles are exhibited in Figure [5](#page-13-1) and Table [5.](#page-14-0) There are two dimensions to analyze the features of the solution schemes: (i) the regulated number of stops and (ii) the budget constraint. The former refers to high-quantity, medium-quantity, low-quantity stop schemes while the latter works out high-price, medium-price, low-price scenarios. Explicitly, the optimal scheme is the high-quantity stop scheme that possesses the greatest number of stops: 25 stops with eliminating 3 stop candidates. During the full-life cycle of the PT operation, the greater number of stops means the lower walking times. Although both the building cost of stops (USD 1606) and operation cost (USD 15,012) are the maximum values, the passengers' travel cost are a minimum, which is USD 71,383. Clearly, compared with the medium-quantity (USD 74,815), and low-quantity stop schemes (USD 83,109), the results also indicate that the high-quantity scheme achieves a 4.81% and 16.43% improvement in total passengers' travel cost, respectively. The fewer number of stops yield the more walk/travel cost, leading to the inferior results in medium- and low-quantity stop schemes, as expected.

On the other hand, given the overall budget threshold, the high-quantity stop scheme employed more big-type vehicles, whose frequency is 58 veh/day (about 31.82%) more than 44 veh/day in the other two schemes. Thus, its average load factor, i.e., 68.03%, is less than that of the medium-quantity stop scheme (70.97%), and low-quantity stop scheme (73.97%) with a 5.94% and 2.94% reduction. Yielding to the budget constraint, in high-price, medium-price, and low-price scenarios, the optimal high-quantity stop scheme can save 3.04%, 3.11%, and 3.38% in overall costs compared with the medium-quantity stop scheme, respectively; their cost savings are 8.53%, 8.70%, and 9.09% more than the costs of the low-quantity stop scheme.

Subject to the budget for building the stops, we provide three scenario-based schemes regarding financial uncertainties. It is assumed that high, medium, and low budgets refer to USD 100,000, USD 97,000, and USD 94,000, respectively. This hypothesis can be changed as the parameter input according to the real economy finances. The predetermined budgets play a key role in the number of stops. The case results indicate that 25, 22, and 18 stops are allowed to be built for the three budgets. During the full-life cycle of the PT operation, the greater number of stops means lower walking times, which indeed facilitates the public transport service to be more attractive. Thus, real equilibrium between budget investment and users' travel cost determines the optimal trade-off for supply and demand.

The results demonstrate that the increase in passenger travel time is approximately 1% according to reducing one stop location candidate. The total travel times of users is closely related to the number of stops to build. Thus, there is a trade-off for the policymakers to consider between travel efficiency and service accessibility. Herein, the hard constraint is derived from the budget. Along with the increasing number of passengers, the majority of users' travel cost during the long-term operation horizon, more stops (if the budget is available) are expected to achieve the service accessibility for the local residents for a sustainable PT career. Even though the incumbent PT company stakeholders are anticipated to incur a high expense, from the perspective of users, they benefit from the PT service convenience (i.e., the short walking distance) over the long haul.

# <span id="page-15-0"></span>**5. Conclusions**

# *5.1. Findings*

In this paper, we developed a mixed-integer linear programming (MILP) model for facilitating an optimal strategic decision involved in the launch of a new public transport (PT) line. The formulation problem covered the number and the locations of stops, and the vehicle types, even though the two problems were often independent in the literature. Our study attained their consistency in a uniform objective function to minimize the passengers' travel costs, stop building costs, and operation costs. Explicitly, yielding to different budget scenarios, the relationship between the two main stakeholders inspired us to seek out the high-quantity, medium-quantity, and low-quantity stop schemes in high-price, mediumprice, and low-price scenarios. Oriented by a real project, this study generated and analyzed a multi-dimensions budget-based stop location plan.

The case results demonstrated the effectiveness of CPLEX solver and the proposed MILP model. One study challenge we combated was how to build a programming model according to the passenger demands and transport resources; simultaneously, another challenge was what solution method to handle with this model. In an actual urban circumstance, public transport line formulation yields urban master planning. Our study provided a line scheme candidate set to urban designers. Considering the existing transportation systems, the new PT mode shared a reasonable percentage for all-mode travels. The percentage was derived from the empirical survey, as a new task. The socio-economic impact analysis/estimation was a valuable insight, which could help in reducing congestion, facilitating low carbon, saving travel cost, and so on. It was a significant concept which was considered in future studies. New PT line formulation in other regions was expected to share a compatible modelling approach for this study.

Drones were used to investigate the probable demand magnitude for assisting the PT passenger flow forecast. Then, a real case in Nantong, China was performed to testify the effectiveness of the model and CPLEX. The calculation results indicated the CPLEX solver availably handled the proposed MILP model within a reasonable commutating time. It is worth noting that the formulation model innovatively integrated with stop planning and multi-type vehicle frequency optimization together. The outcomes reproduced as per the model, i.e., the number and the locations of stops, the service frequency in hourly operation, and the three sub-objective function values were useful estimation criteria for decision makers. In light of different stop schemes that apply to different budget scenarios, the PT operation company was able to make system-optimal decisions about purchasing the multi-type fleet sizes, determining the locations of stops, and number.

Planning activities and practical operations interacted with each other. The practical operation effectiveness could be reflected at the level of service, i.e., the waiting times of passengers, the number of left-behind users, and the load rates. This direct feedback would motivate the service frequency adjustment and future formulation.

## *5.2. Future Work and Discussions*

In the real world, the user demand is dynamic rather than fixed. However, the land status (i.e., land use property) is determined a priori. We can know an approximate order of magnitudes for potential passengers by investigating the data from the area of lands, plot ratio, and the number/property of buildings. We are focusing on a brand-new public transport line formulation, so the real passenger flow has not been generated yet. Essentially, the current passenger flow demand survey is a kind of forecast. In fact, we cannot assert the passenger demand is 100% accurate. However, we propose a good way: drones-based investigation, so as to enhance the precision of the survey/forecast. On the other hand, the contribution of this study is a 'urban public transport line formulation model'. The user demand data are the input of the model. In fact, the proposed model/solution is proven valid based on the input data. In future work, we can pay more attention to improving the forecast precision further.

The model that was built was not special for public transport line structure, vehicle type (electric or fuel-driven), city size, and so on. Thus, for a new public transport line formulation in a medium-sized city, the model and solution is well matched, because Nantong, China belongs to a medium-sized city. Without testing as per other cities' data, it is uncertain to assess how adaptable our model would be, especially for megacities. From the authors' perspective, the detailed information of the different land modules as the demand points (or the so-called interest points) should be principally known. An empirical survey strategy is necessary to carry out. A feasible real survey that needs fulfilling before the formulation provides a data estimation basis for the PT passenger demand. It can facilitate an optimal formulation solution by the means of the proposed model.

## *5.3. Limitations*

With regards to the limitations from our incumbent work developed, future work will focus on the following aspects to enrich the programming study.

- Limitation (i): A PT hub or transfer stop allows the feeder accessibility to connect aircraft terminals and high-speed rail stations. Thus, multi-type PT stops and multitype PT vehicles would be taken into consideration in the study plan.
- Limitation (ii): Our study aimed to determine the reasonable number and locations of stops along a fixed bus route. To be clear, PT routing does not belong in this work. The current formulating stage is that only the routing trend is determined and the locations/number of stops are not yet determined. A line formulation study has the potential to extend to a network level plan.
- Limitation (iii): Multi-type and multi-depot vehicle scheduling integration optimization is not taken into account. In addition, for the proposed MILP model, as the scale of the resolving problem increases, the computation efficiency tends to be lower.
- Limitation (iv): Robustness and flexibleness are not taken into consideration for the study. A special circumstance (such as COVID-19 pandemic) impact was not observed for the current research work. The model proposed in the study is a generalized approach for ordinary operation in the environment. The robustness study for special events will be considered in the next paper.

**Author Contributions:** Conceptualization, S.Z. (Silin Zhang), B.Z., Y.Z., S.Z. (Shun Zhang) and Z.C.; methodology, S.Z. (Silin Zhang) and Z.C.; software, S.Z. (Silin Zhang) and Z.C.; validation, S.Z. (Silin Zhang), B.Z., Y.Z., S.Z. (Shun Zhang) and Z.C.; formal analysis, S.Z. (Silin Zhang), B.Z., Y.Z. and Z.C.; investigation, S.Z. (Silin Zhang), B.Z., Y.Z., S.Z. (Shun Zhang) and Z.C.; resources, S.Z. (Silin Zhang), B.Z., Y.Z., S.Z. (Shun Zhang) and Z.C.; data curation, S.Z. (Silin Zhang), B.Z., Y.Z., S.Z. (Shun Zhang) and Z.C.; writing—original draft preparation, S.Z. (Silin Zhang) and Z.C.; writing—review and editing, S.Z. (Silin Zhang) and Z.C.; visualization, S.Z. (Silin Zhang), B.Z., Y.Z., S.Z. (Shun Zhang) and Z.C.; supervision, S.Z. (Silin Zhang), B.Z., Y.Z., S.Z. (Shun Zhang) and Z.C.; project administration, S.Z. (Silin Zhang), B.Z., Y.Z. and Z.C.; funding acquisition, S.Z. (Silin Zhang), B.Z., Y.Z. and Z.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by National Natural Science Foundation of China (72101126) and (72101127), by Hubei Key Laboratory of Transportation Internet of Things open funding, Wuhan University of Technology (WHUT-IOT-001), and by Opening Foundation of Zhejiang Intelligent Transportation Engineering Technology Research Center (2023ERCITZJ-KF14).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Due to signed confidentiality agreements, the dataset is not publicly available. It can only be accessed under conditional terms by contacting the authors when there is a scholarly need.

**Conflicts of Interest:** Buhao Zhang is from Zhejiang Intelligent Transportation Engineering Tech-**Conflicts of Interest:** Buhao Zhang is from Zhejiang Intelligent Transportation Engineering Technology Research Center. Yi Zhao is from China Academy of Railway Sciences Corporation Limited. nology Research Center. Yi Zhao is from China Academy of Railway Sciences Corporation Limited. Shun Zhang is from Jiangsu CRRC Mechanical & Electric Technology Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

available. It can only be accessed under conditional terms by conditional terms by contacting the authors when the auth

# <span id="page-17-15"></span>**Appendix A**

The appearance and interior of three vehicle types are exhibited in Figure A1. The appearance and interior of three vehicle types are exhi[bite](#page-17-14)d in Figure A1.

<span id="page-17-14"></span>(i) The appearance of Type 1 (ii) The appearance of Type 2 (iii) The appearance of Type 3

(iv) The interior of Type 1

(v) The interior of Type 2

(vi) The interior of Type 3

**Figure A1.** Stop planning problem formulation scenario to accommodate surrounding demands. **Figure A1.** Stop planning problem formulation scenario to accommodate surrounding demands.

# **References**

- <span id="page-17-0"></span>1. Ceder, A. Syncing sustainable urban mobility with public transit policy trends based on global data analysis. *Sci. Rep.* **2021**, *11*, 1. Ceder, A. Syncing sustainable urban mobility with public transit policy trends based on global data analysis. *Sci. Rep.* **2021**, *11*, 14597. 14597. [\[CrossRef\]](https://doi.org/10.1038/s41598-021-93741-4)
- <span id="page-17-1"></span>2. Yang, L.; Qi, J.; Li, S.; Gao, Y. Collaborative optimization for train scheduling and train stop planning on high-speed railways. *Omega* **2016**, *64*, 57–76. *Omega* **2016**, *64*, 57–76. [\[CrossRef\]](https://doi.org/10.1016/j.omega.2015.11.003)
- <span id="page-17-2"></span>3. Repolho, H.M.; Church, R.L.; Antunes, A.P. Optimizing station location and fleet composition for a high-speed rail line. *Res. Part E Logist. Transp. Rev.* **2016**, *93*, 437–452. *Transp. Res. Part E Logist. Transp. Rev.* **2016**, *93*, 437–452. [\[CrossRef\]](https://doi.org/10.1016/j.tre.2016.06.006)
- <span id="page-17-3"></span>4. Yuan, Y.; Li, S.; Liu, R.; Yang, L.; Gao, Z. Decomposition and approximate dynamic programming approach to optimization of train timetable and skip-stop plan for metro networks. Transp. Res. Part C Emerg. Technol. 2023, 157, 104393. [\[CrossRef\]](https://doi.org/10.1016/j.trc.2023.104393)
- <span id="page-17-4"></span>5. Chang, Y.H.; Yeh, C.H.; Shen, C.C. A multiobjective model for passenger train services planning: Application to Taiwan's speed rail line. *Transp. Res. Part B Methodol.* **2000**, *34*, 91–106. high-speed rail line. *Transp. Res. Part B Methodol.* **2000**, *34*, 91–106. [\[CrossRef\]](https://doi.org/10.1016/S0191-2615(99)00013-2)
- <span id="page-17-5"></span>6. Tang, C.; Shi, H.; Liu, T. Optimization of single-line electric bus scheduling with skip-stop operation. Transp. Res. Part D Transp. *Environ.* **2023**, *117*, 103652. *Environ.* **2023**, *117*, 103652. [\[CrossRef\]](https://doi.org/10.1016/j.trd.2023.103652)
- <span id="page-17-6"></span>7. Cacchiani, V.; Qi, J.; Yang, L. Robust optimization models for integrated train stop planning and timetabling with passenger demand uncertainty. *Transp. Res. Part B Methodol.* **2020**, *136*, 1–29. [\[CrossRef\]](https://doi.org/10.1016/j.trb.2020.03.009)
- <span id="page-17-7"></span>8. Shao, J.; Xu, Y.; Sun, L.; Kong, D.; Lu, H. Equity-oriented integrated optimization of train timetable and stop plans for suburban railways system. *Comput. Ind. Eng.* **2022**, *173*, 108721. [\[CrossRef\]](https://doi.org/10.1016/j.cie.2022.108721)
- <span id="page-17-8"></span>9. Canca, D.; Barrena, E.; De-Los-Santos, A.; Andrade-Pineda, J.L. Setting lines frequency and capacity in dense railway rapid transit networks with simultaneous passenger assignment. *Transp. Res. Part B Methodol.* **2016**, *93*, 251–267. [\[CrossRef\]](https://doi.org/10.1016/j.trb.2016.07.020)
- <span id="page-17-9"></span>10. Shi, J.; Yang, J.; Yang, L.; Tao, L.; Qiang, S.; Di, Z.; Guo, J. Safety-oriented train timetabling and stop planning with time-varying and elastic demand on overcrowded commuter metro lines. *Transp. Res. Part E Logist. Transp. Rev.* **2023**, *175*, 103136. [\[CrossRef\]](https://doi.org/10.1016/j.tre.2023.103136)
- <span id="page-17-10"></span>11. De Weert, Y.; Gkiotsalitis, K. A COVID-19 public transport frequency setting model that includes short-turning options. *Future Transp.* **2021**, *1*, 3–20. [\[CrossRef\]](https://doi.org/10.3390/futuretransp1010002)
- <span id="page-17-11"></span>12. Liang, S.; He, S.; Zhang, H.; Ma, M. Optimal holding time calculation algorithm to improve the reliability of high frequency bus route considering the bus capacity constraint. *Reliab. Eng. Syst. Saf.* **2021**, *212*, 107632. [\[CrossRef\]](https://doi.org/10.1016/j.ress.2021.107632)
- <span id="page-17-12"></span>13. Sadrani, M.; Tirachini, A.; Antoniou, C. Optimization of service frequency and vehicle size for automated bus systems with crowding externalities and travel time stochasticity. *Transp. Res. Part C Emerg. Technol.* **2022**, *143*, 103793. [\[CrossRef\]](https://doi.org/10.1016/j.trc.2022.103793)
- <span id="page-17-13"></span>14. Fei, F.; Sun, W.; Iacobucci, R.; Schmöcker, J.D. Exploring the profitability of using electric bus fleets for transport and power grid services. *Transp. Res. Part C Emerg. Technol.* **2023**, *149*, 104060. [\[CrossRef\]](https://doi.org/10.1016/j.trc.2023.104060)
- <span id="page-18-0"></span>15. Tong, P.; Du, W.; Yan, Y.; Li, J. Quantifying Bus Accessibility and Mobility for Urban Branches: A Reliability Modeling Approach. *Sustainability* **2023**, *15*, 15770. [\[CrossRef\]](https://doi.org/10.3390/su152215770)
- <span id="page-18-1"></span>16. Kim, M.; Kim, E. Joint Optimization of Distance-Based Fares and Headway for Fixed-Route Bus Operations. *Sustainability* **2023**, *15*, 15352. [\[CrossRef\]](https://doi.org/10.3390/su152115352)
- <span id="page-18-2"></span>17. Yang, H.; Liang, Y. Examining the Connectivity between Urban Rail Transport and Regular Bus Transport. *Sustainability* **2023**, *15*, 7644. [\[CrossRef\]](https://doi.org/10.3390/su15097644)
- <span id="page-18-3"></span>18. Nadimi, N.; Zamzam, A.; Litman, T. University Bus Services: Responding to Students' Travel Demands? *Sustainability* **2023**, *15*, 8921. [\[CrossRef\]](https://doi.org/10.3390/su15118921)
- <span id="page-18-4"></span>19. Li, X.; Yang, Z.; Lian, F. Optimizing On-Demand Bus Services for Remote Areas. *Sustainability* **2023**, *15*, 7264. [\[CrossRef\]](https://doi.org/10.3390/su15097264)
- <span id="page-18-5"></span>20. Risso, C.; Nesmachnow, S.; Faller, G. Optimized Design of a Backbone Network for Public Transportation in Montevideo, Uruguay. *Sustainability* **2023**, *15*, 16402. [\[CrossRef\]](https://doi.org/10.3390/su152316402)
- <span id="page-18-6"></span>21. Su, H.; Li, M.; Zhong, X.; Zhang, K.; Wang, J. Estimating Public Transportation Accessibility in Metropolitan Areas: A Case Study and Comparative Analysis. *Sustainability* **2023**, *15*, 12873. [\[CrossRef\]](https://doi.org/10.3390/su151712873)

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