

Article

A Multiobjective Land Use Design Framework with Geo-Big Data for Station-Level Transit-Oriented Development Planning

Shihai Dong ¹, Yandong Wang ^{1,2,3,*}, Mingxuan Dou ¹, Yanyan Gu ¹, Peiqi Zhang ¹ and Jianya Gong ^{1,4}

- ¹ State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; shihaid@whu.edu.cn (S.D.); mxdou@whu.edu.cn (M.D.); yyg@whu.edu.cn (Y.G.); pzhang13@umd.edu (P.Z.); gongjy@whu.edu.cn (J.G.)
² Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, China
³ Faculty of Geomatics, East China University of Technology, Nanchang 330013, China
⁴ School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China
* Correspondence: ydwang@whu.edu.cn; Tel.: +86-27-6877-8969

Abstract: Transit-oriented development (TOD) is among the most feasible strategies for relieving urban issues caused by the unbalanced development of transportation and land use. This study proposes a multiobjective TOD land use design framework for the optimization of the land use layout in station catchments. Given the high density and diverse development in Chinese megacities, a planning model that considers nonlinear impacts on ridership, land use efficiency, quality of life, and the environment is constructed. The model applies fine-grained geo-big data to fill gaps in the empirical and statistical data and improve practicability. A genetic multiobjective optimization approach without reliance on objective weighting is used to generate alternative land use schemes. A metro station in Shanghai is applied as a case study. The results indicate that the proposed ridership objective outperforms the commonly used linear function, and the optimization method has superior extreme optima and convergence to baseline models. We also discuss the consistencies and conflicts in the objectives and provide a balanced land use scheme considering local policies. This work provides suggestions for sustainable urban design with coordinated land use and transportation.

Keywords: geo-big data; land use planning; multiobjective optimization; nonlinear impact; transit-oriented development

Citation: Dong, S.; Wang, Y.; Dou, M.; Gu, Y.; Zhang, P.; Gong, J. A Multiobjective Land Use Design Framework with Geo-Big Data for Station-Level Transit-Oriented Development Planning. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 364. <https://doi.org/10.3390/ijgi11070364>

Academic Editors: Wolfgang Kainz, Maria Antonia Brovelli

Received: 10 May 2022
Accepted: 23 June 2022
Published: 25 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Urban mass rail transit (MRT) systems provide an effective tool for residents and promote a shift from private vehicles to a low-carbon public travel mode [1,2]. Nonetheless, cities face problems such as the inappropriate distribution of facilities and wasted social resources caused by unbalanced land use and MRT development. Therefore, comprehensive strategies that integrate MRT and relevant land use are critical to sustainable urban planning [3]. The transit-oriented development (TOD) theory is considered one of the most feasible integrated land use and transportation theories [4,5]. TOD planning can improve the unbalanced state of transportation and land use and offers a sustainable approach to urban development.

Numerous recent studies have focused on TOD planning and generated various effective strategies and methods within this topic [6–8]. Particularly, TOD planning with mathematical models is prominent in generating sustainable land use plans and an intuitive evaluation [9,10]. Quantitative TOD planning methods involve the significant branch of multiobjective optimization, which generates alternative land use schemes

according to objectives based on TOD principles and strategies [11,12]. However, mathematical TOD models for China remain limited, although one noteworthy example of a TOD model for designing land use layout given a type and density around a metro station in China was proposed by Li et al. [13–15]. First, the parameters of the objectives mainly comprise coarse-grained statistical data and empirical indicators, which lead the objective functions to have insufficient expressive ability. Second, most objectives have been constructed with linear and straightforward functions, thereby simplifying the complicated relationship between objectives and land use, thereby reducing model reliability. Finally, although multiobjective problems have been solved in previous studies by weighting objectives, determining the relative weights is challenging because of uncertainty, and rarely have studies used actual multiobjective optimization in TOD planning. Fortunately, the appearance of urban geo-big data and the development of the learning model provide excellent potential for the refined construction and optimization of planning objectives.

This paper presents a multiobjective TOD land use design framework to generate alternative land use schemes in the Chinese context, which consists of a TOD planning model and a multiobjective optimization approach for land use design. Traditional TOD planning models always consider several objectives, including the ridership of rail transit, compactness, land use mix, land use conflicts, and environmental effects [12–14]. In this work, elements of the built environment are added to the objective of ridership, and land use conflict is replaced by a more specific function of land use density. In addition, the objective of the total distance to the station is designed for the convenience of traveling by public transit. The proposed method makes several contributions. (1) We apply geo-big data to describe high-density and high-diversity development in the Chinese context and provide abundant information for constructing objective functions and validating the availability of the method. (2) We explore the nonlinear impact of land use on ridership and represent the objectives of ridership with an ensemble learning method to ensure that the ridership can approach the expectation of the model after optimization to enhance the practicability of the planning model. (3) The method generates alternative land use schemes via a multiobjective genetic algorithm with a designed mix coding, which eliminates the limits of objective weighting and provides further suggestions for decision-makers with a unique reference point structure and normalization process. A case study of a metro station in Shanghai is presented to demonstrate the feasibility of the proposed method. The objective function of ridership is compared with the linear function to show its superiority. The employed optimization method presents a reliable ability to achieve optimization and convergence via comparison with two commonly used genetic algorithms. In addition, an alternative land use layout obtained by the planning method appears to be balanced and consistent with TOD strategies.

The remainder of the paper is organized as follows. Section 2 reviews the related literature on TOD planning. Section 3 describes the research problem in detail and demonstrates the proposed TOD land use planning method. Section 4 introduces a case study in Shanghai and presents land use schemes to verify the effectiveness of the method. Section 5 discusses the relationship between the TOD planning objectives and provides several policy implications. Finally, the conclusions and limitations of the current work are discussed in Section 6.

2. Related Work

TOD was first introduced for development design in a mixed-use community with high density and diversity that encourages residents to live near a transit station and decreases their dependence on private vehicles [16]. Cervero [17] proposed the basic “3D” principles of “density, diversity and design” for TOD planning and argued that land use with high density and diversity and a pedestrian-oriented design would encourage non-auto travel [18]. Since then, the principles and strategies of TOD have attracted numerous scholars and have increased in complexity to fulfill the requirements of

urban sustainability [3,6,19–21]. As a green large-capacity travel mode, MRT is an appropriate research object for TOD planning, and it is acceptable to build stable MRT systems based on TOD strategies.

Although the principles and strategies of TOD have been comprehensively identified, there remain gaps in understanding how to transform TOD theory into a detailed analytical model for sustainable urban planning. The land use design problem (LDP) refers to a mathematical programming model for optimizing the locations, types, and sizes of urban land use [22]. Because of the tight relationship between transportation systems and land use, recent studies have increasingly focused on planning the integration of land use and transit design [23–25]. In this context, TOD strategies have been applied to land use design. Lin & Gau [11] proposed a TOD model to increase MRT ridership, improve living and environmental quality, and ensure social equity, but only commercial and residential land use was involved. As an extension, Lin & Li [12] developed a city-region level model that allocated recreational land use and added the objective of maximizing access to nonresidential activities. Li et al. [13] first established a TOD model for China based on the characteristics of Chinese megacities to maintain the sustainable development of the MRT system in China. On this basis, Ma et al. [14] considered the contribution of transportation to station-level planning, and local accessibility was appended to the objectives. Sahu [9] introduced a TOD planning model for modifying land use in the city of Naya Raipur that was based on the global TOD parameters of density, diversity, and distance to transit.

A notable limitation of previous TOD planning studies was that they used only statistical and empirical data, which were insufficient for refined planning and yielded impractical results for specific planning tasks. Recently, geo-big data such as points of interest, mobile sensors, and smart card data containing abundant information about human mobility and urban spatial structures have been widely applied to urban planning [26,27]. Geo-big data provide new opportunities to analyze TOD scenarios quantitatively and are capable of describing real-time urban spatial structures accurately for TOD planning [28]. In addition, ridership indicates the usage degree of the transit services and mobility triggered by land use, which is one of the most important indicators in TOD. Previous TOD planning models have generally estimated ridership with statistical trip generate/attraction ratio and linear (log-linear) functions. However, several studies have illustrated that the impact of land use on ridership is nonlinear, and the influence of ridership cannot be entirely explained by land use [29]. Therefore, a more comprehensive nonlinear model with multiple factors is necessary for constructing the ridership objective in the TOD planning model.

Even when a TOD model is constructed, optimization is challenging because the model involves multiple objectives with conflicting goals, a vast search space for the optimal solution, and a series of constraints. The genetic algorithm (GA) has been indicated to be appropriate for solving such large-scale problems, as it is adept at automatically and optimally searching in vast solution dimensions [30]. Most importantly, the GA has already been modified to solve multiobjective problems [31,32]. Consequently, various GAs have been applied to optimize the LDP and TOD models [9,33,34]. For multiobjective optimization, most TOD planning includes in its design a comprehensive function, $F = \sum_{k=1}^{obj} w_k F_k$, with weights to simplify the optimization [13,14,35]. However, the weights of the objectives are defined subjectively and different objectives have different dimensions, thereby resulting in models suffering from bias and sensitive outliers.

In summary, although the land use and TOD planning model have been constructed with all-round perspective, station-level mathematical TOD can be improved in terms of data and function representation. An improved reasonable TOD land use planning method is still required to be developed for sustainable urban design with the balanced development of the MRT system and land use in the Chinese context.

3. Methodology

3.1. Problem Statement

This paper presents a TOD planning method based on strategies appropriate to the Chinese context to solve the land use design problem. As depicted in Figure 1, considering the MRT station as the central point, the land use design problem aims to determine the type, density, and location of the surrounding undeveloped land use cells and to obtain an optimal land use layout based on the objectives relevant to TOD. Finally, alternative land use sketch maps are generated to display land use layouts [35,36]. Inspired by work by [14], seven types of land use are presented in this work, namely: public (e.g., school, government agency, hospital), industry (e.g., factory, warehouse), commercial (e.g., shopping mall, restaurant), economic (e.g., office building, corporate business, finance organization), residential (e.g., apartment, villa), road, and water. Notably, only the first five types of land use are applied to land use design because of the high cost of rectifying roads and water.

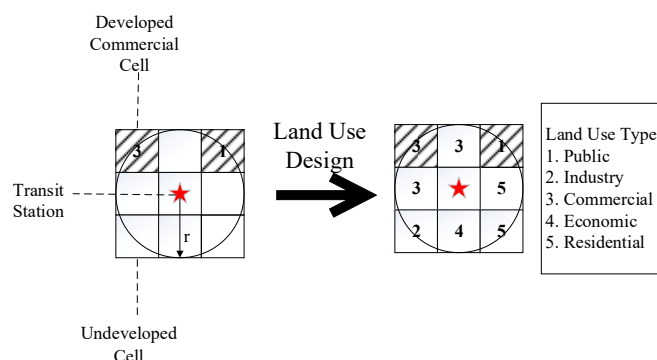


Figure 1. Description of the land use design problem.

In general, this study proposes a multiobjective TOD land use design framework. As shown in Figure 2, the method comprises two parts of a TOD planning model and a multiobjective optimization approach for land use design in a station catchment area. By using emerging geo-big data and traditional statistical and empirical data, a TOD planning model is constructed with linear and nonlinear objectives. Then, a robust heuristic algorithm is applied for generating alternative land use layouts. The details are introduced in the next sections.

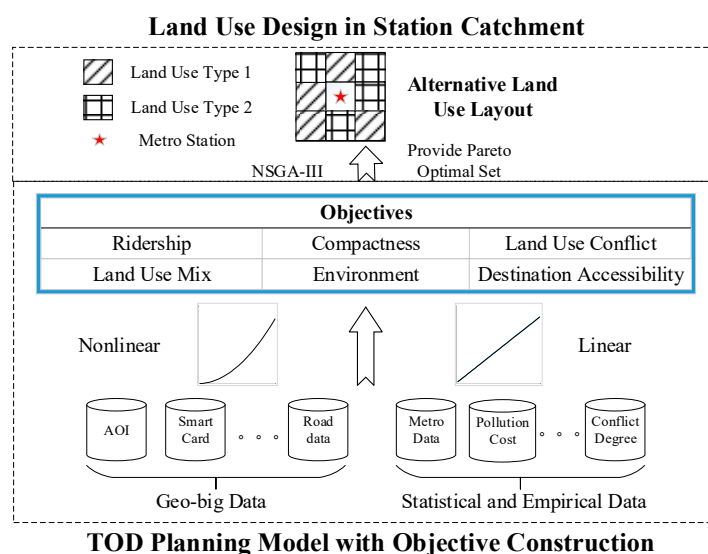


Figure 2. Work flowchart of the TOD land use design framework.

3.2. TOD Planning Model

A dependable TOD planning model needs to consider numerous factors, including ridership, land value, land use efficiency, environmental effects, quality of life, and social equality [11,37]. Nonetheless, different contextual characteristics result in specific applications of TOD strategies [3]. In contrast to most American cities, which are characterized by low density, car orientation, and multicenter layouts [38], most Chinese cities are characterized by high density and high transit ridership. To channel megacity growth in MRT corridors, TOD in China should contain high-density development, mixed land use, and pedestrian-friendly environments to be consistent with the features of Chinese cities. Based on the Chinese context and the concept of TOD, six objectives by referring to previous studies [11–15] are designed to promote the sustainable development of MRT and related land use: the perspectives of ridership, compactness, land use conflict, land use mix, environmental effects, and destination accessibility. Moreover, a set of constraints are considered in the model.

3.2.1. Parameter Definition

The proposed TOD planning model contains the following parameters.

I : set of land use cells; cell $i \in I$.

K : set of land use types; type $k \in K$.

BE : built environment variables in the station area.

T : transit service variables in the station area.

G : demographic variables in the station area.

N_i : cells within 8 neighbors of cell i .

c_{kl} : conflict degree between adjacent cells of types k and l .

$Side$: length of cells.

$Area$: area of cells.

N_{cell} : number of cells.

P_k^r : r -type pollutants generated by unit density in a k -type land use cell.

$cost^r$: cost of treating r -type pollutants.

L_i : distance from cell i to the metro station along the road.

$Attr_k$: passenger attraction of unit k -type cell.

$lowerEII_k/upperEII_k$: lower/upper bound, respectively, of the exploitative intensity index of each type of land use in each cell.

$lowerRatio_k/upperRatio_k$: lower/upper bound, respectively, of the total percentage of each type of land use in the overall layout.

3.2.2. Decision Variables

Two decision variables are applied to the allocation of land use type and density.

X_{ik} : binary variable. $\{X_{ik}=1$ if cell i is assigned to k -type; otherwise, $X_{ik}=0\}$, which describes the land use type of cells with one-hot encoding.

D_{ik} : k -type land use density of cell i ; the unit of land use density is 100 m²/cell, and the value of the density is a positive integer for each cell.

3.2.3. Objectives

1. MRT ridership

The first objective aims to maximize MRT ridership, which increases the utilization efficiency and profit for MRT and results in economic sustainability. Before the construction of the TOD objectives, an available model is required to determine the nonlinear relationship between land use and ridership. Tree ensemble models such as gradient boosting regression tree (GBRT) have been applied for the description of ridership [39,40]. In this study, as one of the best tree ensemble methods, extreme gradient boosting (XGBoost) [41] is used for nonlinear modeling. Compared with GBRT, XGBoost has important advantages in two main respects: XGBoost employs the second-order Taylor

expansion to deal with the cost function, resulting in high model accuracy; and penalty coefficients are added to the cost function to increase the generalization ability. XGBoost consists of numerous classification and regression trees (CARTs), and the result is predicted by the summation of the continuous score in the leaves of the CARTs. The prediction function is as follows:

$$\hat{y}_i = \phi(x_i) = \sum_{k=1}^K \varphi(k) * f_k(x_i), f_k \in F \quad (1)$$

where $\varphi(k)$ and $f_k(*)$ are the learning rate and function of the k -th CART, respectively, and x_i is the i -th leaf in the tree. Following the training of XGBoost, the learning model from the selected indicators becomes available to simulate the impact of multiple factors on MRT ridership, as shown in Equation (2).

$$Ridership = XGBoost_r(D; BE; T; G) \quad (2)$$

where D, BE, T, G are the indicators of land use, built environment, transit service, and demographics, respectively [39]. The detailed data and indicators are presented in Section 4.2. Owing to the goal of this work, only the land use factors are changeable and the other factors are regarded as fixed control variables. Under these circumstances, the maximum MRT is shown as Equation (3):

$$maxf_1 = XGBoost \left(\sum_{i \in I} X_{i1} D_{i1}, \sum_{i \in I} X_{i2} D_{i2}, \dots, \sum_{i \in I} X_{ik} D_{ik} \right) \{k \in K\} \quad (3)$$

2. Compactness

Compactness is a hot spot in urban planning and is significant for constructing resource-saving and environment-friendly cities [42,43]. The second objective aims to maximize compactness among land use cells, which improves land utilization efficiency and living convenience. There is no unified measurement of compactness owing to its diverse definitions and conceptual frameworks [44–46]. Inspired by previous TOD planning applications, this study measures compactness as the number of neighboring cells with the same land use type [9,15]. For example, the compactness of cell i with k -type land use can be computed by the number of cells assigned to k -type land use within the eight cells surrounding i . The objective is formulated as follows:

$$maxf_2 = \sum_{i \in I} \sum_{k \in K} \sum_{j \in N_i} X_{jk} X_{ik} \quad (4)$$

3. Land use conflict

Conflicts are commonly generated in adjacent land use cells; for example, industrial lands may bring about noise and pollution and lead to negative effects for people living in nearby residential lands. The third objective aims to minimize the conflict between different adjacent land parcels to improve the quality of life of residents around the station. Nevertheless, the conflict between adjacent cells is formidable to quantitatively describe, so an empirical indicator, conflict degree c_{kl} , is employed according to previous studies. In addition, considering that differently dense regions result in varying conflicts, this study introduces the identification of “adjacency” to the function establishment [47]. The objective is formulated as follows:

$$minf_3 = \frac{\sum_{i \in I} \sum_{k \in K} X_{ik} D_{ik} \sum_{j \in Adj_i} X_{jl} \frac{Side}{\sqrt{D_{ik}}} c_{kl}}{\sum_{i \in I} \sum_{k \in K} X_{ik} D_{ik}} \quad (5)$$

4. Land use mix

Mixed land use not only contributes to stimulating activities but also reduces travel costs and promotes travel by walking and bicycling, which is a significant feature for

compact and vital neighborhoods. The third objective aims to increase the degree of mixed land use to design a functionally diversified neighborhood consistent with TOD's basic concept. The degree of mixed land use is measured by the entropy index [48], and the objective is formulated as follows:

$$\max f_4 = \frac{-\sum_{k \in K} A_k \ln(A_k)}{\ln(N_{cell})} \quad (6)$$

where A_k is the density percentage of k -type land use in the station area; and $A_k = \sum_{i \in I} D_{ik} / \sum_{i \in I} \sum_{k \in K} D_{ik}$.

5. Environmental effects

The local environmental burden will be heavier as developed land use density and human activities increase, and reducing the environmental footprint in the precondition of TOD principles is important for sustainable planning [12]. By reference to previous studies, the fourth objective aims to reduce the pollution treatment cost, which reflects a decrease in the negative environmental influence of development and maintains environmental quality [13,14]. A high density of industrial land cells may increase the cost of pollution; thus, controlling the allocation of high-pollution industries is important for green planning. The treatment cost is highly correlated with the land use density and needs to consider various types of pollution. The minimum pollution treatment cost is formulated as follows:

$$\min f_5 = \sum_{i \in I} \sum_{k \in K} \sum_{r \in R} X_{ik} D_{ik} P_k^r \text{cost}^r \quad (7)$$

6. Destination accessibility

Destination accessibility, which is measured by the walkable distance from the origin to the transit station, is an essential principle of TOD [49]. This last objective aims to reduce the total walking distance of residents from the MRT station to their destination to support travel by MRT. Specifically, there are significant divergences in residents' attraction to various land uses. Highly attractive land use should be close to the MRT station, and dense development will create a high cost in terms of walking distance. The objective is formulated as follows:

$$\min f_6 = \sum_{i \in I} \sum_{k \in K} L_i X_{ik} D_{ik} \text{Attr}_k \quad (8)$$

3.2.4. Constraints

In addition to the above objectives, several constraints exist during the optimization process. First, as the minimum units in the study, each cell must be allocated to one type of land use, and mixed land use is not acceptable. Therefore, the land use type is encoded into a binary variable, X_{ik} , and the constraint is presented as follows:

$$X_{ik} = \{0,1\}, \forall i \in I, k \in K \quad (9)$$

$$\sum_{k \in K} X_{ik} = 1, \forall i \in I \quad (10)$$

Second, intensive development will exceed environmental capacity; therefore, land use density is limited by the exploitative intensity index, such as the floor area ratio provided by the Urban Management and Planning Ordinance. The land use density of each cell must be constrained as follows:

$$X_{ik} * \text{lowerEII}_k \leq \frac{X_{ik} D_{ik}}{\text{Area}} \leq X_{ik} * \text{upperEII}_k, \forall i \in I, k \in K \quad (11)$$

Finally, as regional planning varies across different station locations, such as planning for a commercial center or a cultural center, the percentage of the land use density of different land use types is constrained by the government's overall urban planning:

$$lowerRatio_k \leq \frac{\sum_{i \in I} D_{ik}}{\sum_{k \in K} \sum_{i \in I} D_{ik}} \leq upperRatio_k, \forall k \in K \quad (12)$$

3.3. Optimization Approach

Land use design is a multiobjective programming problem, and it is impossible to achieve the best value for all of the objectives because of the conflict among them. Optimization aims to find the Pareto optimal solutions of land use maps. Therefore, a method based on the nondominated sorting genetic algorithm III (NSGA-III)[32], which is skilled in searching for multiobjective optima without weights, is employed to generate alternative land use maps. Specifically, land use layout A is prior to being chosen only when A performs better than B among all objectives (A dominates B). Therefore, the land use layouts optimize the multiple objectives simultaneously and approach the optimum in the consecutive generations. Finally, a set of alternative land use layouts are obtained for further selection by decision-makers. The flowchart of the optimization model is shown in Figure 3, and the main operators are illustrated below.

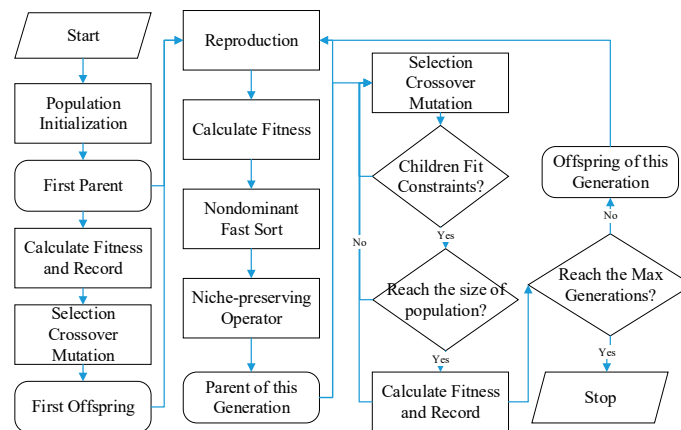


Figure 3. Flowchart of the optimization approach.

1. Representation and Initialization

In this work, we design a mixed coding for fitting the problem, where real coding is applied to describe the decision variables, and each undeveloped cell in the land use sketch map is transformed into two genes to represent the land use type (integer 0–4) and density (positive floating number). Consequently, the length of the chromosome in an individual is twice the number of undeveloped cells. In the initialization step, individuals are generated randomly, and unqualified individuals are removed by a constraining judgment. The initialization ends when the number of eligible individuals reaches the population size.

2. Fitness

Fitness is a crucial indicator for evaluating the performance of individuals and is the fundamental element for genetic operators. In this work, six fitness values that correspond to the objectives are calculated: $\{f1, f2, -f3, f4, -f5, -f6\}$.

3. Preselection

For each iteration of the reproduction, a set of candidates (Q_t) consists of both the parent (P_t) and offspring (C_t) in the last iteration, and the purpose of preselection is to choose the best and most diverse individuals (P_{t+1} with the population size) for the next

iteration. The procedure of the operator is shown in Figure 4, and there are two steps in the preselection.

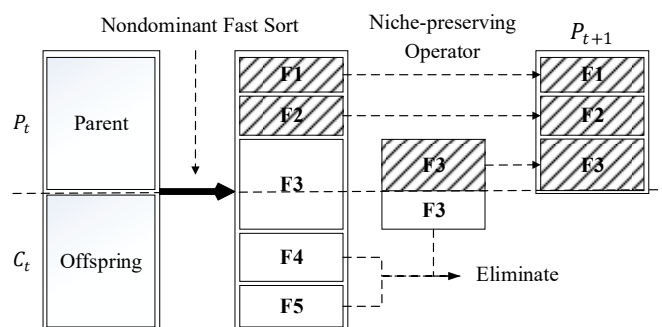


Figure 4. Illustration of the preselection operator.

Nondominant fast sort: This operator is used to classify individuals into several levels based on the dominant relation. As shown in Figure 4, F1 is the best level. Individuals in F1 are not dominated by any other individuals, and individuals in F2 are dominated only by individuals belonging to F1. Individuals at a higher level are more likely to be selected.

Niche-preserving operator: This operator is applied to choose individuals at the same level with the principle of individual diversity (such as F3 in Figure 4). The sketch map of the operator is shown in Figure 5. First, two-layer *reference points* (112) uniformly distributed on the unit simplex of the solution space are generated. Second, the *ideal point* is determined by the minimum value of each objective, and the *extreme points* are chosen as the points nearest to the axes. Third, the fitness of the objectives is normalized based on the ideal point and extreme point, and each *solution* is associated with the reference point whose *reference line* (the line that connects the reference point to the ideal point) is closest to the solution. Finally, when the reference point contains fewer solutions, the solutions associated with this reference point have a greater potential to be selected. Significantly, the reference point intrinsically represents the preference for objectives, which can be regarded as a novel “weight”, and the fitness with normalization displays a comparable “score” for all objectives, which helps further decision making.

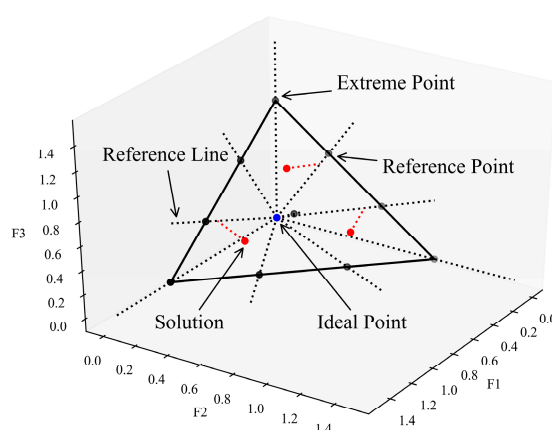


Figure 5. Sketch map of the niche-preserving operator.

4. Conventional Genetic Operators

As excellent individuals are selected by preselection, the model randomly selects pairs of parents from P_{t+1} . This work applies uniform crossover and uniform mutation

operators, and each gene decides whether to reciprocally interchange and mutate based on constants P_s and P_m . In this study, the genes of the land use type are directly exchanged in the crossover and are randomly mutated to integers (0–4), while the genes of land use density employ a simulated binary crossover and polynomial mutation in the genetic operators. Finally, each individual must satisfy the constraints.

5. Elitism and Termination

This work also introduces elitism to the model, which preserves the individuals with the best performance for each objective to force the solutions to approach the optima. The algorithm ends when the given maximum iteration is reached; otherwise, the algorithm returns to reproduction.

4. Case Study

4.1. Study Area

Shanghai is considered to be among the cities with a highly developed MRT system. To date, 16 lines with 416 stations are in operation, and the total length of the Shanghai Metro is 705 km. There is a substantial need for planning to integrate the MRT system and land use to relieve urban issues, and significantly, the government has adopted TOD as a crucial tool in land use planning. As a previous study found, the results of the TOD typologies are valuable for site selection [7]. In this study, the area around Fanghua Road Station is selected as the study area because the development of land at this station has been proven to lag behind the development of rail transit in the TOD typology [6], and the station has the potential to develop in coordination with land use optimization. Fanghua Road station is at the end of line 7 and near the subcenter of Century Park, which is located in the suburbs of Shanghai city (as shown in Figure 6a). The allocation of land use around the station is shown in Figure 6b, where residential regions occupy the largest area in the station catchment, and the diversity and density of land use are insufficient compared to MRT development. Specifically, a distance of between 400 m and 800 m (10-min walk time) is appropriate for TOD planning, and a square circumscribed in the 800 m radius of the station is used for planning in this case study, which is consistent with previous work [13]. Furthermore, the planned area is divided into 20×20 cells of uniform size; the size of each cell is $80 \text{ m} \times 80 \text{ m}$. On the one hand, we find that almost all the grids in this size contain a specific land use type, which distinguishes the different types of land cells and accurately describes the land use status. On the other hand, the grid size guarantees the efficiency of the solutions and generates appropriate land use sketch maps for regional planning.

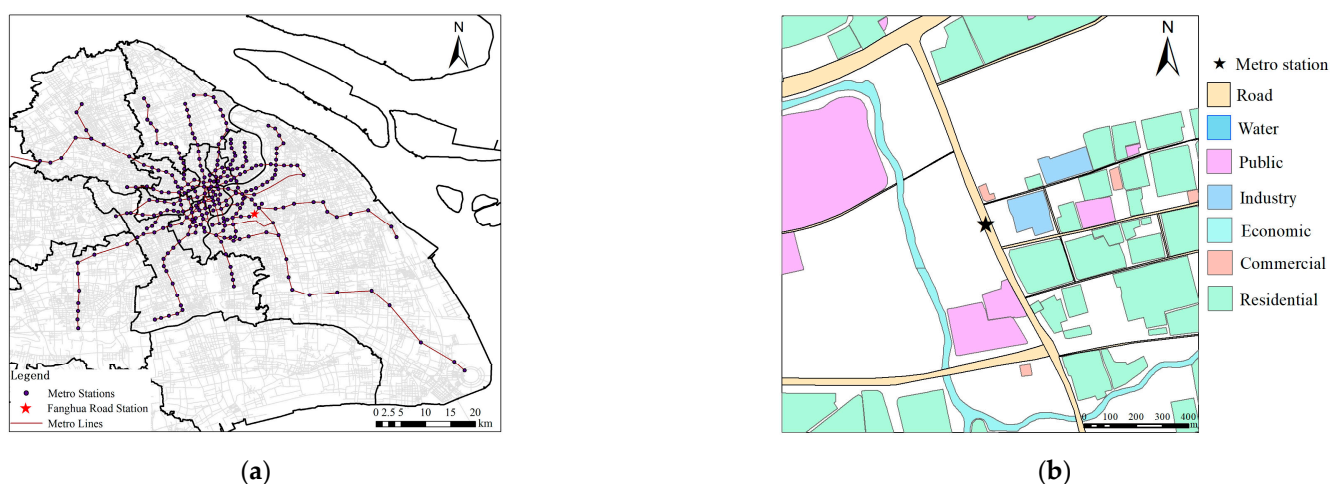


Figure 6. Map of the study area: (a) Distribution of the Shanghai Metro and Fanghua Road station; (b) Allocation of the land use around Fanghua Road Station.

4.2. Data Description and TOD Planning Model Construction

Geo-big data have a high spatial resolution, fast updating, and convenient access compared to traditional survey information and are suitable for describing complicated scenarios around stations. The proposed method uses both geographical data with a high spatial resolution and empirical and statistical data for objective construction.

Land use data: As substantial variables, land use allocation around the station is determined by the area of interest (AOI), which is a type of geographic data that comprises land parcels with exact types, locations, and outlines. Land use density is estimated mainly by the floor area ratio (FAR) from planning data and housing websites for residential land and building outlines with floor information for other types of land.

Data for ridership: Ridership is estimated by the inbound and outbound passenger flow using smart card data that contain abundant information on individual travel, including the user, time, stop, and line. In this study, the average workday ridership is calculated from 266 metro stations in January 2018, with a total of 258,896,211 records for modeling. The data that correspond to the variables of ridership are illustrated in Table 1 (buffer with 800 m).

Empirical data: Several parameters are employed in the planning model. An artificial indicator, conflict degree c_{kl} (0–8), is designed to estimate the conflict between adjacent parcels because quantitative evaluation involving numerous factors faces significant challenges [13,14]. The total cost of pollution treatment is calculated by pollution emissions P_k^r and unit cost $cost^r$ for different pollutant types, which is estimated by the relevant policy [50]. Passenger attraction $Attr_k$ determines the distribution of residents' destinations and is estimated according to the Technical Standards of the Traffic Impact Analysis of the Shanghai Construction Project [51].

Table 1. Data description for the variables of the ridership objective.

Categories	Variables	Source
Land use	Residential density	The area of interest (AOI) of Amap, building outlines with floors, and the floor area ratio from planning information and housing websites (such as https://m.ke.com/sh/ , accessed on 31 October 2020)
	Economic density	
	Commercial density	
	Public density	
	Industry density	
Built environment	Length of the street network	Open Street Map (OSM)
	Population density	Landscan 2018 [52]
	Number of road intersections	Open Street Map (OSM)
Transit service	Number of metro lines	The official website of Shanghai Metro
	Number of bus stops	The points of interest (POI) of Amap
	Terminal station (dummy)	The official website of Shanghai Metro
	Departure interval	The official website of Shanghai Metro
Demographics	Household	Housing websites
	Housing price	Housing websites
Ridership	MRT ridership (dependent variable)	Smart card data

The case study is conducted with the data mentioned above. The initial land use layout can be acquired based on the *land use data* by mapping the land use allocation of AOI data to grids and determining the density of the existing land use development from FARs. Then, the TOD planning model is established for optimization. With the count of the total (both boarding and alighting) ridership from smart card data and the preparation of the control variables and explanatory variables listed in Table 1, the XGBoost model is trained with *data for ridership* (all metro stations) for the first Objective of MRT ridership. The parameters of conflict degree c_{kl} , pollution emissions P_k^r and unit cost of different kinds of pollutants $cost^r$ are estimated from the *empirical data*, which

are used to build up the objectives of land use conflict, and environmental effects, respectively. To construct the objective of destination accessibility, passenger attraction $Attr_k$ are estimated with empirical data and the distances from land use cells to the metro station L_i are calculated based on the OSM. Finally, the objectives of the land use mix and compactness are calculated with the land use layout, which does not involve additional information.

4.3. Results and Analysis

4.3.1. Ridership Modeling

Fivefold cross-validation (5-CV) is applied to search for the optimal structure of XGBoost to ensure a robust model, and the total sample is divided into a training set with 80% of the data, and a test set with the other 20% of the data for evaluation. Consequently, XGBoost for nonlinear modeling is developed with a maximum of 10,000 CARTs and a learning rate of 0.005. In addition, the training is stopped early to improve the generalization of the model. The average goodness of fit (R^2) and root mean square error (RMSE) of the 5-CV are used to evaluate the model. Table 2 displays a comparison of the results of XGBoost and those of the linear model. The result of the ridership regression shows that XGBoost obtains a better result than the linear model in the training step, and 61% of the ridership of the test samples can be explained by the XGBoost model, which is slightly higher than that of linear regression. In addition, this study compares the result of XGBoost with another two popular tree models, random forest and gradient boosting decision tree (GBDT), which indicate that the employed method is fitter at depicting the objective of MRT ridership than the other tree models. Nonetheless, the accuracy of the test set does not achieve the expected value for several reasons: there is a lack of substantial impact factors, such as car ownership, because of data accessibility; the size of the samples is too small to support the XGBoost model; and the data have deviation and noise. In summary, the model provides an appropriate setting for constructing the ridership objective.

Table 2. Model performance of XGBoost, linear model, and tree models.

Objective	Model	R^2		RMSE	
		Training	Test	Training	Test
Ridership	Linear (OLS)	0.64	0.53	1160	2543
	Random Forest	0.77	0.48	869	3190
	GBDT	0.95	0.56	340	2484
	XGBoost	0.96	0.61	330	2340

4.3.2. Optimized Land Use Layouts

The NSGA-III is used for multiobjective optimization, and the population size is set to 112. The maximum number of generations is 2000 because it is necessary to search a large number of generations for optimal solutions in such a large solution space. The initial land use map is acquired by transforming the land use allocation in the planning area to 20×20 cells, which consists of 233 developed cells and 167 undeveloped cells (Figure 7a). Finally, 105 alternative non-dominated land use layouts are obtained after 2000 iterations. The extreme optimal solutions (best performance in six objectives) are selected to validate the effectivity of the optimization approach, and the employed method is compared with the elitist genetic-based algorithm (EGA) [53] and the nondominated genetic algorithm-II (NSGA-II) [31]. Table 3 shows that the optimal land use layouts obtained by NSGA-III achieve a maximum ridership of 129,418 persons per day (A), maximum compactness of 1520 (B), minimum land use conflict of 3.328 (C), maximum land use mixed degree of 0.696 (C), lowest pollution treatment cost of 8.917×10^5 RMB per year (D) and shortest walking distance of 2.557×10^8 meters (E). The employed method performs

better than the EGA in most objectives (ridership, land use mix, and pollution cost), suggesting that the extreme optimization of the model is as excellent as that of the EGA. However, the EGA is not powerful enough to solve multiobjective problems. In addition, the nondominant solution set of NSGA-III is compared to that of NSGA-II by a set coverage metric (C-metric) [54]. The comparison shows that 0.8% of the solutions in the Pareto solution set of the NSGA-III are dominated by the solutions in the Pareto solution set of the NSGA-II; and inversely, 2.7% of the solutions in the Pareto solution set of the NSGA-II are dominated by the solutions in the Pareto solution set of the NSGA-III, which shows that the applied model has remarkable convergence. Therefore, the NSGA-III is an appropriate optimization method with no empirical weights for the TOD planning model.

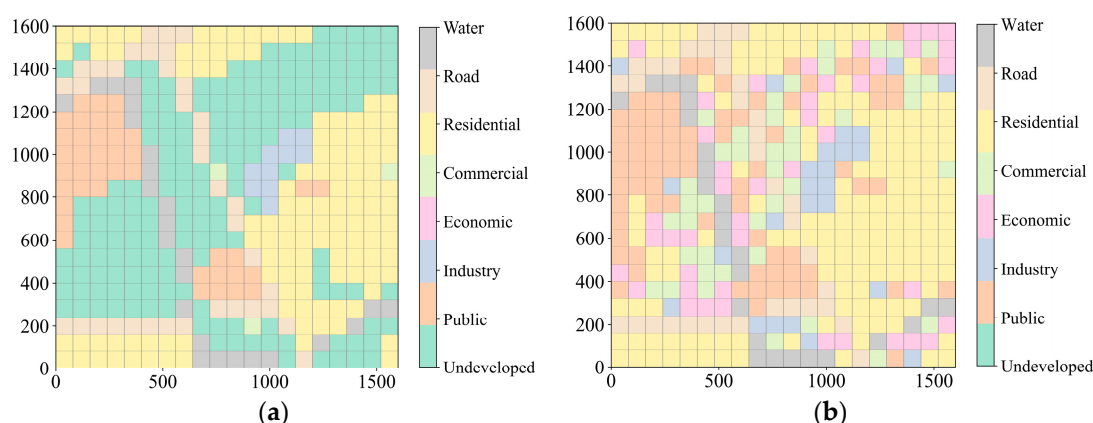


Figure 7. Land use sketch maps of initialization and optimization: (a) Initial land use maps of the station area; (b) Balanced optimized land use maps.

Table 3. Performance of the extreme optimal solutions of the NSGA-III and EGA.

Algorithm	Solution	Ridership (persons/day)	Compactness	Land Use Conflict	Land Use Mix	Pollution Cost (RMB/year)	Walking Distance (m)
NSGA-III	A	129,418	1258	4.739	0.259	4.971×10^6	8.810×10^8
	B	59,073	1520	3.579	0.242	2.447×10^6	5.586×10^8
	C	56,554	1506	3.328	0.237	2.120×10^6	4.653×10^8
	D	73,299	1196	5.493	0.268	7.047×10^6	7.120×10^8
	E	53,369	1354	4.127	0.227	8.917×10^5	4.699×10^8
	F	50,381	1308	5.895	0.224	5.432×10^6	2.557×10^8
EGA	A	126,273	1304	4.403	0.246	3.178×10^6	7.976×10^8
	B	56,381	1572	4.021	0.241	3.671×10^6	4.476×10^8
	C	66,794	1482	3.246	0.240	2.273×10^6	6.347×10^8
	D	66,886	1294	4.947	0.266	6.608×10^6	7.012×10^8
	E	56,950	1340	3.937	0.231	9.468×10^5	6.322×10^8
	F	44,823	1350	5.372	0.214	5.401×10^6	2.436×10^8

4.3.3. Further Selection: A Feasible Solution for the Study Station

As discussed above, several consistencies and conflicts among the objectives need to be considered by decision-makers. Although the extreme optimal solutions achieve the best performance for each objective, a balanced solution is more favorable for practical planning. According to the “comprehensive plan and general land use plan of Pudong New Area, Shanghai” [55], the Fanghua Road Station of rail transit line 7 is regarded as the core, and the construction of both cultural parks and cultural facilities is planned to improve the comprehensive support of regional services and maintain ecological sustainability. Therefore, the percentage of public land use has to be set to a high value to

satisfy this requirement, and the planning solution focuses more on land use conflict, land use mix, and pollution cost. Based on the purposes of the planning, an alternative balanced solution, which is selected from the aforementioned 105 alternative solutions, achieves the highest score of 0.455 after normalization and is associated with the reference point having the preference of (0.083333, 0.083333, 0.25, 0.25, 0.25, 0.083333) in each objective dimension. The solution for ridership is 58,616 persons per day, compactness is 1298, land use conflict is 4.742, land use mix is 0.256, pollution cost is 3.751×10^6 RMB per year, and total walking distance is 6.157×10^8 meters. Figure 7b shows the optimized land use sketch map and Figure 8 depicts the land use density of the selected solution around the station. Significantly, commercial land use is vigorously developed and mainly concentrated in the internal and intermediate areas of the station to attract the attention of residents. Economic land use is uniformly allocated in the planning area to offer plenty of employment. Residential land use is mainly aggregated in the periphery of the planning area, which is not influenced by the noisy station, and residents can walk home within a comfortable distance. Industrial land use is dispersed in the area with a small number of cells to maintain the land use mix and environmental protection. Finally, public land use including parks, cultural places, and public services, occupies a large space in the area, which is consistent with regional planning. The selected land use layout largely achieves the objectives of TOD planning and conforms to local political strategies [55], which provides suggestions regarding station-level land use development.

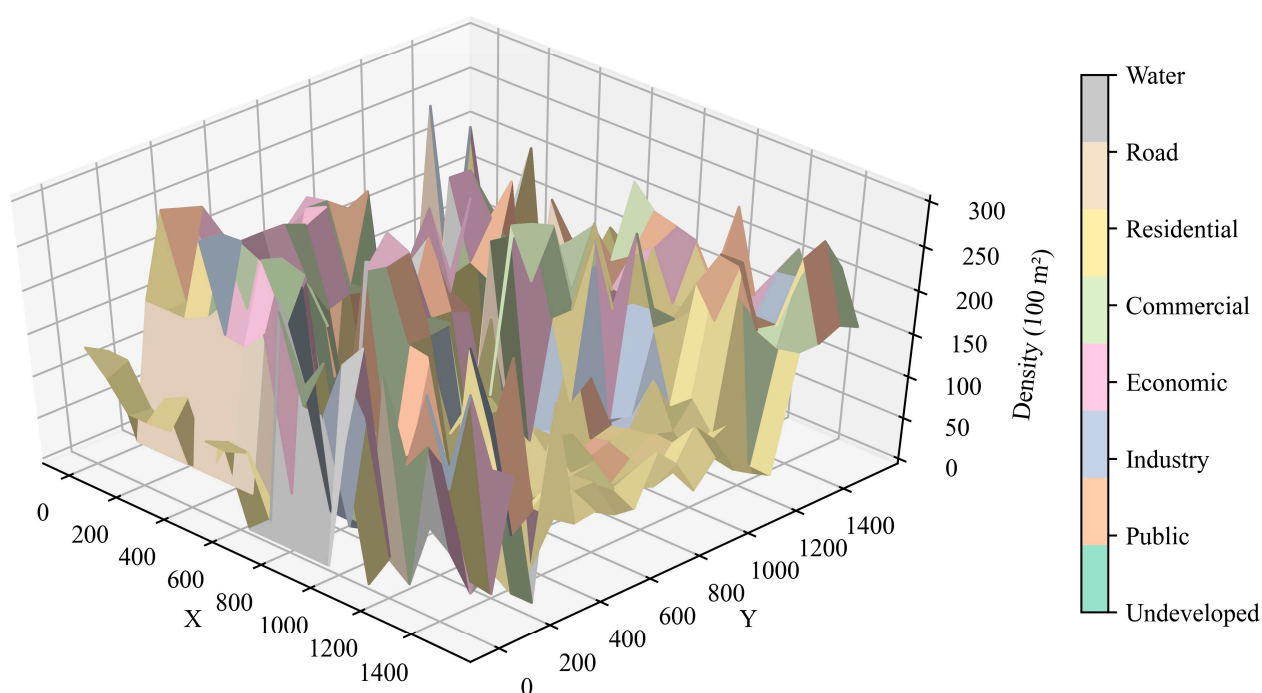


Figure 8. Optimized land use density of the station catchment.

5. Discussion

Analyzing the TOD planning objectives is imperative for integrated development of transportation and land use, and highlights valuable policy implications. The coefficient confusion matrix is constructed with the Pearson correlation coefficients of the objectives in pairs to seek out the relationship between the planning objectives. The correlation coefficients are shown in Figure 9 and several interesting findings are acquired.

Firstly, similarly to the results of previous studies [39,40], we find that a high degree of mixed land use results in an increase in ridership, suggesting that functionally diverse land use provides abundant space for various activities and generates substantial travel

demands for MRT systems. Therefore, promoting functional diversity for land use is crucial for planning the undeveloped or developing metro station catchments.

Secondly, high values of compactness enhance the local land use efficiency and decrease the conflict between adjacent land cells, but it negatively affects the degree of mixed land use. This is consistent with findings from a previous study [14]. For urban planners, it is important to keep a balance simultaneously between regional global land use functional diversity and compact local land use when making decisions for TOD planning.

Thirdly, the local pollution treatment cost increases as ridership increases because high-density land development both increases residents' travel demands and pollution emissions. This result has also been proved by a previous study [13]. Therefore, policies should advocate multi-mode public and low-carbon transportation and expand efforts to develop sustainable buildings, energy, and waste practices to construct green and livable TODs.

Finally, the total walking distance increases along with land use mix and ridership, as shown in Figure 9a, because high-density and high-diversity development raises passage flows. Furthermore, we calculate the average walking distance of each person (the total distance divided by ridership). Figure 9b shows the relationship between average personal walking distance and ridership (both of them are standardized by max-min normalization), which indicates that the average walking distance of each person does not increase, confirming that an appropriate land use allocation could shorten the walkable distance and enhance travel convenience to MRT stations, both of which increase the attractiveness of rail transit for residents. −0.2–

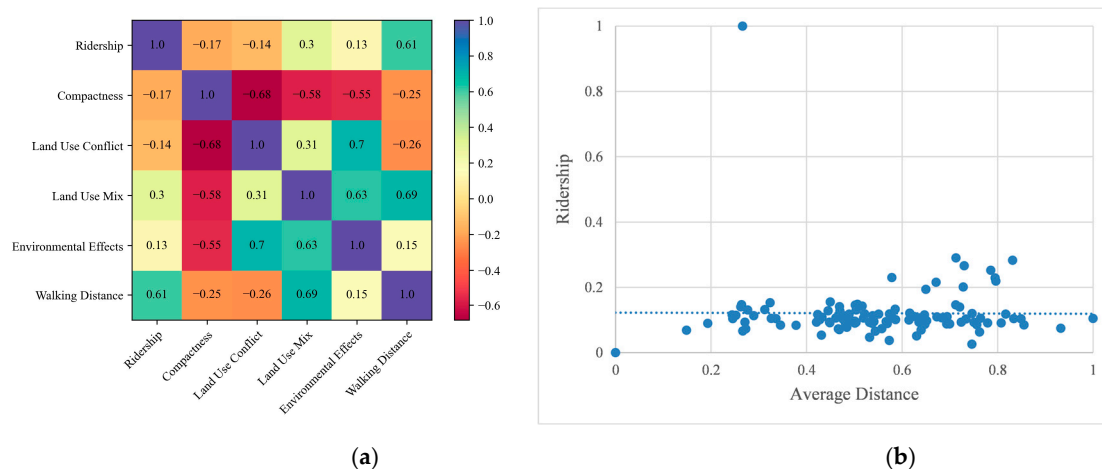


Figure 9. Relationships between the planning objectives: (a) Coefficient confusion matrix; (b) Relationship between the average walking distance and ridership.

6. Conclusions

This paper proposes an improved multiobjective TOD land use design framework to finely optimize the land use layout surrounding MRT stations and promote coordinated development between public transportation and land use in Chinese megacities. The method is constructed via a TOD planning model and a multiobjective optimization approach. Based on TOD strategies in the Chinese context, the planning model focuses on economic rail transit, efficient and functionally diverse land use, high-quality and convenient living, and a low-polluting environment. Specifically, the planning model introduces geo-big data and explores the practical nonlinear impacts of land use on MRT ridership. An improved genetic algorithm is used to solve the multiobjective programming model, and it generates land use schemes with specific land use allocations and density features without objective weighting. Significantly, the optimized approach pro-

vides objective preference information and normalized scores of the solutions for further selection to satisfy the local policies. The proposed method is validated in the case study of Shanghai. The results indicate that (1) the TOD planning model, constructed with geo-big data and statistical and empirical data with the consideration of nonlinear relationships between the land use and TOD objectives, is closer to the actual situation than other models; (2) The employed optimization approach is effective in solving the complicated TOD planning model and in generating a set of alternative land use schemes with favorable extreme optima and convergence. Furthermore, this study discovers several interesting findings based on analyzing the relationships between the TOD planning objectives, and provides meaningful suggestions for urban planning and policy-making. Most of the data in this framework are from open sources and the objective functions can be established; thus, the method is applicable for other Chinese megacities with some revisions (e.g., different study areas may use different grid sizes, model parameters, and constraints corresponding to local policies), which would provide valuable insights from a micro perspective for the future balanced development of MRT and station-level land use.

Although the proposed method solves problems regarding TOD planning, there are several features to be improved and extended in the future. Overall, this study applies fixed-dimension cells for designing, allocates a pure land use category to each cell, and does not consider the mixed uses of buildings, which reduces the diversity of the realistic urban environment. Therefore, it would be significant to incorporate different land parcel shapes and mixed-use buildings into planning. Secondly, the alternative land use sketch maps from the proposed model are scattered and discrete, which creates problems in practical urban planning. Accordingly, critical spatial constraints will be introduced to the model. Moreover, some objectives, such as environmental effects, are simplifications that construct relationships only with land use and do not consider other elements of the built environment; thus, combining with more factors that contribute to the objectives can be beneficial for improving the TOD planning model. The multiple factors change with the transformation of the land use, so more complicated dynamic models can be considered for planning as an extension of this study. Finally, it is difficult to practically measure the attraction of land use with empirical data, but mobile sensor data such as cellular signaling data offer the possibility of precisely estimating the distribution of passenger flow, which will be used in future works.

Author Contributions: Conceptualization, S.D. and Y.W.; methodology, S.D. and M.D.; validation, S.D.; data curation, S.D., Y.G., P.Z.; writing—original draft preparation, S.D.; writing—review and editing, S.D. and Y.W.; visualization, S.D., M.D., Y.G.; supervision, Y.W., J.G.; project administration, Y.W.; funding acquisition, Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Key Research Program of China (2016YFB0501403), the National Natural Science Foundation of China (Grant No. 41271399), China Special Fund for Surveying, Mapping and Geoinformation Research in the Public Interest (Grant No. 201512015).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data. Smart card data was provided by SODA 2018 (<http://soda.shdataic.org.cn/>, accessed on 15 August 2018). The other data source was described within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yang, J.; Quan, J.; Yan, B.; He, C. Urban rail investment and transit-oriented development in Beijing: Can it reach a higher potential? *Transp. Res. Part A Policy Pract.* **2016**, *89*, 140–150.

2. Chen, E.; Ye, Z.; Wang, C.; Zhang, W. Discovering the spatio-temporal impacts of built environment on metro ridership using smart card data. *Cities* **2019**, *95*, 102359.
3. Lyu, G.; Bertolini, L.; Pfeffer, K. Developing a TOD typology for Beijing metro station areas. *J. Transp. Geogr.* **2016**, *55*, 40–50.
4. Motieyan, H.; Mesgari, M.S. An Agent-Based Modeling approach for sustainable urban planning from land use and public transit perspectives. *Cities* **2018**, *81*, 91–100.
5. Zhang, Y.; Marshall, S.; Manley, E. Network criticality and the node-place-design model: Classifying metro station areas in Greater London. *J. Transp. Geogr.* **2019**, *79*, 102485.
6. Li, Z.; Han, Z.; Xin, J.; Luo, X.; Su, S.; Weng, M. Transit oriented development among metro station areas in Shanghai, China: Variations, typology, optimization and implications for land use planning. *Land Use Policy* **2019**, *82*, 269–282.
7. Ibraeva, A.; Correia, G.H.D.A.; Silva, C.; Antunes, A.P. Transit-oriented development: A review of research achievements and challenges. *Transp. Res. Part A Policy Pract.* **2020**, *132*, 110–130.
8. Motieyan, H.; Mesgari, M.S. Towards Sustainable Urban Planning Through Transit-Oriented Development (A Case Study: Tehran). *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 402.
9. Sahu, A. A methodology to modify land uses in a transit oriented development scenario. *J. Environ. Manag.* **2018**, *213*, 467–477.
10. Liu, L.; Zhang, M.; Xu, T. A conceptual framework and implementation tool for land use planning for corridor transit oriented development. *Cities* **2020**, *107*, 102939.
11. Lin, J.J.; Gau, C.C. A TOD planning model to review the regulation of allowable development densities around subway stations. *Land Use Policy* **2006**, *23*, 353–360.
12. Lin, J.; Li, C. A grey programming model for regional transit-oriented development planning. *Pap. Reg. Sci.* **2008**, *87*, 119–138.
13. Li, Y.; Guo, H.L.; Li, H.; Xu, G.H.; Wang, Z.R.; Kong, C.W. Transit-oriented land planning model considering sustainability of mass rail transit. *J. Urban Plan. Dev.* **2010**, *136*, 243–248.
14. Ma, X.; Chen, X.; Li, X.; Ding, C.; Wang, Y. Sustainable station-level planning: An integrated transport and land use design model for transit-oriented development. *J. Clean. Prod.* **2018**, *170*, 1052–1063.
15. Huang, X.; Liang, Q.; Feng, Z.; Chai, S. A TOD Planning Model Integrating Transport and Land Use in Urban Rail Transit Station Areas. *IEEE Access* **2021**, *9*, 1103–1115.
16. Calthorpe, P. *The Next American Metropolis: Ecology, Community, and the American Dream*; Princeton Architectural Press: New Jersey, USA, 1993.
17. Cervero, R. *The Transit Metropolis: A Global Inquiry*; Island press: Washington, D.C., USA, 1998.
18. Cervero, R.; Kockelman, K. Travel demand and the 3Ds: Density, diversity, and design. *Transp. Res. Part D Transp. Environ.* **1997**, *2*, 199–219.
19. Cervero, R.; Murakami, J. Rail + Property Development: A Model of Sustainable Transit Finance and Urbanism. Institute of Transportation Studies, UC Berkeley, 2008. <https://escholarship.org/uc/item/6jx3k35x> (accessed on 31 October 2020)
20. Higgins, C.D.; Kanaroglou, P.S. A latent class method for classifying and evaluating the performance of station area transit-oriented development in the Toronto region. *J. Transp. Geogr.* **2016**, *52*, 61–72.
21. Dou, M.; Wang, Y.; Dong, S. Integrating Network Centrality and Node-Place Model to Evaluate and Classify Station Areas in Shanghai. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 414.
22. Cao, K.; Batty, M.; Huang, B.; Liu, Y.; Yu, L.; Chen, J. Spatial multi-objective land use optimization: Extensions to the non-dominated sorting genetic algorithm-II. *Int. J. Geogr. Inf. Sci.* **2011**, *25*, 1949–1969.
23. Lin, J.J.; Feng, C.M. A bi-level programming model for the land use—Network design problem. *Ann. Regional Sci.* **2003**, *37*, 93–105.
24. Yim, K.K.W.; Wong, S.C.; Chen, A.; Wong, C.K.; Lam, W.H.K. A reliability-based land use and transportation optimization model. *Transp. Res. Part C Emerg. Technol.* **2011**, *19*, 351–362.
25. Levi, Y.; Bekhor, S.; Rosenfeld, Y. A multi-objective optimization model for urban planning: The case of a very large floating structure. *Transp. Res. Part C Emerg. Technol.* **2019**, *98*, 85–100.
26. Abdul-Rahman, M.; Chan, E.H.W.; Wong, M.S.; Irekponor, V.E.; Abdul-Rahman, M.O. A framework to simplify pre-processing location-based social media big data for sustainable urban planning and management. *Cities* **2021**, *109*, 102986.
27. Kandt, J.; Batty, M. Smart cities, big data and urban policy: Towards urban analytics for the long run. *Cities* **2021**, *109*, 102992.
28. Zhou, J.; Yang, Y. Transit-based accessibility and urban development: An exploratory study of Shenzhen based on big and/or open data. *Cities* **2021**, *110*, 102990.
29. van Wee, B.; Handy, S. Key research themes on urban space, scale, and sustainable urban mobility. *Int. J. Sustain. Transp.* **2016**, *10*, 18–24.
30. Holland, J.H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*; MIT Press: Massachusetts, United States, 1992.
31. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T.A.M.T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197.
32. Deb, K.; Jain, H. An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach, Part I: Solving Problems With Box Constraints. *IEEE Trans. Evol. Comput.* **2014**, *18*, 577–601.
33. Stewart, T.J.; Janssen, R.; van Herwijnen, M. A genetic algorithm approach to multiobjective land use planning. *Comput. Oper. Res.* **2004**, *31*, 2293–2313.

34. Schwaab, J.; Deb, K.; Goodman, E.; Lautenbach, S.; van Strien, M.J.; Grêt-Regamey, A. Improving the performance of genetic algorithms for land-use allocation problems. *Int. J. Geogr. Inf. Sci.* **2018**, *32*, 907–930.
35. Feng, C.; Lin, J. Using a genetic algorithm to generate alternative sketch maps for urban planning. *Comput. Environ. Urban Syst.* **1999**, *23*, 91–108.
36. Li, X.; Parrott, L. An improved Genetic Algorithm for spatial optimization of multi-objective and multi-site land use allocation. *Comput. Environ. Urban Syst.* **2016**, *59*, 184–194.
37. Wey, W.; Zhang, H.; Chang, Y. Alternative transit-oriented development evaluation in sustainable built environment planning. *Habitat Int.* **2016**, *55*, 109–123.
38. Elgar, I.; Kennedy, C. Review of optimal transit subsidies: Comparison between models. *J. Urban Plan. Dev.* **2005**, *131*, 71–78.
39. Ding, C.; Cao, X.; Liu, C. How does the station-area built environment influence Metrorail ridership? Using gradient boosting decision trees to identify non-linear thresholds. *J. Transp. Geogr.* **2019**, *77*, 70–78.
40. Gan, Z.; Yang, M.; Feng, T.; Timmermans, H.J.P. Examining the relationship between built environment and metro ridership at station-to-station level. *Transp. Res. Part D Transp. Environ.* **2020**, *82*, 102332.
41. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016.
42. Cao, K.; Huang, B.; Wang, S.; Lin, H. Sustainable land use optimization using Boundary-based Fast Genetic Algorithm. *Comput. Environ. Urban Syst.* **2012**, *36*, 257–269.
43. Yuan, H.; He, Y.; Zhou, J.; Li, Y.; Cui, X.; Shen, Z. Research on compactness ratio model of urban underground space and compact development mechanism of rail transit station affected area. *Sustain. Cities Soc.* **2020**, *55*, 102043.
44. Newman, P.; Kenworthy, J. *Sustainability and Cities: Overcoming Automobile Dependence*; Island Press: Washington, D.C., USA, 1999.
45. Ewing, R.; Pendall, R.; Chen, D.D.T. Measuring Sprawl and Its Impact. Volume I. 2002. Available online: <http://smartgrowthamerica.us/documents/MeasuringSprawlTechnical.pdf> (accessed on 31 October 2020).
46. Kotharkar, R.; Bahadure, P.; Sarda, N. Measuring Compact Urban Form: A Case of Nagpur City, India. *Sustainability* **2014**, *6*, 4246–4272.
47. Eom, S.; Suzuki, T.; Lee, M. A land-use mix allocation model considering adjacency, intensity, and proximity. *Int. J. Geogr. Inf. Sci.* **2020**, *34*, 899–923.
48. Manaugh, K.; Kreider, T. What is mixed use? Presenting an interaction method for measuring land use mix. *J. Transp. Land Use* **2013**, *6*, 63–72.
49. Ewing, R.; Cervero, R. Travel and the Built Environment. *J. Am. Plan. Assoc.* **2010**, *76*, 265–294.
50. Standards for Environmental Protection Tax, 2017. Available online: <http://www.shanghai.gov.cn/nw2/nw2314/nw2319/nw12344/u26aw54490.html?date=2017-12-22> (accessed on 31 October 2020).
51. Technical Standards of Traffic Impact Analysis of Shanghai Construction Project, 2015. Available online: <http://zjw.sh.gov.cn/jsgl/20180912/0011-26053.html> (accessed on 31 October 2020).
52. Rose, A.N.; McKee, J.J.; Urban, M.L.; Bright, E.A.; Sims, K.M.; LandScan Global 2018 [Data set].. Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2019. <https://doi.org/10.48690/1524213>
53. Rudolph, G. Convergence analysis of canonical genetic algorithms. *IEEE Trans. Neural Netw.* **1994**, *5*, 96–101.
54. Deb, K.; Thiele, L.; Laumanns, M.; Zitzler, E. In *Scalable Test Problems for Evolutionary Multiobjective Optimization*; Abraham, A., Jain, L., Goldberg, R., Eds.; Springer London: London, UK, 2005.
55. Comprehensive Plan and General Land-Use Plan of Pudong New Area, Shanghai, 2019. Available online: <http://www.shanghai.gov.cn/nw2/nw2314/nw2319/nw12344/u26aw63249.html?phlnohdjmglngdbi> (accessed on 31 October 2020).