Research on Dynamic Ride and Drop-off Site Setting for Customized Passenger Transport Based on Spatial-temporal Clustering

Jing Yuan, Xiaojing Li, Yihao Qin, Yijian Li, and Bojun Yang

***Abstract*—Traditional customized passenger transport (CPT) systems often rely on fixed stops derived from historical trip data and struggle to accommodate dispersed spatial-temporal demands. This paper presents a novel method for setting dynamic ride and drop-off sites based on spatial-temporal clustering. Firstly, a three-objective model is constructed, aimed at minimizing deviations between actual and expected locations and times of these sites while also reducing their overall number. Secondly, we design a refined NSGA-III algorithm that integrates the ST_DBSCAN algorithm to achieve Pareto optimal solutions for dynamic ride and drop-off sites, incorporating capacity and time considerations. Finally, experiments conducted using CPT data from Chongqing to Tongliang District, China, indicate that the refined algorithm effectively reduces the noise point ratio of the ST_DBSCAN algorithm to 5% and identifies 89 ride and drop-off sites across four categories. It also decreases the deviations of actual riding time and location from expectations to 12.97% and 12.62%, respectively. Additionally, employing CPLEX to solve the first type of dynamic ride and drop-off sites results in a 4.7% reduction in computation time. Compared to the unrefined algorithm, the total time for setting dynamic sites is shortened by 30.02% while optimizing the objective values. The refined algorithm exhibits high convergence efficiency, particularly with large-scale data.**

***Index Terms*—Customized passenger transport, dynamic ride and drop-off site, three-objective optimization model, spatial-temporal clustering, NSGA-III algorithm**

I. Introduction

At the dawn of the 20th century, road travel was predominantly reliant on public transit and private cars, with personalized travel options being virtually non-existent. The advent of mobile Internet technology in the 1990s ushered in a new era for on-demand travel services, leading to the emergence of customized passenger transport (CPT), customized buses (CB), flexible-route buses (FRB), and demand responsive transit (DRT). These innovations have significantly transformed personal mobility [1]. According to Armando [2], demand responsive transportation is particularly effective in low-density areas, has shown promise in suburban and inter-regional contexts. Within this context, CPT emerges as a pioneering model of personalized transport, leveraging Internet technology to align passenger travel demands with available capacity resources. By accounting for individual preferences regarding journey origins, destinations, and desired stops, CPT identifies optimal ride and drop-off locations, thus providing a regional demand responsive service that caters to the flexible and convenient travel preferences of modern passengers [3].

The strategic placement of CPT ride and drop-off sites is critical, impacting both passenger convenience and the economic viability of transport operators. When passenger demand is dispersed across time and space, transport companies must employ a flexible and comprehensive system for dynamically establishing these sites. Recent advancements in CB have enhanced service quality in urban areas, offering differentiated and high-quality services. While CPT provides agile, small-scale, and personalized road transport services [4], it shares similarities with CB. The key distinction lies in their customization approaches: CB employs a three-step process involving passenger demand surveys, route recruitment, and booking for route personalization [5]. In contrast, CPT bases service station locations solely on traveler demand, reflecting specific passenger needs.

Despite the growing importance of dynamic ride and drop-off site setting for CPT (CPT-DRDSS), this area remains under-researched. Li [6] laid the groundwork by utilizing historical travel data to determine customized passenger ride and drop-off sites, but this data primarily reflects temporal and spatial patterns of passenger travel. In the context of site setting for CB services, several notable studies have been conducted. Seyed [7] focused on identifying suitable bus stations based on students' decision-making for school bus routes, while Ren [8] identified candidate bus stops by mapping the relationship between demand locations and existing bus stops in demand-responsive scheduling. However, methods reliant on individual or enterprise decision-making often succumb to

Manuscript received December 10, 2024; revised April 21, 2025.

This research was funded by the National Natural Science Foundation of China Grant No. 71861023, the Key Research and Development Program Project of the Department of Science and Technology of Gansu Province No. 23YFFA0058.

Jing Yuan is a postgraduate student at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (e-mail: 2075827855@qq.com).

Xiaojing Li is an associate professor at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (Corresponding author, phone: +86 13389404962, e-mail: lixjing@mail.lzjtu.cn).

Yihao Qin is a postgraduate student at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (e-mail: 1733879683@qq.com).

Yijian Li is a postgraduate student at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (e-mail: 18707210692@163.com).

Bojun Yang is a postgraduate student at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (e-mail: 1296642023@qq.com).

irrational influences. Cortenbach [9] proposed a methodology for setting ride and drop-off sites that aims to reduce operating costs in route planning for the Dial-a-Ride problem, leveraging route intersections to establish potential pick-up and drop-off points. However, this approach, which considers the actual road network, adds complexity to the planning process. Thus, further empirical investigation into dynamic site setting within CPT is essential to enhance effectiveness and reliability.

Zheng [10] developed a service unit generation model to determine riding locations and times for CB route planning. However, once the service area is defined, passengers at the boundaries may experience service loss. To better accommodate the personalized travel needs of CB passengers, Li [11] and Qiu [12] focused on optimizing individual travel experiences without clustering passenger demands, which can result in excessive stops, decreased comfort, and inefficient operations. In response, some researchers have employed clustering algorithms to group passenger travel demands, alleviating the spatial constraints of original ride and drop-off sites. For example, studies utilizing the K-means algorithm have spatially clustered similar origin-destination (OD) demands, with Lei [13] and Shen [14] designating cluster centers as ride and drop-off locations. Ayman [15] used clustering results as initial solutions for the mixed-integer programming (MIP) site coverage problem, subsequently designing an algorithm to identify optimal locations. Hu [16] and Yu [17] applied the DBSCAN algorithm to detect demand noise points, followed by K-means to establish ride and drop-off sites. Xue [18] and Hu [19] employed a hierarchical clustering algorithm to group reservation sites based on similar travel times, using DBSCAN to eliminate spatially isolated sites. Li [20] combined DBSCAN and the Partitioning Around Medoids (PAM) algorithms to spatially and temporally cluster passenger travel demands, generating synthetic stations for dynamic routes. Yuan [21] constructed a recognition model using a graph semi-supervised learning algorithm to identify potential CB demand from labeled internet taxi data and unlabeled cab data, clustering results with DBSCAN and further subdividing them to create CB stops based on reachable walking distances.

Despite these advancements, the determination of CB stations often overlooks the integrated spatial-temporal attributes of reservation demands. Hu [22] proposed an improved BIRCH clustering algorithm based on large-scale historical carpooling data, addressing challenges in big data clustering. However, limitations remain when clustering three-dimensional spherical data. Wang [23] introduced an improved ST_DBSCAN algorithm for clustering spatial-temporal data, but it does not directly address noise sites, potentially leading to inadequate analysis of spatial-temporal dispersed data. Erdmann [24] explored response strategies for on-demand travel services, emphasizing heuristic-based timely responses and global re-optimization. Zhang [25] proposed an optimization framework for dynamic ride and drop-off sites based on real-time travel demand, aiming to enhance the operational efficiency of flexible route buses. This approach seeks to maximize the fulfillment of passenger requests while minimizing time expenditure, transcending the spatial constraints of traditional ride and drop-off sites and significantly reducing the rate of rejected passenger bookings without incurring additional operational costs. Lu [26] addressed the issue of passengers lacking suitable travel options by eliminating bookings for those with viable alternatives, ensuring that dial-a-ride services effectively target individuals who truly need them. He highlighted the implementation of an active request rejection mechanism with dynamic thresholds to prevent system overloads in DRT systems. This proactive approach helps maintain service quality and safeguards against the degradation of standards that can occur with excessive demand.

When constructing models for passenger station siting, two distinct perspectives emerge: passenger-centric models prioritize minimizing walking distances [27], while enterprise-focused models aim to optimize total route length and maximize profits [28]. These conflicting objectives complicate multi-objective optimization efforts. Cui [29] introduced the NSGA-II algorithm, which utilizes population average distance clustering to select the best individuals during the selection phase. However, this approach encounters reduced computational efficiency as population sizes increase.

In conclusion, most existing studies primarily focus on CB stop configurations, personalized passenger routes, and CB route planning, with relatively little attention devoted to dynamic ride and drop-off sites for personalized passenger services. Additionally, many studies neglect actual passenger booking demands and fail to jointly consider the spatial and temporal characteristics of OD pairs and booking demands. This oversight results in insufficient consideration of both enterprise costs and passenger satisfaction. Furthermore, the algorithms used often exhibit inefficiencies when dealing with large-scale datasets.

This paper aims to address these gaps by incorporating multi-attribute spatial-temporal demands of passengers and developing a multi-objective model for CPT-DRDSS. Spatial-temporal clustering techniques are employed to aggregate passenger booking needs into clusters, thereby reducing the proportion of noise sites and ensuring that the needs of all passengers are met. Furthermore, the individual initialization phase of the NSGA-III algorithm is enhanced through the integration of the ST_DBSCAN algorithm, facilitating effective model solutions. The effectiveness of this approach is validated using Point-of-Interest (POI) data within the service area of a customized passenger line, serving as a case study.

II. DESCRIPTION OF THE PROBLEM

This study systematically introduces the concept of CPT before addressing the associated issues. CPT is a passenger service that utilizes online ticketing provided by shuttle bus operators through e-commerce platforms. These platforms offer detailed information on passenger transport routes, allowing for flexible agreements on departure times and stopping places. The primary objective of this model is to deliver tailored transportation solutions that cater to the increasingly diverse and personalized travel needs of passengers. Given the complexity of high passenger reservation demand and significant spatial and temporal

TABLE I
RESERVATION REQUEST INFORMATION

No.	Ride location (Lon, Lat)	Drop location (Lon, Lat)	Expected riding time
1	(x_1, y_1)	(X_1, Y_1)	T_1
2	(x_2, y_2)	(X_2, Y_2)	T_2
…	…	…	…
i	(x_i, y_i)	(X_i, Y_i)	T_i

variations in ride and drop-off locations, this study analyzes travel demand for CPT from urban centers to surrounding counties. It identifies a subset of dynamic ride and drop-off sites characterized by capacity and spatiotemporal attributes from a broad pool of demands.

Within this framework, it is assumed that passengers submit their reservation details via a CPT information service platform. Based on the provided reservation demand information, the platform efficiently transports passengers from designated locations in the departure area to specified locations in the destination area. The CPT platform collects essential demand information, including the reservation site number, ride location (longitude and latitude), drop-off location (longitude and latitude), and expected riding time, as detailed in TABLE I.

To account for the multitude of complex factors influencing the setting of dynamic ride and drop-off sites in practical scenarios, this study incorporates several assumptions regarding the setting of dynamic ride and drop-off sites:

1) The setup cost for each dynamic ride and drop-off site is assumed to be uniform.

2) Passenger reservation demands are considered independent and are assumed to remain constant.

3) Travel conditions are deemed to be accessible at every reservation demand site as well as at the potential ride and drop-off sites.

4) The distance between the reservation site and the dynamic site is calculated using Euclidean distance, without accounting for potential detours that passengers may encounter while traveling to the dynamic site.

III. MODELING OF THE CPT-DRDSS PROBLEM

Given the importance of minimizing deviations between actual and desired ride and drop-off locations, as well as between actual and expected riding times, effective implementation of door-to-door services in CPT is critical. Additionally, the number of dynamic stop settings impacts not only the operational efficiency of CPT companies but also, if excessive, can degrade the passenger travel experience.

To address these challenges, this paper establishes a multi-objective optimization model for CPT-DRDSS. The model notations are summarized in TABLE II.

$$\min Z_1 = \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij} \times x_{ij} \tag{1}$$

$$\min Z_2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \left|T_j - T_i\right| \times x_{ij} \tag{2}$$

$$\min Z_3 = \sum_{i=1}^{n} y_i \tag{3}$$

$$s.t.\ \ 1 \le \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} \le n \tag{4}$$

$$\sum_{i=1}^{n} x_{ij} = 1,\ j = 1, 2, \ldots, n \tag{5}$$

$$d_{ij} \times x_{ij} \le \mathrm{R},\ i = j = 1, 2, \ldots, n \tag{6}$$

$$\left(T_j - T_i\right)^2 \times x_{ij} \le \mathrm{T}^2,\ i = j = 1, 2, \ldots, n \tag{7}$$

Equations (1) - (3) define the objective functions of our model. Specifically, equation (1) minimizes the distance deviation, while equation (2) minimizes the time deviation, respectively; both are critical indicators for maximizing passenger satisfaction. Equation (3) aims to minimize the number of vehicle stops, which is a key factor in reducing operational costs for the enterprise.

The constraints are detailed in equations (4) - (7). Equation (4) specifies the range for the number of dynamic sites. Equation (5) ensures that each reservation site *j* is associated with exactly one dynamic site *i*. Equation (6) guarantees that at least one demand site should be covered within the service

TABLE II
RELATED SYMBOLS

	Symbol	Definition
Decision variable	x_{ij}	0–1 decision variable, $x_{ij} = 1$ indicates that potential ride and drop-off site *i* covers reservation site *j*, otherwise it is 0
	y_i	0–1 decision variable, $y_i = 1$ indicates that potential ride and drop-off site *i* is selected as a dynamic ride and drop-off site, otherwise it is 0
Intermediate variables and other parameters	n	Number of reservation sites or potential ride drop-off sites
	R	Maximum distance deviation acceptable to passengers
	T	Maximum time deviation acceptable to passengers
	T_i	Time attributes of potential ride site *i*
	T_j	Time attributes of reservation ride site *j*
	d_{ij}	Distance from potential ride site *i* to reservation site *j*
	Z_1	Total distance from each reservation site to the corresponding dynamic sites
	Z_2	Total time deviation from each reservation sites to the dynamic sites
	Z_3	Number of dynamic site settings

radius R of any dynamic site. Finally, equation (7) stipulates that the deviation of actual departure time from the expected time must not exceed a predefined threshold T.

IV. Model Solution

A. Algorithm Design

The location selection problem is fundamentally NP-hard. Research has demonstrated that the NSGA-III algorithm consistently outperforms MOEA/D algorithms across a range of problems with objectives varying from 3 to 15 [29]. In scenarios where minimizing objective Z_1 may result in an increase in objectives Z_2 and Z_3, this paper employs the NSGA-III algorithm to solve the constructed three-objective model for CPT-DRDSS.

In the realm of clustering algorithms, the ST_DBSCAN algorithm extends DBSCAN by adding an additional dimension, transforming the two-dimensional circular search area into a three-dimensional spherical search area. This enables density-based spatial-temporal clustering of spatial point data with temporal attributes [30]. For CPT demands that are temporally and spatially dispersed and involve large volumes of data, it is challenging to cluster them into groups with similar attributes using a single criterion [31]. Such difficulties can lead to an increased number of rejected passenger reservation requests.

Algorithm 1 Generation *t* of NSGA-III procedure

Input: NSGA-III related parameters, including *MaxGen*, *Pm*, *Pc*, and *H* (Structured reference sites Z^s)
ST_DBSCAN related parameters, including dataset D_1, and clustering parameter Eps_1, Eps_2, Eps_3, *MinPts, and* $\Delta\varepsilon$

Output: P_{t+1}

1. $S_t = \emptyset$, $i = 1$
2. When $t = 1$:
3. Three rounds of ST_DBSCAN procedure
4. $P_t = H$ individual initialization + Population initialization
5. $(F_1, F_2, \ldots)$ = Non-dominated-sort (P_t)
6. Q_t = Selection + Crossover + Mutation (P_t)
7. When $t > 1$:
8. Q_t = Selection + Crossover + Mutation (P_t)
9. $R_t = P_t \cup Q_t$
10. $(F_1, F_2, \ldots)$ = Non-dominated-sort (R_t)
11. **repeat**
12. $S_t = S_t \cup F_i$ and $i = i + 1$
13. **until** $|S_t| \geq H$
14. Last front to be included: $F_l = F_i$
15. **if** $|S_t| = H$ **then**
16. $P_{t+1} = S_t$, break
17. **else**
18. $P_{t+1} = \cup_{j=l}^{l-l} F_j$
19. Sites to be chosen from F_l: $K = H - |P_{t+1}|$
20. Normalize objectives and create reference set Z_r
21. Associate each member *s* of S_1 with a reference site
22. Compute niche count of reference site
23. Choose *K* members one at a time from F_l to construct P_{t+1}
24. **end if**

To tackle the challenges associated with data scale and noise points identified by density-based clustering algorithms, this study employs the ST_DBSCAN algorithm in three iterations during the individual initialization phase of the NSGA-III algorithm. This process clusters the mapping relationships between potential dynamic ride and drop-off sites (core objects) and reservation sites (cluster points) extracted from the reservation demand data, effectively dividing them into smaller, distinct sets of ride and drop-off sites. This categorization streamlines individual encoding, reduces chromosome length, and enhances the performance of the NSGA-III algorithm. The detailed process is outlined in Algorithm 1. Here, D_1 represents the dataset of passenger reservation demands, while D_2 and D_3 correspond to the noise sets derived from the first and second rounds of spatial-temporal clustering, respectively, and are consistent with the data type of D_1.

B. Three Rounds of ST_DBSCAN Algorithm

This study employs the ST_DBSCAN algorithm, which is characterized by its multidimensional approach to evaluating both spatial and non-spatial similarities within datasets. Specifically, the algorithm defines the Eps_1 neighborhood for reservation sites in the departure area and the Eps_2 neighborhood for those in the arrival area to quantify spatial similarity. Non-spatial similarity, based on the temporal dimension, is assessed through the Eps_3 neighborhood associated with departure times.

The density parameter *MinPts* establishes the minimum number of points required within the neighborhood of a core object for it to be classified as such. The parameter $\Delta\varepsilon$ is crucial for fragmenting clusters that are spatially contiguous but exhibit significant disparities in non-spatial attributes. Additionally, the algorithm employs a stack-based iterative methodology to aggregate all density-reachable data points from the core objects, ensuring a comprehensive and robust clustering process.

1) Initial Spatial-Temporal Clustering

The initial spatial-temporal clustering focuses on the geographic locations of ride and drop-off sites, as well as the departure times, considering the joint origins and destinations. This process aggregates passengers whose ride and drop-off locations, along with their departure times, are in close geographic and temporal proximity. Following this procedure, the reservation information $(x_{1,2,3,\ldots}, y_{1,2,3,\ldots}, x'_{1,2,3,\ldots}, y'_{1,2,3,\ldots}, T_{1,2,3,\ldots})$ undergoes a clustering process, resulting in the formation of distinct clusters denoted as $(x_{\mathrm{I},\ldots}, y_{\mathrm{I},\ldots}, x'_{\mathrm{I},\ldots}, y'_{\mathrm{I},\ldots}, T_{\mathrm{I},\ldots})$, along with a set of noise points. In this context, identifiers such as 1, 2, 3, etc., represent the reservation sites, while I, etc., denote the cluster identifiers assigned post-clustering. The attributes of the noise points maintain their status as reservation information. The specific procedures for the initial spatial-temporal clustering are detailed as follows.

Input: D_1, Eps_1, Eps_2, Eps_3, and *MinPts*.

Output: D_2, C, M, $N_{Eps_1,Eps_2,Eps_3}(p)$, t_{pq}, and d^1_{pq}.

Let C denote the set of clusters, p represent a core object, and M signify the set of all core objects. Let set $N_{Eps_1,Eps_2,Eps_3}(p)$ be defined as the collection of sites within the neighborhood of site p; specifically, this notation reflects the inverse mapping relationships between p and q (where $q \in N_{Eps_1,Eps_2,Eps_3}(p)$). Additionally, t_{pq} and d_{pq}^1 represent the spatial-temporal distances between sites p and q. The neighborhood set $N_{Eps_1,Eps_2,Eps_3}(p)$ is determined by equation(8), while the t_{pq} and d_{pq}^1 are calculated using equation (9).

$$\begin{cases} N_{Eps_1,Eps_2,Eps_3}(p)=\begin{Bmatrix} q\in D_1 \,\middle|\, dist_1(p,q)\le Eps_1, \\ dist_2(p,q)\le Eps_2, \\ dist_3(p,q)\le Eps_3 \end{Bmatrix} \\ dist_1(p,q)=\sqrt{(x_p-x_q)^2+(y_p-y_q)^2} \\ dist_2(p,q)=\sqrt{(X_p-X_q)^2+(Y_p-Y_q)^2} \\ dist_3(p,q)=\left|T_p-T_q\right| \end{cases} \tag{8}$$

$$\begin{cases} t_{pq}=dist_3(p,q) \\ d_{pq}^1=dist_1(p,q)+dist_2(p,q) \end{cases} \tag{9}$$

Here, $dist_1$, $dist_2$, and $dist_3$ represent the spatial deviation associated with the riding location, the spatial deviation associated with the dropping location, and the temporal deviation associated with the riding time for sites p and q, respectively.

2) Second and Third Spatial-Temporal Clustering

The second spatial-temporal clustering phase focuses on the pre-response of ride and drop-off sites within the departure region regarding their spatial-temporal dynamics. In this framework, the clustering algorithm categorizes the spatial-temporal attributes associated with ride actions at the noise sites. Following this step, the noise points dataset $(x_{1,2,3,\ldots}, y_{1,2,3,\ldots}, x'_{1,2,3,\ldots}, y'_{1,2,3,\ldots}, T_{1,2,3,\ldots})$ is subjected to a clustering process, resulting in the segregation into distinct clusters $(x_{1,\ldots}, y_{1,\ldots}, x'_{1,2,3,\ldots}, y'_{1,2,3,\ldots}, T_{1,\ldots})$ and a set of noise points. The significance of the numbering remains consistent with the previous step, while the attributes of the noise sites continue to be retained as reservation information.

The third spatial-temporal clustering then focuses on the spatial latitude pre-response of the destination region. In this step, the same clustering algorithm is applied to cluster the drop-off locations and departure times of the noise points. Upon completion of this step, the noise points dataset $(x_{1,2,3,\ldots}, y_{1,2,3,\ldots}, x'_{1,2,3,\ldots}, y'_{1,2,3,\ldots}, T_{1,2,3,\ldots})$ is subjected to a clustering process, resulting in segregation into distinct clusters $(x_{1,2,3,\ldots}, y_{1,2,3,\ldots}, x'_{1,\ldots}, y'_{1,\ldots}, T_{1,\ldots})$ and a set of noise points with the attributes of the noise points continuing to be retained as reservation information.

The second and third instances of spatial-temporal clustering are analogous to the initial clustering process, with differences in the input and output datasets. Specifically, the second clustering uses dataset D_2 as input and produces dataset D_3 and set $N_{Eps_1,Eps_3}(p)$ as output, while the third clustering takes D_3 as input and yields the noise site dataset D_4 and set $N_{Eps_2,Eps_3}(p)$ along with the final output. Within the output results, the sets $N_{Eps_1,Eps_3}(p)$ and $N_{Eps_2,Eps_3}(p)$ are calculated using equation (10), and the spatial distances are determined by equation (11).

$$\begin{cases} N_{Eps_1,Eps_3}(p)=\begin{Bmatrix} q\in D_2 \,\middle|\, dist_1(p,q)\le Eps_1, \\ dist_3(p,q)\le Eps_3 \end{Bmatrix} \\ N_{Eps_2,Eps_3}(p)=\begin{Bmatrix} q\in D_3 \,\middle|\, dist_2(p,q)\le Eps_2, \\ dist_3(p,q)\le Eps_3 \end{Bmatrix} \end{cases} \tag{10}$$

$$\begin{cases} d_{pq}^2=dist_1(p,q) \\ d_{pq}^3=dist_2(p,q) \end{cases} \tag{11}$$

Here, d_{pq}^2 and d_{pq}^3 denote the spatial distance between p and q during the second and third stages of spatial-temporal clustering.

3) Encoding and Decoding Rules

This study employs a binary hybrid structure for encoding chromosomes, which represent the decision variables of the model. The hybrid structure consists of two layers of numerical codes. The upper layer comprises two additional codes, encoded using natural numbers, representing sets M and C, respectively. The lower layer consists of two variable codes that are encoded in binary notation.

The chromosome length corresponding to y_i equals the number of elements p in M, with the sum of the digits indicating the number of dynamic site setups. In contrast, the chromosome length for x_{ij} corresponds to the number of sites in set C, where each gene segment's length matches that of y_i. A sum of the digits in a gene segment equal to 1 signifies that each reservation site belongs to a single dynamic site, while a sum of 0 indicates that the reservation site is a core site within the dynamic site structure. A schematic illustration of the chromosome encoding is provided in Fig. 1.

As depicted in Fig. 1, the demand dataset used in this case is defined as follows: set M includes the elements {24, 111, 75, 89, 193, 147}, while set C consists of {24, 99, 111, 75, 89, 193, 147, 53}. The association between potential dynamic sites and reservation sites is represented in the format {core site: [set of reservation sites]} , specifically as follows: {24: [24, 75, 99, 53, 111, 147, 193], 111: [111, 24, 89, 53], 75: [75, 24, 89], 89: [89, 75, 111, 193], 193: [193, 24, 89], 147: [147, 24, 75, 53, 99]}. Analyzing this example yields the decoding result, revealing the set of dynamic sites as {24, 89, 147}. The relationship between these dynamic sites and the reservation sites they serve is expressed as follows: {24: [24,

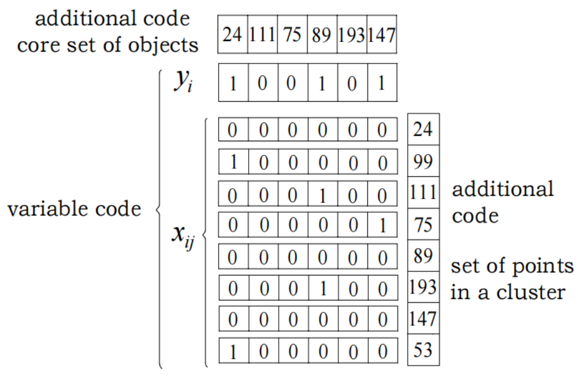


Fig. 1 Schematic diagram of chromosome encoding

99, 53], 89: [89, 111, 193], 147: [147, 75]}.

4) Crossover and mutation rules

Initially, two individuals are randomly selected along with the genetic segments of the lower y_i variable code chromosomes. Subsequently, a crossover operation is meticulously executed on these selected gene segments in conjunction with their corresponding upper layer codes. This process modifies the lower x_{ij} variable code chromosome segments while ensuring that the changes align with the association mapping between potential dynamic sites and reservation sites. The mutation operation follows similar principles to those of the crossover operation.

A graphical illustration of the chromosome crossover process is detailed in Fig. 2. In this figure, shaded squares indicate the selected crossover segments, whereas squares with scattered dots represent the segments that have been modified as a result of the crossover operation. In this example, the additional code for the selected crossover segment is 75. After the crossover, reservation site 75 in Individual 1 is transformed into a dynamic site, necessitating adjustments to its associated reservation sites, which change to {75: [75, 24]}. As a result, dynamic site 24 is reclassified as a reservation site, and the reservation sites within its service range, {99, 53}, are randomly reassigned to the service range of dynamic site 147 based on the existing set of relationships. This adjustment rule is similarly applied to modify the encoding of Individual 2.

Following the crossover, the dynamic sites of Individual 1 are {75, 89, 147}, with the relationship between dynamic sites and the reservation sites they serve expressed as {75: [75, 24], 89: [89, 111, 193], 147: [147, 99, 53]}. For Post-crossover Individual 2, the dynamic sites are {111, 147}, with the corresponding relationship being {111: [111, 89, 193, 24], 147: [147, 53, 99, 75]}.

C. Improved Algorithm Interpretation

The depiction in Fig. 3 elucidates the improvement process of the NSGA-III algorithm through the application of the ST_DBSCAN algorithm in three rounds to address the model effectively. This figure highlights the methodological integration of the spatial-temporal clustering approach with the multi-objective optimization algorithm, showcasing the synergy between these computational techniques.

In the figure, the spatial distribution of demands is represented, where sites in City A denote departure demand, and sites in City B signify arrival demand, collectively forming complete trips. In the improved algorithm, sites sharing the same shape indicate travel reservation OD demand from similar passengers. The ellipses represent clustered data that maintain consistent attributes among sites within these clusters. Dashed circles indicate the coverage areas surrounding reservation sites categorized as dynamic sites. The circular, square, and triangular symbols within Cities A and B correspond to the outcomes of the first, second,

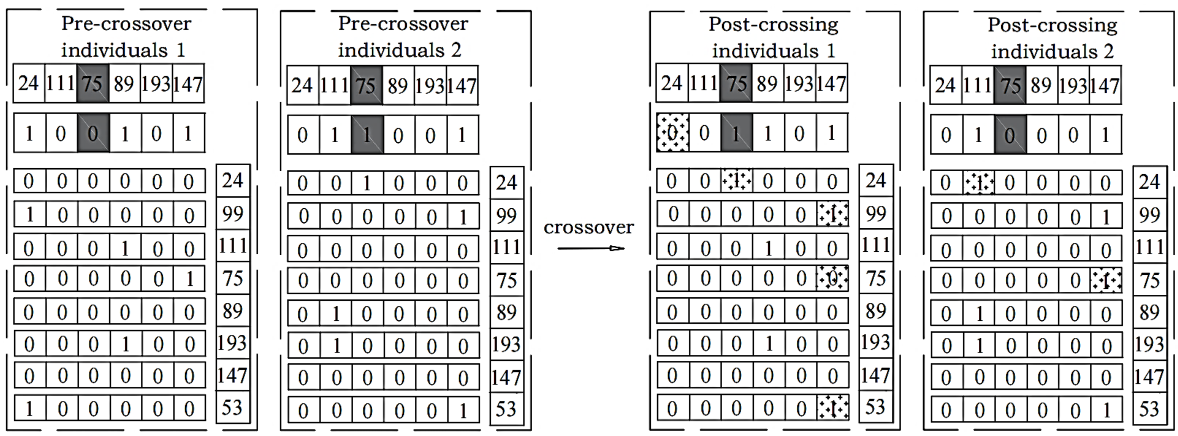


Fig. 2 Chromosome crossover schematic

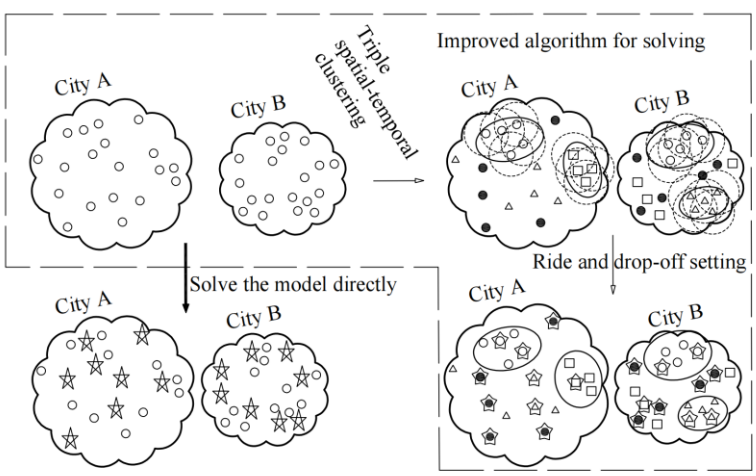


Fig. 3 Improved algorithm interpretation

and third rounds of spatial-temporal clustering, respectively. Black dots denote noise points that remain unclustered after three rounds of clustering.

In the results of the improved algorithm, dynamic ride and drop-off sites are represented by a combination of pentagram markers and original reservation site indicators. Distinct shapes and colors are used to differentiate the four types of reservation sites. For the direct solution component of the model, five-pointed stars represent the determined dynamic ride and drop-off sites.

A comparative analysis of the two solution methods reveals that the improved NSGA-III algorithm effectively reduces the processed data through spatial-temporal clustering, establishing ride and drop-off sites within these clusters. This approach maximizes the number of served reservation sites while minimizing the number of ride and drop-off sites, thereby avoiding scenarios where a single ride and drop-off site serves only one reservation site. Consequently, this methodology enhances the efficiency and effectiveness of the algorithm.

V. EXAMPLE VALIDATION AND RESULT ANALYSIS

A. Example Data

The mathematical model developed within the scope of this study work is implemented utilizing the Python 3.7 programming environment, along with the CPLEX 22.1.0.0 optimization solver, facilitating both algorithmic coding and resolution processes. The computational experiments were conducted on a personal computer equipped with an Intel Core i5-8265U CPU operating at 1.60 GHz, coupled with 8 GB of random-access memory (RAM).

This study aims to explore a CPT service model designed to transport passengers directly from their designated pick-up sites at the origin to their designated drop-off sites at the destination. The research is contextualized within Chongqing, one of China's four municipalities, focusing specifically on its significant district, Tongliang. The study utilizes CPT ticketing data from the "Chongqing Yukexing" platform for the year 2021. By analyzing the spatial and temporal distribution characteristics of this data and integrating it with actual POI data from residential areas within the region, we simulated the demand for CPT bookings from Chongqing's main urban area to Tongliang District over a single day, selecting a sample of 200 demands. Specific details of the related data are presented in TABLE III.

Furthermore, this paper provides a detailed examination of the current operational status of CPT routes from Chongqing's main urban area to Tongliang District, with specific data available in TABLE IV. Through this research, we aim to assess the efficiency and feasibility of the CPT service model in actual operations, thereby offering theoretical support and practical guidance for urban transportation planning and management.

B. Model Solving Results

1) Parameterization

Based on the data provided in TABLE IV, this study derived the current scheduling results for the example case, along with the deviations between passengers' actual departure locations times compared to their expectations. Under the current operational mode, the deviation of the actual drop-off locations from expectations is not considered due to the strategy of dropping off passengers within the regional service area. The detailed analysis results are presented in TABLE V.

According to TABLE V, the total time deviation on the current routes is approximately 260 hours, indicating relatively low sensitivity to time among passengers. From this analysis, a specific reservation site was selected, and the time distances between this site and other reservation sites were calculated and ranked in ascending order. A significant inflection point was identified at a 2-hour time difference, leading to the establishment of the Eps_3 parameter to 2 hours as a critical threshold for the time distance.

Furthermore, considering that the average speed of a customized passenger vehicle for riding and dropping passengers in the origin and destination areas is 40 km/h, and

TABLE III
DATA OF RESERVATION DEMAND SITE

No.	Departure	Destination	Ride location	Drop location	Expected riding time
1	Royal View Riverside	Sunshine Waterfront	(106.588687, 29.569077)	(106.057794, 29.850683)	9:42
2	34 Dahuang Rd.	The Garden of the Tilling of the Heart	(106.529592, 29.542067)	(106.011960, 29.856476)	7:42
…	…	…	…	…	…
200	Kunyu Mansion	Shihe Village 13	(106.565501, 29.563165)	(106.002172, 29.856307)	17:00

TABLE IV
CURRENT OPERATION OF THE LINE

Line name	No.	Ride site	Drop-off site	Vehicle type	Total number of shift
Chongqing ⇒ Tongliang Highway (G93)	1	Longhu Paradise Street, Daping → Daping Hospital → Daping Elementary School → The First Hospital of Chongqing Medical University → Chenjiaping Passenger Station	Ride and drop within the service area	9-seater commercial vehicle	36
	2	Hongqihegou passenger Station → North District Public Transportation Service Site → Ranjiaba Maternal and Child Health Center	Ride and drop within the service area		
	3	Chongqing West Railway Station Passenger Station → Xinqiao Hospital Outpatient Clinic → Southwest Hospital Outpatient Clinic → Shapingba Yankou Bus Stop	Ride and drop within the service area		

TABLE V
RESULTS OF EXAMPLE DATA LOADING ON CURRENT ROUTES AND SCHEDULING

Shift	Line No.	Ride time	Passenger ridding Service Sequence	Distance deviation Z_1 (km)	Time deviation Z_2 (h)
1	1	6:50 → 6:53 → 6:55 → 7:00 → 7:20	183 → (24, 43, 69, 122, 6, 23) → ( ) → ( ) → ( )	119.88	1.65
2	2	8:20 → 8:26 → 8:30	(137, 142, 67, 36) → (108, 26, 21) → 126	118.90	4.14
…	…	…	…	…	…
35	3	19:45 → 19:50 → 19:55 → 19:55	18 → 179 → (101, 45, 89) → (91, 96, 172)	365.18	15.87
36	1	20:30 → 20:33 → 20:35 → 20:40 → 21:00	(75, 167, 156) → (57, 85) → ( ) → (120, 71) → 178	53.98	6.40
Total	—	—	—	4723.23	262.98

TABLE VI
RESULTS OF THREE-TIME SPATIAL-TEMPORAL CLUSTERING

Spatial -temporal clustering	Cluster No.	{Core site: [Reservation site]}	Number of core sites: Cluster inside	Number of core sites: Total	Number of noise site: Noise site	Number of noise site: Total
No.1	1	{188: [0, 28, 121, 61, 127], …, 101: [153]}	14	100	[5, 10, …, 195, 196]	100
	…	…	…			
	28	{178: [198], 198: [178]}	2			
No.2	29	{5: [81, 19], …, 75: [59]}	25	78	[24, 32, …, 192, 195]	22
	…	…	…			
	37	{72: [27, 82, 99], …, 77: [68]}	6			
No.3	38	{24:[49],49:[24]}	2	12	[50, 52, …, 192, 195]	10
	…	…	…			
	42	{158:[46],46:[158]}	2			

based on the data in TABLE V, we assumed an upper limit of 4 sites per trip. Consequently, we set the density parameter *MinPts* to 2 to ensure that the density requirement in the clustering process is satisfied. Additionally, since passenger ride and drop-off services need to be completed within 40 minutes, we set both the Eps_1 and Eps_2 parameters to 7 kilometers as thresholds for spatial distance. This facilitates the definition of the geographic distribution of passengers and the service range of reservation sites in the clustering analysis. These parameters are designed to optimize the scheduling strategy for customized passenger services, aiming to reduce passenger waiting times, improve service efficiency, and enhance the overall travel experience.

2) Improved algorithm solving results

For the population operation, we divided the axes of the three objectives into 12 equal segments (p = 12), resulting in H = 91 reference points. We set the population size (*PopSize*) to H, the mutation probability (*Pm*) to 0.9, the crossover probability (*Pc*) to 0.1, and the maximum number of generations (*MaxGen*) to 500. This configuration allowed us to generate three sets of spatial-temporal clustering results, as detailed in TABLE VI.

As indicated in TABLE VI, the three rounds of clustering effectively partitioned the reservation demands. The initial clustering yielded 28 clusters comprising 100 reservation sites and 100 noise sites, successfully clustering 50% of the reservation demands. The second round produced 9 clusters containing 78 reservation sites and 22 noise sites, while the third round identified 5 clusters with 12 reservation sites and 10 noise sites. Collectively, the three rounds of the ST_DBSCAN algorithm partitioned 95% of the reservation demands, thereby reducing the proportion of noise sites to 5%. This process established mapping relationships between potential dynamic ride and drop-off sites and reservation sites, effectively decreasing the data scale of the CPT-DRDSS model.

The improved algorithm, when applied to solve the model, yields a total of 31 Pareto-optimal solutions, with their distribution in the solution space illustrated in Fig. 4. The minimum and average values of these 31 Pareto-optimal solutions across the three objective dimensions are presented in TABLE VII.

As evidenced in TABLE VII, an inverse relationship is observed between the growth trend of Z_3 (the number of

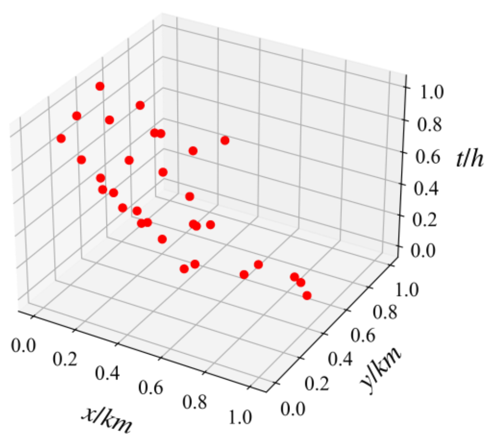


Fig. 4 The distribution of 31 Pareto optimal solutions

TABLE VII
PARTIAL RESULTS OF THE PARETO OPTIMAL SOLUTION ARE SHOWN

	Z_1 (km)	Z_2 (h)	Z_3
Z_1 minimum	490.38	113.32	79
Z_2 minimum	497.08	106.52	80
Z_3 minimum	518.08	120.12	74
Average value	511.85	112.83	76

TABLE VIII
RESULTS OF DYNAMIC RIDE DROP SITE SETTING

Dynamic site	Cluster No.	{Dynamic site: [reservation site]}	Number of sites	
			Cluster inside	Total
Type 1	1	{188: [0], 153: [101], …, 28: [1, 61, 121]}	6	43
	…	…	…	
	28	{178: [198]}	1	
Type 2	29	{5: [81], 76: [19], …, 59: [75, 32]}	10	31
	…	…	…	
	37	{72: [27], 68: [77]}	3	
Type 3	38	{24:[49]}	1	5
	…	…	…	
	42	{158: [46]}	1	
Type 4	—	[50, 52, …, 192, 195]	—	10

dynamic ride and drop-off sites) and Z_1 (the deviation of actual ride and drop-off locations from expectations) as well as Z_2 (the deviation of passengers' actual riding times from expectations). Specifically, an increase in Z_3 corresponds to a decrease in both Z_1 and Z_2. This finding suggests that increasing the number of dynamic ride and drop-off sites can lead to a reduction in the deviations of actual ride and drop-off locations and passengers' riding times, thereby enhancing the overall quality of passenger travel services.

Furthermore, the results for the four types of dynamic ride and drop-off site settings corresponding to the minimum in the first target direction are shown in TABLE VIII. The NSGA-III algorithm, improved by three rounds of spatial-temporal clustering, effectively identified four distinct categories of dynamic ride and drop-off sites during the model-solving process. The initial round of spatial-temporal clustering yielded 43 dynamic sites, the second round produced 31 dynamic sites, and the third round resulted in 5 dynamic sites. Furthermore, 10 dynamic ride and drop-off sites were classified under the fourth category.

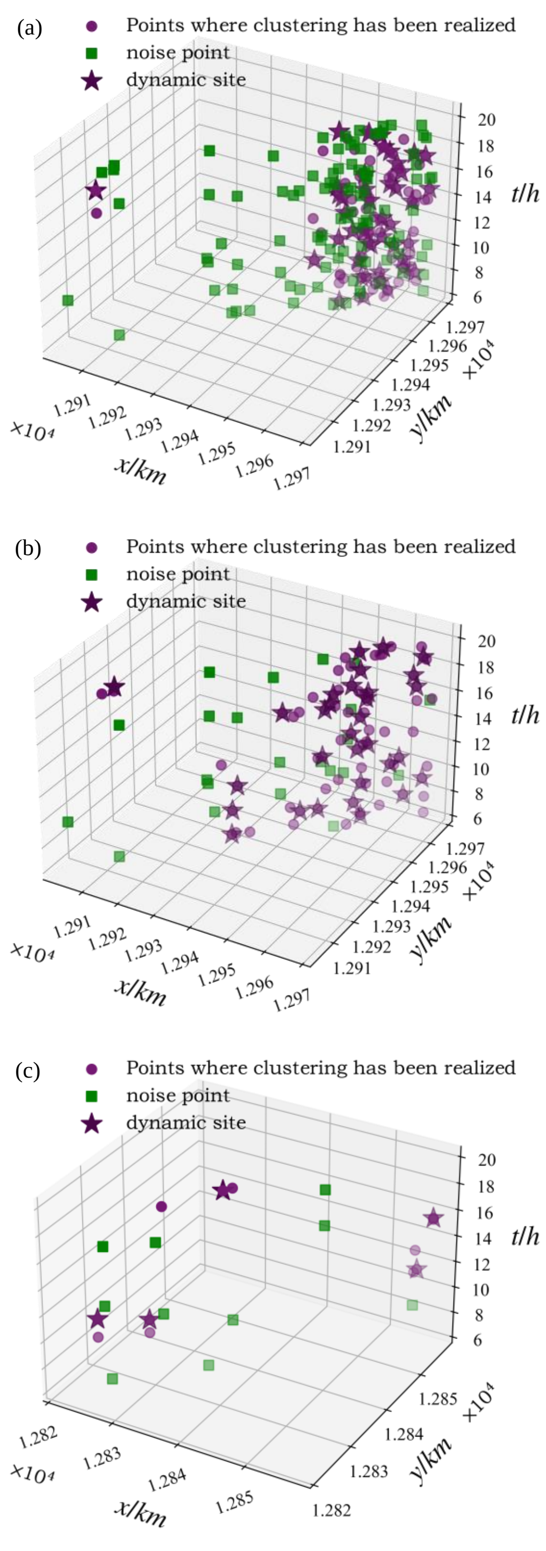
For the dynamic ride and drop-off sites of this fourth category, we considered the objectives embedded within the model and the prevailing operational paradigm of the case line within the drop-off service area. These sites were assigned to those exhibiting minimal deviation in spatial and temporal attributes of departure times. Consequently, these sites are serviced accordingly, optimizing allocation and service delivery within the dynamic ride and drop-off system.

The study also visualizes the effects of the first, second, and third spatial-temporal clustering in Fig. 5 (a) - (c), respectively, while the dynamic site setup is summarized in TABLE VIII. This visualization provides additional insights into the clustering process and highlights the effectiveness of the improved algorithm in enhancing the CPT service model.

C. Comparative Analysis of Algorithms

1) Validation of the validity of the algorithm

To verify the effectiveness of the proposed algorithm, we designed two experiments. The first experiment involved comparing the results of scheduling case data on the current line with the scheduling outcomes derived from the methods presented in this paper. This simulation of the algorithm's scheduling, based on the frequency of the current line,



(a) The initial round; (b) The second round; (c) The third round
Fig. 5 Improved NSGA- III algorithm solution (Z_1 minimum)

employed the following systematic approach:

a) Reservation demands within the same cluster are allocated to a single vehicle for service delivery.

b) Reservation demands across different clusters are consolidated and scheduled onto the same vehicle by calculating the average spatial-temporal distances among them.

c) Clusters with a substantial number of reservation sites are directly assigned vehicles for operation.

d) Clusters with a limited number of sites are considered for demand consolidation, wherein reservation demands from different clusters are amalgamated and scheduled onto the same vehicle.

Using this method, we calculated the deviations between actual ride locations and times compared to passengers' anticipated preferences. The results of this analysis are detailed in TABLE IX.

Operational data from Shift 2 revealed that the customized passenger vehicle reached dynamic stop 'B Chun Court' (numbered 128) at 7:54, successfully picking up passengers with reservation demand numbers 6, 83, and 128. Subsequently, at 8:12, the vehicle arrived at dynamic site 'Kafu Community' (numbered 8) to collect passengers with reservation demand numbers 7 and 8. The vehicle then proceeded to dynamic site 'Jade Ark of Chongqing Tiandi' (numbered 136) at 8:24 to pick up passengers with reservation demand numbers 68 and 136. For this trip, the mean deviation between the passengers' actual and desired ride locations was recorded as 22.92 kilometers, while the mean deviation between actual and desired boarding times was 1.5 hours.

A comparative analysis of the results from the proposed algorithm with those detailed in TABLE V indicates a significant reduction in the frequency of trips. Specifically, the deviation of passengers' actual ride locations from their expected ones decreased to 12.97%, representing a reduction of 4,110.67 kilometers compared to the previous scenario. Similarly, the deviation of actual ride times from the anticipated times dropped to 12.62%, leading to a total reduction of 229.78 hours, or an average decrease of 1.15 hours per individual. These findings strongly support the effectiveness of the algorithm.

TABLE IX
ARITHMETIC SIMULATION OF SCHEDULING RESULTS

Shift No.	Passenger riding service sequence {Dynamic site: [reservation site] (time)}	Fitness	
		Z_1 (km)	Z_2 (h)
1	153: [101] (6:42) → 42: [161] (6:48) → 188: [0, 61, 121] (8:36)	28.13	1.80
2	128: [6, 83] (7:54) → 8: [7] (8:12) → 136: [68] (8:24)	22.92	1.50
…	…	…	…
33	179: [118, 108] (19:30) → 126: [124] (20:00) → 88: [100] (20:10)	17.46	0.80
Total	—	612.56	33.20

TABLE X
COMPARISON RESULTS OF THE IMPROVED ALGORITHM WITH CPLEX

Comparison parameters	Improved algorithm	CPLEX	Difference value
Z_1 (km)	328.57	403.27	-74.70
Z_2 (h)	50.41	59.33	-8.92
Z_3	40	43	-3
Computation time (s)	59.79	1,267.56	-1,207.77

To further validate the accuracy and efficacy of the improved algorithm, a comparative analysis was conducted between the outcomes of the first category of dynamic ride and drop-off site configurations and the computational results obtained from solving the 0-1 integer programming problem using CPLEX. The comparison metrics included the fitness values Z_1, Z_2, and Z_3, as well as the computational time required for solution convergence. The detailed comparative results are presented in TABLE X. This analysis provides a quantitative basis for assessing the algorithm's performance, confirming its superiority in practical operational contexts.

In a second experiment, we compared the validation results of the first type of dynamic site setup using CPLEX to solve the 0-1 integer programming problem with the results obtained from the method presented in this paper. The experimental results, as depicted in TABLE X, demonstrate that the improved algorithm significantly outperforms the CPLEX solution in both solution quality and efficiency. Notably, the improved algorithm reduces the computation time to just 4.7% of that required by CPLEX. This substantial reduction in computation time enhances the algorithm's practicality and better aligns it with the demands of real-world applications.

2) Efficient validation of algorithms

To verify the efficiency of our algorithms, we conducted two experiments. Given that the NSGA-II algorithm can handle simple three-objective models, the first experiment compares the performance of the improved strategy integrated into the NSGA-II framework for solving the CPT-DRDSS model with that of the refined NSGA-III algorithm.

Through iterative execution of the program, we consistently obtained approximately three Pareto-optimal solutions, all of which are part of the set of 31 Pareto-optimal solutions derived from the refined NSGA-III algorithm presented in this study. This observation highlights the instability and incompleteness of the optimal solutions generated by the improved NSGA-II algorithm. To compare the operational outcomes of the improved NSGA-III and the improved NSGA-II algorithms, we employed two widely recognized performance metrics: Inverted Generational Distance (IGD) and Hypervolume (HV).

a) Inverted Generational Distance (IGD) [33]. This metric measures the distance between the obtained solution set and the real or ideal Pareto front. It reflects the coverage and separation between the solution set and the ideal set by calculating the average of the minimum distances from a set of reference points to the target solution set. Smaller values indicate a closer distance between the solution set from the algorithm and the Pareto front, signifying better algorithm performance.

b) Hypervolume (HV) [34]. This metric assesses the performance of a multi-objective optimization algorithm by measuring the volume of the objective space enclosed by the

TABLE XI
COMPARISON OF 30 RUNS OF IGD AND HV BETWEEN THE IMPROVED NSGA-III AND THE IMPROVED NSGA-II

Name	IGD			HV		
	Min	Median	Max	Min	Median	Max
NSGA-III	0.1746	0.1927	0.1998	0.3605	0.4671	0.7202
NSGA-II	0.2165	0.2305	0.3583	0.0788	0.2974	0.4418

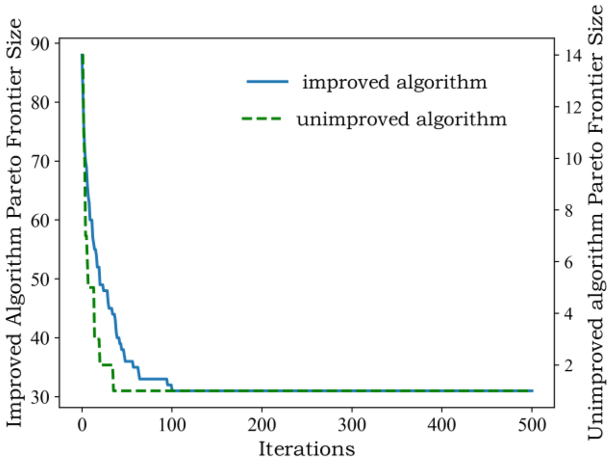


Fig. 6 Comparison of Pareto front-end size change

Pareto front. A larger HV value indicates better quality and diversity of the solutions.

We performed a comparative analysis of the IGD and HV metrics over 30 independent runs for both the improved NSGA-III and NSGA-II algorithms to evaluate their performance and efficacy in addressing the CPT-DRDSS model. The results of these runs are presented in TABLE XI. Upon examination of the data in TABLE XI, it is evident that the improved NSGA-III algorithm significantly outperforms the improved NSGA-II algorithm in terms of both IGD and HV metrics.

The second experiment aimed to validate the efficiency of the algorithm through a comparative analysis between the solution outcomes from the enhanced algorithm—utilizing the mean of Pareto-optimal solutions—and those from the unimproved NSGA-III algorithm, which, while lacking spatial-temporal clustering, follows the methodologies presented in this study.

Consistent with the approach of the unimproved algorithm, we incrementally input reservation data of varying scales, labeled as D_1, D_2, and D_3, into the model to determine the first three types of dynamic stations. The comparative outcomes of the two algorithms are delineated in TABLE XII.

Upon examining the objective values derived from the algorithmic solutions in TABLE XII, it is evident that the improved algorithm outperforms the unimproved counterpart across the Z_1, Z_2, and Z_3 metrics. Notably, the Z_1 metric was reduced by 60.9 kilometers, while the Z_2 metric constituted 24.84% of that of the unimproved algorithm. This indicates that the improved algorithm is more effective in minimizing the deviation between actual and desired passenger drop-off locations, as well as in reducing ride time deviations.

The improvement is particularly pronounced in the establishment of the second category of dynamic sites, where the Z_1 metric decreases by 59.33 kilometers, while the Z_2 metric accounts for 18.20% of that obtained using the unimproved algorithm. This suggests that the improved algorithm is most effective in reducing passenger distance and time deviations associated with the configuration of the second category of dynamic sites. Furthermore, the Z_3 metric obtained from the improved algorithm slightly outperforms that of the unimproved algorithm, thereby substantiating the accuracy of the enhancements made.

Regarding algorithmic convergence time, the improved algorithm demonstrates expedited performance in both incremental and cumulative time requirements for resolving the initial three types of dynamic stops, with the total time being 30.02% of that required by the unimproved algorithm. This suggests that the improved algorithm can significantly reduce the duration for establishing dynamic sites. When addressing the first and second types of dynamic stops with input data sizes of 200 and 100, respectively, the convergence time for the improved algorithm is 29.39% and 29.37% of that for the unimproved algorithm. Conversely, when solving for the third type of dynamic stops with an input of 22 reservation sites, the convergence time is 64.21% of that of the unimproved algorithm. This indicates that the improved algorithm's high convergence efficiency is more pronounced with larger datasets and less noticeable with smaller ones. The input data for the first type of dynamic stops is more substantial, making it more indicative of real-world scenarios.

As shown in Fig. 6, both algorithms rapidly converge to the Pareto optimal solutions of the model. However, the improved algorithm demonstrates a higher frequency of convergence and a shorter overall convergence time, indicating a more efficient convergence rate and more pronounced effects compared to the unimproved algorithm. This underscores the enhanced efficiency of the improved algorithm, particularly in the context of large-scale datasets.

To further elucidate the convergence efficiency of the improved algorithm, we compared the two algorithms in resolving the first type of dynamic stops, with the improved algorithm utilizing a dataset structured into 28 clusters and the unimproved algorithm using a dataset comprising 200 reservation sites. The variation in the size of the Pareto front is illustrated in Fig.6.

TABLE XII
COMPARISON RESULTS OF THE IMPROVED ALGORITHM WITH UNIMPROVED ALGORITHM

Dynamic type	Results of the improved algorithm				Results of the unimproved algorithm			
	Z_1 (km)	Z_2 (h)	Z_3	Convergence time (s)	Z_1 (km)	Z_2 (h)	Z_3	Convergence time (s)
Type 1	328.57	50.41	40	59.79	329.45	131.88	40	203.42
Type 2	178.23	57.52	31	42.31	237.56	316.00	35	144.05
Type 3	5.05	4.90	5	4.18	5.74	6.30	5	6.51
Total	511.85	112.83	76	106.28	572.75	454.18	80	353.98

VI. Conclusion and Future Work

The main conclusions drawn from this paper, along with future outlooks, are as follows:

1) Current operational models in CPT enterprises typically establish fixed service sites based on historical passenger travel data. Vehicle scheduling results are announced through platforms, allowing passengers to purchase tickets according to their needs. However, this approach results in passive engagement from passengers, failing to fully address their personalized spatial and temporal needs. In contrast, the model and solution algorithm proposed in this study effectively cater to these personalized demands, thereby leveraging the full advantages of CPT.

2) This study proposes a novel CPT-DRDSS method, grounded in actual reservation needs. This method comprehensively considers both departure and arrival locations, analyzing data from spatial and temporal dimensions through three rounds of ST_DBSCAN spatial-temporal clustering. The NSGA-III algorithm is enhanced by reducing the initial individual size, which allows for the classification of dispersed and large volumes of CPT spatial-temporal demand into clusters with similar attributes. This approach effectively minimizes the impact of noise sites in the density clustering algorithm and categorizes dynamic ride and drop-off sites into four distinct types.

3) The results of the refined algorithm analysis show that three rounds of ST_DBSCAN spatial-temporal clustering can classify 95% of the reservation needs into four types of clusters. Based on these classifications, the number of dynamic ride and drop-off sites is set at 43, 31, 5, and 10, respectively. Compared to current route results, the proposed method reduces the deviation of passengers' actual ride locations from their expectations to 12.97%, translating to a total reduction of 4,110.67 kilometers. Additionally, the deviation of actual ride times is reduced to 12.62%, equating to a total reduction of 229.78 hours, or an average of 1.15 hours per passenger. When comparing with CPLEX solution results, the proposed algorithm reduces computation time to just 4.7% of that required by CPLEX, demonstrating its suitability for practical applications. Moreover, when comparing the settings of the first three types of dynamic stations with those of the unimproved algorithm, the proposed method shortens algorithm convergence time to 30.02% of that of the unimproved algorithm, optimizing the objective function value. The algorithm exhibits greater efficiency, particularly when the customized passenger ride and drop-off sites are spatial-temporally dispersed and the data scale is large.

4) While this paper introduces a novel method for solving the CPT-DRDSS problem by considering the time and distance of passengers arriving at dynamic stations, it is limited by the number of ride and drop-off sites set by enterprises. The layout of these sites significantly impacts the order of rides and the distances traveled by vehicles, which in turn affects route planning. Future studies should aim to further optimize route design in conjunction with the layout of ride and drop-off sites. This optimization should take into account both the total time passengers spend on the vehicle and the economic efficiency of the enterprise during the planning process. Additionally, exploring the integration of real-time data and advanced predictive analytics could enhance the responsiveness and adaptability of CPT systems.

References

[1] Song C Y, Wang H L, Tian Z H, and Niu X Q, “A Review of Demand-Responsive Bus Vehicle Scheduling Models and Algorithms,” Journal of Beijing Jiaotong University, vol. 47, no. 4, pp. 31-44, 2023.

[2] Armando D, Teresa G D, Jorge P S, and Bruno A P, “Configurations and features of demand responsive transports,” Transportation Research Procedia, vol. 78, pp. 71-78, 2024.

[3] Huang Y J, “Pain Sites and Development Strategies of Customized Passenger Transportation,” China Road Transportation, no. 2, pp. 40-42, 2024.

[4] Wang J C, “Development Path of Customized Passenger Transportation,” China Road Transportation, no. 10, pp. 46-48, 2023.

[5] Tian Y, Yang D, Wang J, Lin W F, Jiang Y, and An S, “Designing and planning sustainable customized bus service for departing attendees of planned special events: A two-phase methodology integrating data-driven and demand-responsive,” Sustainable Cities and Society, vol. 96, p. 104630, 2023.

[6] Li J, Xie C, Wang L, and Gao Z L, “Extraction method of stopping station location of road passenger shuttle buses based on trajectory data,” Transportation Information and Safety, vol. 39, no. 4, pp. 60-67, 2021.

[7] Seyed P P, Mehdi M, and Mostafa S, “A bi-level school bus routing problem with bus stops selection and possibility of demand outsourcing,” Applied Soft Computing, vol. 61, no. 1, pp. 222-238, 2017.

[8] Ren J X, Chang X T, Wu W T, and Jin W, “Demand-responsive feeder bus scheduling considering candidate stops and full-service process,” Transportation Systems Engineering and Information, vol. 23, no. 5, pp. 202-214, 2023.

[9] Cortenbach L E, Gkiotsalitis K, Berkum E C, and Walraven E, “The Dial-a-Ride problem with meeting sites: A problem formulation for shared demand‑responsive transit,” Transportation Research Part C: Emerging Technologies, vol. 169, p. 104869, 2024.

[10] Zheng H, Zhang X C, and Wang Z M, “Research on the customization problem of hybrid model demand response bus service,” Transportation Systems Engineering and Information, vol. 18, no. 2, pp. 157-163, 2018.

[11] Li W Y, Yu Z W, Wei S, Lu X Y, Zhang Y Q, and Wang T, “Custom bus optimization of docking stations based on partition path selection,” Advances in Mechanical Engineering, vol. 9, no. 10, pp. 1-6, 2017.

[12] Qiu F, Li W Q, and Zhang J, “A dynamic station strategy to improve the performance of flex-route transit services,” Transportation Research Part C: Emerging Technologies, vol. 48, no. 12, pp. 229-240, 2014.

[13] Lei Y W, Lin P Q, and Yao K B, “Network scheduling model of internet customized bus and its solution algorithm,” Transportation Systems Engineering and Information, vol. 17, no. 1, pp. 157-163, 2017.

[14] Shen C, Sun Y, Bai Z J, and Cui H J, “Real-time customized bus routes design with optimal passenger and vehicle matching based on column generation algorithm,” Physica A, vol. 571, p. 125836, 2021.

[15] Ayman A, Pieter L, Tobias B, and Wolfgang K, “Balancing convenience and sustainability in public transport through dynamic transit bus networks,” Transportation Research Part C: Emerging Technologies, vol. 151, p. 104100, 2023.

[16] Hu D, and Jin W Z, “Demand response bus scheduling based on station optimization,” Journal of Shenzhen University (Science and Technology Edition), vol. 39, no. 2, pp. 209-215, 2022.

[17] Yu Z, “Demand-responsive transit planning considering passenger travel experience,” Transportation Technology and Economy, vol. 22, no. 2, pp. 32-37, 2020.

[18] Xue H N, and Wang J, “Demand response mechanism of customized public transportation based on Spatial-temporal clustering,” Highway Transportation Science and Technology, vol. 38, no. 3, pp. 113-121, 2021.

[19] Hu L G, An T, Wang J, and Liu X, “Layout study of urban customized bus carpool stops,” Journal of Xuzhou Engineering College (Natural Science Edition), vol. 31, no. 1, pp. 27-32, 2016.

[20] Li W G, Zheng L J, Liao L Q, Yang X Z, Sun D H, and Liu W N, “A Multiline Customized Bus Planning Method Based on Reinforcement Learning and Spatiotemporal Clustering Algorithm,” IEEE

Transactions on Computational Social Systems, vol. 11, no. 3, pp. 3691-3705, 2023.
[21] Yuan T, Yang D, Wang J, Lin W F, Jiang Y, and An S, “Designing and planning sustainable customized bus service for departing attendees of planned special events: A two-phase methodology integrating data-driven and demand-responsive,” Sustainable Cities and Society, vol. 96, p. 104630, 2023.
[22] Hu D, Wang C, and Chen J W, “Intercity customized passenger transportation service plan optimization design with spatial-temporal accessibility based on BIRCH-VNS,” Neural Computing and Applications, vol. 36, no. 21, pp. 13127-13154, 2024.
[23] Wang H D, Song Y L, and Tian Y J, “An analytical study of a Spatial-temporal data clustering algorithm with multiple attributes,” Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), vol. 33, no. 4, pp. 661-668, 2021.
[24] Erdmann M, Dandl F, and Bogenberger K, “Combining immediate customer responses and car–passenger reassignments in on-demand mobility services,” Transportation Research: Part C, vol. 126, p. 103104, 2021.
[25] Zhang J, Li W Q, Zheng Y, and Guo R R, “Dynamic Clustering Meeting Sites Strategy to Improve Operational Service Capability of Flex-Route Transit,” Journal of Transportation Engineering, Part A: Systems, vol. 149, no. 6, pp. 1-11, 2023.
[26] Lu C Q, Tiwari S, Nassir N, and Nagel K, “Efficient operation of demand-responsive transport systems: Active requests rejection,” Procedia Computer Science, vol. 238, pp. 81-90, 2024.
[27] Feng S M, Hu B Y, Ma D C, and Wang D X, “Setting method of intercity expressway passenger station,” Journal of Harbin Institute of Technology, vol. 48, no. 3, pp. 71-76, 2016.
[28] Deng J H, Song R, Cheng X H, Wang T S, and Wang B H, “Integrated design of micro-circulation stops and line network for public transportation in open neighborhoods,” Journal of Beijing Jiaotong University, vol. 44, no. 1, pp. 34-41, 2020.
[29] Cui Z H, Zhang M Q, Chang Y, Zhang J J, Wang H, and Zhang W S, “NSGA-II based on average distance clustering,” Journal of Automation, vol. 47, no. 5, pp. 1171-1182, 2021.
[30] Deb K, and Jain H, “An Evolutionary Many-Objective Optimization Algorithm Using Reference-Site-Based Nondominated Sorting Approach, Part I: Solving Problems with Box Constraints,” IEEE transactions on evolutionary computation: A publication of the IEEE Neural Networks Council, vol. 24, no. 2, pp. 138-143, 2016.
[31] Birant D, and Kut A, “ST-DBSCAN: An algorithm for clustering spatial-temporal data,” Data & Knowledge Engineering, vol. 60, no. 1, pp. 208-221, 2006.
[32] Zheng L, Gao L P, Shen J X, and Li W Q, “Optimization strategy of variable route type bus operation service capacity,” Journal of Harbin Institute of Technology, vol. 53, no. 9, pp. 107-115, 2021.
[33] Leonardo C T, Bezerra Manuel L, and Thomas S, “An Empirical Assessment of the Properties of Inverted Generational Distance on Multi- and Many-Objective Optimization,” Evolutionary Multi-Criterion Optimization, vol. 10173, pp. 31-45, 2017.
[34] Andreia P G, Carlos M F, and Luís P, “The Hypervolume Indicator: Computational Problems and Algorithms,” ACM Computing Surveys, vol. 54, no. 6, pp. 1-42, 2021.

Jing Yuan was born in Chongqing, China, in 2001. She obtained her Bachelor degree from Chongqing Jiaotong University, Chongqing, China, in the year 2023. She is currently pursuing her master degree in Transportation Planning and Management in Lanzhou Jiaotong University. Her research interests include customized passenger operations organization study and urban public transportation planning and management.