Habitat International 150 (2024) 103120

Contents lists available at ScienceDirect

Habitat International

journal homepage: www.elsevier.com/locate/habitatint

Destigmatizing urban villages by examining their attractiveness: Quantification evidence from Shenzhen

Jin Rui[a,*], Yuhan Xu[b], Xiang Li[c]

[a] Department of Spatial Planning, Technical University Dortmund, August-Schmidt-Straße 10, 44227, Dortmund, Germany
[b] School of Architecture, Southeast University, 2 Sipailou, Nanjing, 210096, China
[c] School of Architecture, Urban Planning Department, The University of Hong Kong, 18 Fo Chun Road, Tai Po, Hong Kong

ARTICLE INFO

Keywords:
Destigmatization
Urban villages
Mobility patterns
Job-housing patterns
Spatial attractiveness

ABSTRACT

While existing social and political research has qualitatively discussed the stigmatization of urban villages (UVs), there's a lack of quantitative support. This study innovatively attempts to provide quantitative evidence, employing a spatial perspective for the destigmatization of UVs. We integrated population trajectory, GDP data, employment and residential points of interest, and developed a spatial gravity model to derive the spatial attractiveness index. Concurrently, we selected environmental, spatial structure, and housing economic indicators to construct a regression analysis. Our findings revealed that urban villagers exhibit a preference for shorter daily commutes, underscoring the positive role of UVs in promoting job-housing balance and providing diverse services. The high-frequency mobility of villagers accelerates the categorization of UVs and enhances intra-area circulation. We further discussed the interaction between UV renewal and its destigmatization. To counter spatial stigmatization, enhancing the spatial quality and infrastructure of UVs, as well as promoting diversified land use can reduce the public's stereotypical impressions. Moreover, prioritizing affordable housing and equitable distribution of facilities, along with fostering synergies between urban capital and UVs, can facilitate the destigmatization process. Stigmatization in informal settlements is a widespread issue. Our quantitative approaches, as well as targeted renovation and policy recommendations, can serve as a blueprint for addressing similar challenges in informal settlements globally.

1. Introduction

Driven by capital accumulation, urbanization consolidates labor, capital, technology, and markets within cities. This process fosters the rapid regeneration of labor while capitalizing on spatial imbalances in production to further boost capital accumulation (Wu et al., 2006). However, with limited spatial resources, this intensive accumulation inevitably leads to significant urban housing challenges, most notably the rise of large-scale informal settlements (Gouverneur, 2014). These settlements have become an integral part of the urban landscape in developing countries. Currently, a quarter of the world's population lives in various types of informal settlements (United Nations, 2017).

In pursuit of efficiency, Chinese urbanization chose to bypass traditional rural areas. This led to villages being gradually encircled by newly developed urban zones, resulting in informal residential areas known as "Urban Villages" (UVs) (Rui, 2023a; Wang et al., 2009). China's dual urban-rural economic and social management system is the root cause of the emergence of UVs (Liu & Zhang, 2020; Yang & Qian, 2022). China's household registration system was initially formed in the early 1950s, categorizing the population into agricultural and non-agricultural households (Meng, 2019; Zhang, 2023). The migration of the agricultural population to urban areas and their transformation into a non-agricultural population are a major characteristic of the urbanization process (Zhang, 2023). Since 1978, the government began to loosen controls on the mobile population, and from the 1990s onwards, a large-scale migration of rural agricultural population started moving to cities. Due to the central location of UVs in the city and the provision of affordable housing, many migrants reside in these areas (Pan & Du,

Abbreviations: UV, urban village; ATI, attractiveness index; CBD, central business district; AD, administrative district; POI, point of interest; AOI, area of interest; NDVI, Normalized Difference Vegetation Index; FAR, floor area ratio; BCR, building coverage ratio; IDW, Inverse Distance Weighting; OLS, Ordinary Least Squares; MGWR, Multiscale Geographically Weighted Regression; OD, origin destination; RF, Random Forest.

* Corresponding author.
E-mail addresses: jin.rui@tu-dortmund.de (J. Rui), xuyh_seu@163.com (Y. Xu), yennergy@hku.hk (X. Li).

https://doi.org/10.1016/j.habitatint.2024.103120
Received 17 October 2023; Received in revised form 22 April 2024; Accepted 6 June 2024

2021b). Before the 1990s, these UVs were considered to diminish the urban imageability due to poor environmental quality, safety hazards, outdated facilities, illegal constructions, and heterogeneity among the resident groups (Wu et al., 2013). It can be stated that the implementation of the household registration system led to urban-rural opposition and division, fostering geospatial group consciousness and local protectionism.

Beyond the structural exclusion of the household registration system, implicit social exclusion and stigmatization are among the many challenges faced by rural migrants living in cities (Meng, 2019). Research on the issue of stigmatization is concentrated in psychology (Demirtaş-Madran, 2020) and sociology (Zhang et al., 2021). From the perspective of spatial-cultural interpretation, Firey (1945) proposed that space combined with symbolic social value becomes an essential component in the local cultural system. Beyond its use value and exchange value, space is also endowed with "symbolic" value. Informal residential areas, represented by UVs, are given various labels, with the majority being derogatory, such as slums, illegal constructions, chimney buildings, and shaking hand buildings. This series of spatial discourses realizes the symbolic construction and widespread stigmatization of the living spaces of rural migrants—UVs. The stigmatization of UVs results from an interplay between spatial stigma and the social identity stigma of the residents, leading to a dual predicament of spatial segregation and stigmatization.

Previous studies focused on the demolition and regulation of UVs (Zeng et al., 2019). As UVs began to demonstrate their roles in terms of location, culture, welfare, and social integration during the urbanization process, the academic perspective on UVs started to shift. Instead of advocating for their immediate demolition, there were growing suggestions towards appropriate transformation (Pan & Du, 2021a). Therefore, the current research on urban UVs focuses on elucidating the positive value across spatial, social, economic, and political dimensions and exploring a multitude of transformation policies, strategies, and governance systems (Wei et al., 2020; Yu et al., 2019). Urban space (including UVs) serves as an important platform for social and economic development (Gehl, 2013). Research from a social dimension suggests that the management system of UVs constitutes "semi-formalized" community governance, promoting the integration of rural migrants with various urban stakeholders. From an economic perspective, studies indicate that UVs not only provide affordable rental housing but also play a positive role in urban consumption, job-housing relationships, income augmentation, and potentially facilitating social cohesion (Wu et al., 2020). Transitioning from disorganized rental configurations to well-structured communities with multifaceted functions, the significance of UVs has gradually been recognized (Gu & Zhang, 2021; Pan & Du, 2021a). However, regarding spatial structure and layout, the impact of "stigmatization" still persists, leading to a distinctive characteristic in UV transformation. For instance, resettlement areas based on UVs were treated as "second-class communities", creating spatial disjunctions with surroundings. Therefore, there is a need to provide more quantitative evidence on spatial inequality for destigmatization.

Big data has been widely used as decision support in several domains, enabling a better understanding of behavioural patterns (Gong et al., 2024; Jiang et al., 2022; Li et al., 2022; Rui, 2023b). It provides support for the identification and classification of UVs, as well as the quantification of population mobility and environmental characteristics of UVs. Further exploring the relationship between the spatial attractiveness of UVs and environmental indicators could provide evidence for destigmatizing UVs. Regarding indicator selection, Ewing and Cervero (2001) proposed the 5D framework (i.e., Density, Design, Diversity, Destination accessibility, and Distance), which has been applied to several fields, including transit-oriented development (Niu et al., 2021) and low-carbon initiatives (Zhang et al., 2020). In addition to this, some scholars have used land use mix, building coverage ratio, and street intersections to characterize the physical built environment (Jiang et al., 2022; Li et al., 2022). The absence of a unified built environment standard might be due to significant differences in research aims and subjects. Our study does not aim to outline a comprehensive built environment assessment framework. Instead, it selects from existing environmental indicators to support destigmatization-oriented UV regeneration.

Based on the behaviour-space interaction theory (Chai et al., 2017), this study provides quantitative evidence for the "destigmatization" of UVs (Fig. 1). The foundation of the spatial behaviour interaction theory encompasses the social-spatial theory as epistemology, behavioural geography and time geography as methodologies, and urban planning and management methods as practical guidance. The theory posits that the external environment influences individual behaviour choices and individuals behaviours have the potential to modify the space. Aggregating human spatio-temporal and behaviours spatially allows for an effective understanding and analysis of urban space. Moreover, correlating human behavioural choices with spatial indicators clarifies the mechanisms of space acting upon individuals, further elucidating the interactive relationship between space and behaviour. Specifically, this study utilizes mobile phone data to quantify the weekday mobility patterns of UV residents. Based on mobility patterns, a clustering model is employed to determine the job-housing types in UVs. Next, by refining the urban gravity model, the spatial attractiveness index (ATI) for work-oriented and living-oriented UVs is derived, and factors influencing the ATI of UVs are explored from a spatial perspective. The research conclusions offer recommendations for the UV destigmatization, addressing aspects of village environmental design, job-housing structure and layout, and government policies.

2. Data and methodology

2.1. Study area

Shenzhen, located in Guangdong Province, stands as one of China's first-tier cities (Fig. 2 b). Shenzhen can be broadly categorized into the core areas comprising Nanshan, Futian, and Luohu (District), the suburban regions encompassing Bao'an, Longhua, Yantian, and Longgang, and the outlying suburbs spanning Guangming, Pingshan, and Dapeng. Shenzhen has 1877 UVs, accommodating over 8 million residents (Rui, 2023a). With the continuous growth of the mobile population, there has been an increasing demand for affordable housing UVs, characterized by their low cost of living, are strategically positioned to cater to the requirements of this population segment. The Shenzhen Planning and Natural Resources Bureau (2019) released the "*Shenzhen Urban Village Comprehensive Improvement Master Plan (2019–2025)*", which underscores a shift in UV renewal policies from large-scale demolition to fine-grained transformation. This plan serves as guidelines for the latest UV renewal initiatives.

We classified an area as a UV when UV's building area occupied over 80% of the grid (Fig. 2 c). A total of 1478 UV grids were identified and marked in pink. The blue stars represent the central business district (CBD) of diverse administrative districts (ADs).

2.2. Data sources and processing

This study drew upon seven types of data, including mobile phone data, point of interest (POI) data, Normalized Difference Vegetation Index (NDVI), streets, building footprints, housing economic indicators, and GDP data. The data sources and collection date are presented in Table 1, while all data interpretations and descriptive statistics, including maximum, minimum, average values, and standard deviations, are displayed in Table A 3.

2.2.1. Independent variable cluster I: Village environment characteristics

We used floor area ratio (FAR) and building coverage ratio (BCR) to assess settlement density (Xia et al., 2020). FAR captures the ratio of total floor area to building footprint. BCR is the ratio of the total building

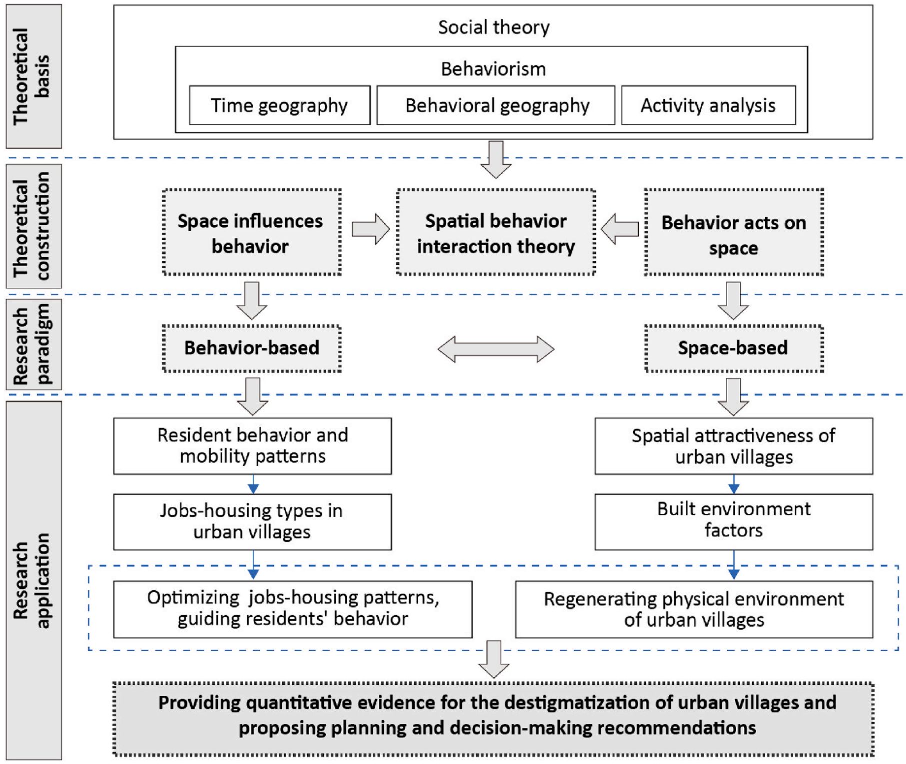

Fig. 1. Theoretical framework.

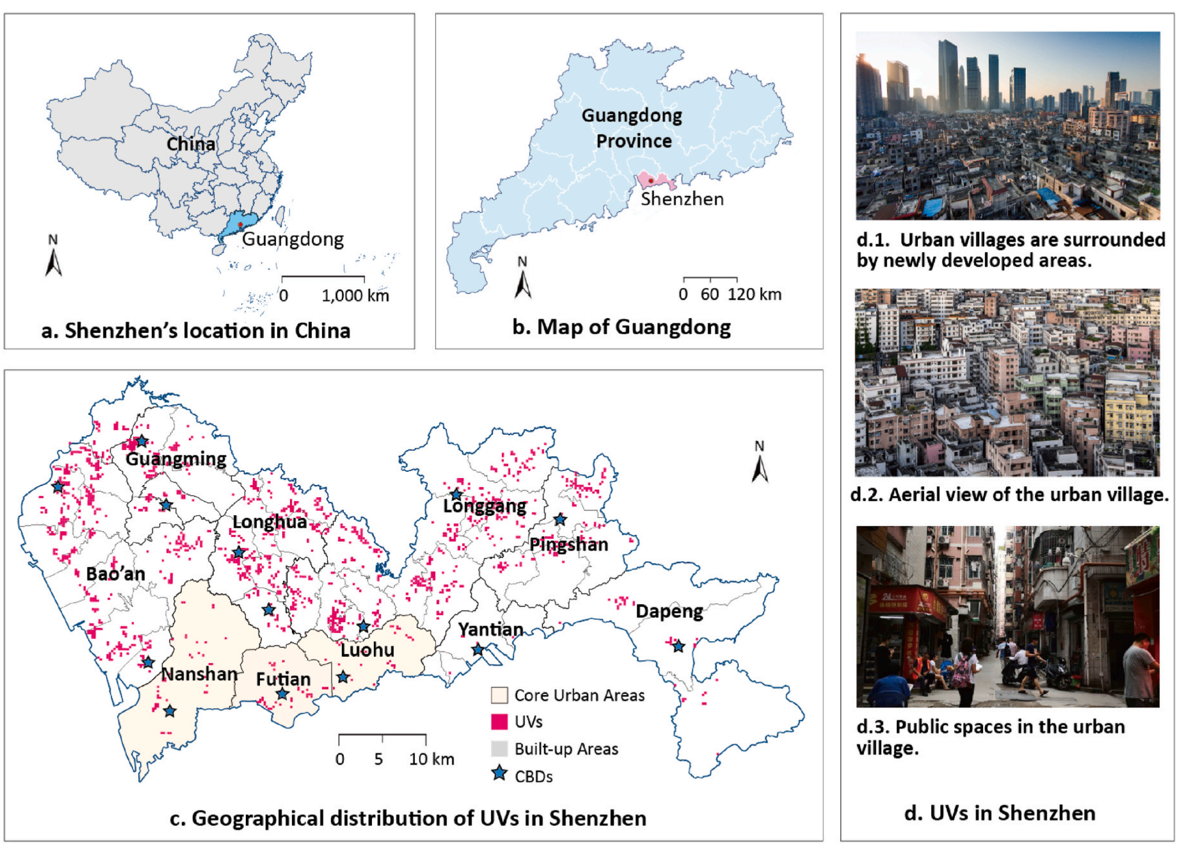

Fig. 2. Study area.

Table 1
Data and data sources.

Data	Sources	Collection date
Mobile phone data	China Mobile Communications Corporation (CMCC) (https://www.chinamobileltd.com/).	October 24, 2022
POI data	AMap (https://www.amap.com/)	January 6, 2023
NDVI	NASA's Earth Observing System Data and Information System Reverb (http://reverb.echo.nasa.gov/)	January 5, 2023
Street	OpenStreetMap (https://www.openstreetmap.org)	January 6, 2023
Building footprint	Shenzhen Municipal Government Open Data Platform (https://opendata.sz.gov.cn/data/dataSet/toDataDetails/29200_00300237)	January 5, 2023
Housing economic indicators	Lianjia (https://m.lianjia.com/) Anjuke (https://www.anjuke.com/)	October 28, 2022
GDP data	Resource and Environment Science and Data Center (https://www.resdc.cn/doi/doi.aspx?DOIid=33)	October 25, 2022

footprint in the neighborhood to the total area of the neighborhood.

$$FAR_i = \frac{F_i}{A_i} \tag{1}$$

$$BCR_i = \frac{BA_i}{A_i} \tag{2}$$

where F_i denotes the gross floor area of study unit i (m^2), A_i denotes the area of study unit i (m^2) and BA_i denotes the gross building footprint of study unit i (m^2).

NDVI, a remote sensing index, provides a measure of the relative intensity of vegetation cover and growth, making it a reliable indicator for assessing urban greenery. When NDVI is close to 1, it means there is abundant green vegetation. When NDVI is close to 0, it means there is no vegetation cover. POI refers to a geographical location, classified based on its relevance to a specific field of study. We obtained POI data from AMap (https://lbs.amap.com/) and classified them into distinct categories. The classification criteria are outlined in Table A 1.

2.2.2. Independent variable cluster II: Spatial structure and layout

Spatial structure and layout include six sub-variables: street network density, street intersection density, bus station density, metro station density, distance to city center, and accessibility. Street network density represents the overall length of streets within each grid. Street intersection density quantifies the number of street intersections within each grid. Distance to the city center indicates the straight-line distance between each grid and the city center. The city center was defined as the Shenzhen Municipal Government. The Two-Step Floating Catchment Area (2SFCA) method was employed to analyze accessibility.

2.2.3. Independent variable cluster III: Housing economic indicators

The housing economic indicators encompass five sub-variables: housing price, building height, community age, average housing area, and rental price. Building height, community building height, community age, and average price were derived by calculating the mean values of the sampling points within each grid. To estimate housing price and rental prices, we employed the Inverse Distance Weighting (IDW) method. This method determined the linear optimum for the sampling points within the community boundary, utilizing a search area of 12 points in the neighborhood and a search radius of 1500 m. The housing price data refers to the price of the entire property set, while rental data represents the monthly rent per square meter, accounting for the unit area. As a result, there is no overlap between these two variables, and the Ordinary Least Squares (OLS) results confirmed the absence of covariance.

2.3. Methodology and model architecture

Our methodology encompasses four steps (Fig. 3). First, we leveraged mobile phone data to map the travel patterns of urban villagers and employed clustering models based on work and sleep hours to categorize the job-housing patterns of UVs. We modified the urban gravity model and calculated the ATI as our dependent variable. Second, we selected village environmental variables, including village environment characteristics, spatial structure and layout, and housing economic indicators, as our independent variables. Third, we explored the correlations between village environmental variables and the ATI using the OLS model. To mitigate the potential biases stemming from spatial effects, we refined our selection of independent variables based on their value and VIF in the OLS model and adopted the Multiscale Geographically Weighted Regression (MGWR) model. To analyze numerical variations, we incorporated a nonlinear regression model using Random Forest (RF). By employing a multi-model-based collaborative interpretation, we discerned both geographical and numerical correlations between variables. Finally, this study identified the job-housing patterns and spatial attractiveness of UVs. Based on this, recommendations for destigmatization.

2.3.1. Cluster analysis of UVs based on personal travel patterns

We acquired the mobile phone data, which encompasses user IDs, start-end locations, and the duration of each stay. To ensure utmost privacy protection for the users involved, all user IDs have been encrypted and cannot be traced. Furthermore, the origin-destination (OD) location information has been discretized into square grids with 250-m sides instead of latitude and longitude. Therefore, we created study units with a side length of 250 m.

To discern the job-housing types of UVs, we extracted the month-long mobile phone data in October 2022. We selected data on October 24, 2022, for individual travel pattern analysis. We adopted an approach based on individuals' working and living locations and duration. We tracked a total of 779,384 data for regular users for identification of the working population. Specifically, individuals who remain in a fixed grid for at least 6 h between 6 a.m. and 8 p.m., continuously for a minimum of 4 days, are classified as working population. To capture the job-housing patterns, we introduced the working-living ratio (F) as a key feature. The subscript "i" refers to the UV grid, WD_i represents the number of individuals in the working population residing in grid i, and LO_i denotes the number of individuals in the living population residing in grid i.n Eq (3), utilizing F_i as the parameter in our k-means clustering model, we determine the appropriate number of clusters (k) using the elbow method (Figure A 3). Our analysis reveals that the sum of squared errors (SSE) stabilizes when k exceeds 3, leading us to categorize the UVs into three groups.

$$F_i = \frac{WD_i - LO_i}{WD_i + LO_i} \tag{3}$$

2.3.2. Calculation of the ATI

The urban gravity model serves as a tool for analyzing urban mobility, drawing upon the gravitational principle that cities' mutual attraction is linked to their population size or economic magnitude, while inversely related to the distance. Initially, the urban gravity model (Eq (4)) was formulated to calculate the gravitational value T_{ij} between two cities, with Q_i and Q_j representing the masses of cities i and j, and d representing the distance between them. As the model evolved, additional parameters were incorporated (Zhao et al., 2021). In urban analyses, Q_i and Q_j are replaced with P_i and G_i, denoting the population and GDP. To calculate the ATI, we modified the urban gravity model (Eq (5)). We incorporated the variable $P2_i$, representing the number of POI at location i. For working-oriented urban villages (WOUVs), POI encompasses the sum of industrial and commercial POIs, while for living-oriented urban villages (LOUVs), it corresponds to residential

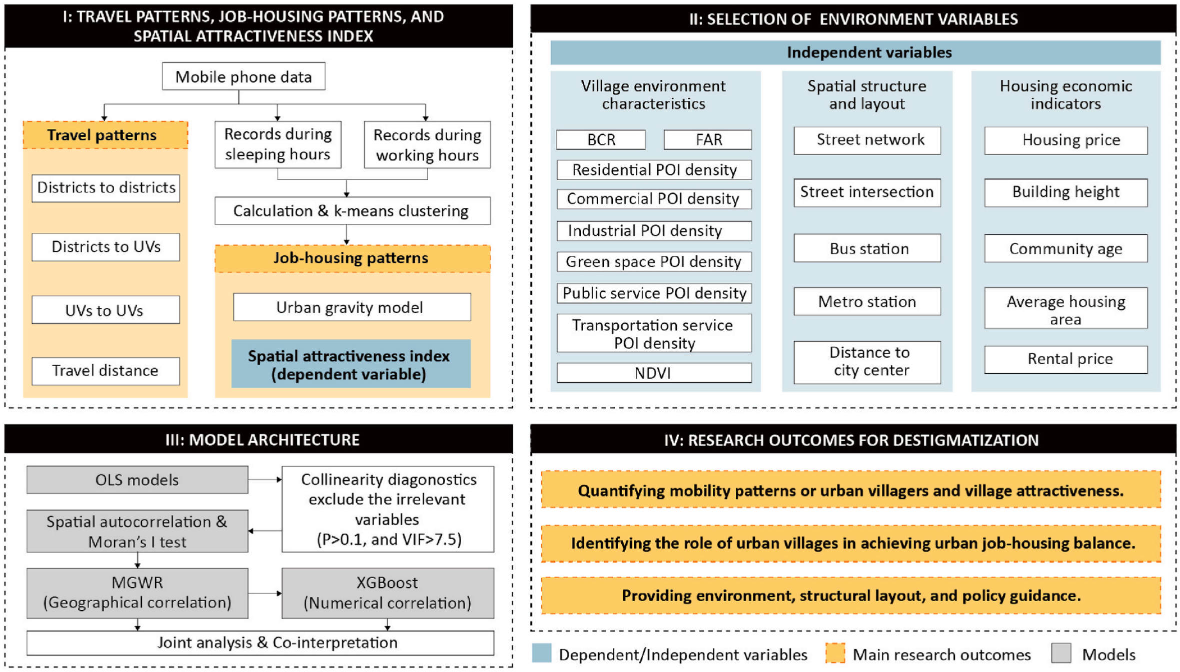

Fig. 3. Method and workflow.

POIs. Additionally, we adopted the average distance of OD points reaching location *i* via a street network fitting as the value for *d*.

$$T_{ij} = \frac{Q_i Q_j}{d^2} \tag{4}$$

$$ATI_i = \frac{P_i G_i P2_i}{d^2} \tag{5}$$

2.3.3. OLS, Moran's I, and MGWR

The OLS is the most commonly used method for analyzing the relationships between two or more variables. However, it neglects the spatial dependence of observations and their independent influences at the local level. Hence, we further employed the MGWR model, renowned for its ability to capture spatial heterogeneity, as a means to enhance our analysis. Prior to applying MGWR, it is imperative to conduct a spatial autocorrelation assessment to unveil underlying spatial patterns and dependencies. Additionally, calculating the Moran's I index allows us to validate the crucial assumption of spatial independence. MGWR differs from the conventional classical GWR model by making the bandwidth specific to each variable. This approach addresses the limitation of GWR where all variables must have the same optimal bandwidth. Furthermore, the specific bandwidth for each variable can be used as an indicator of the spatial scale of each spatial process. Thus, the spatial process model generated by this multi-bandwidth method is more realistic and practical (Shen et al., 2020).

2.3.4. RF and partial dependence plots (PDPs)

We compared six regression models (Table A 2). The results indicate that the RF model exhibits superior performance, particularly evident in its larger errors on the training set compared to smaller errors on the test set, suggesting an absence of overfitting. The training and test sets were split in an 8:2 ratio. In the WOUV model, the sample size was 1126, while in the LOUV model, the sample size was 352. We optimized the parameters of the RF regression model. Initially, a comprehensive parameter set was defined, including the number of trees, the maximum depth of the trees, the minimum number of samples required to split an internal node, the minimum number of samples required to be at a leaf node, and the number of features to consider when looking for the best split. Subsequently, the RandomizedSearchCV method was employed for random search, coupled with 5-fold cross-validation. Specific parameters are displayed in Appendix B. RF parameters.

Additionally, we chose PDPs to visualize complex relationships between variables. PDPs can provide insights into the impact of specific factors on the predicted outcome of a machine-learning model (Greenwell, 2017).

3. Results

3.1. The roles of UVs in job-housing balance and ATI results

We identified types of UVs to demonstrate their role in urban job-housing balance, and the typological classification serves as a theoretical foundation for subsequent calculations of ATI of UVs. Fig. 4 a showcased the clustering results, while Fig. 4 b indicated the geographical distribution of WO- and LOUVs. The greater number of WOUVs compared to LOUVs indicated that UVs, characterized as distinctive informal settlements, have evolved beyond just offering low-cost living. By employing Eq (5), we calculated the ATI (Fig. 4c). The distribution of high and low ATI for WO- and LOUVs demonstrated an interdependence, confirming the geographical interaction and functional complementarity (Fig. 4b). The presence of both high and low values across different ADs underscored the initial polycentric urban structure of Shenzhen. During the early stages of Shenzhen's development, UVs were primarily concentrated in urban core areas. However, as the city faced limitations in terms of land availability within the city center, the expansion of urban construction gradually extended towards the outskirts (Hao et al., 2012, 2013). Consequently, the main body of UVs shifted, as reflected in the flourishing development and stronger ATI of UVs in the suburbs.

3.2. Model results of variables influencing the ATI

3.2.1. OLS results of variables influencing the ATI

Following existing studies (Dong et al., 2023; Han et al., 2023; Li et al., 2022), we incorporated the independent variable clusters I, II, and III for both WOUVs and LOUVs into the OLS model. We excluded variables with VIF greater than 7.5, and the processed data did not have

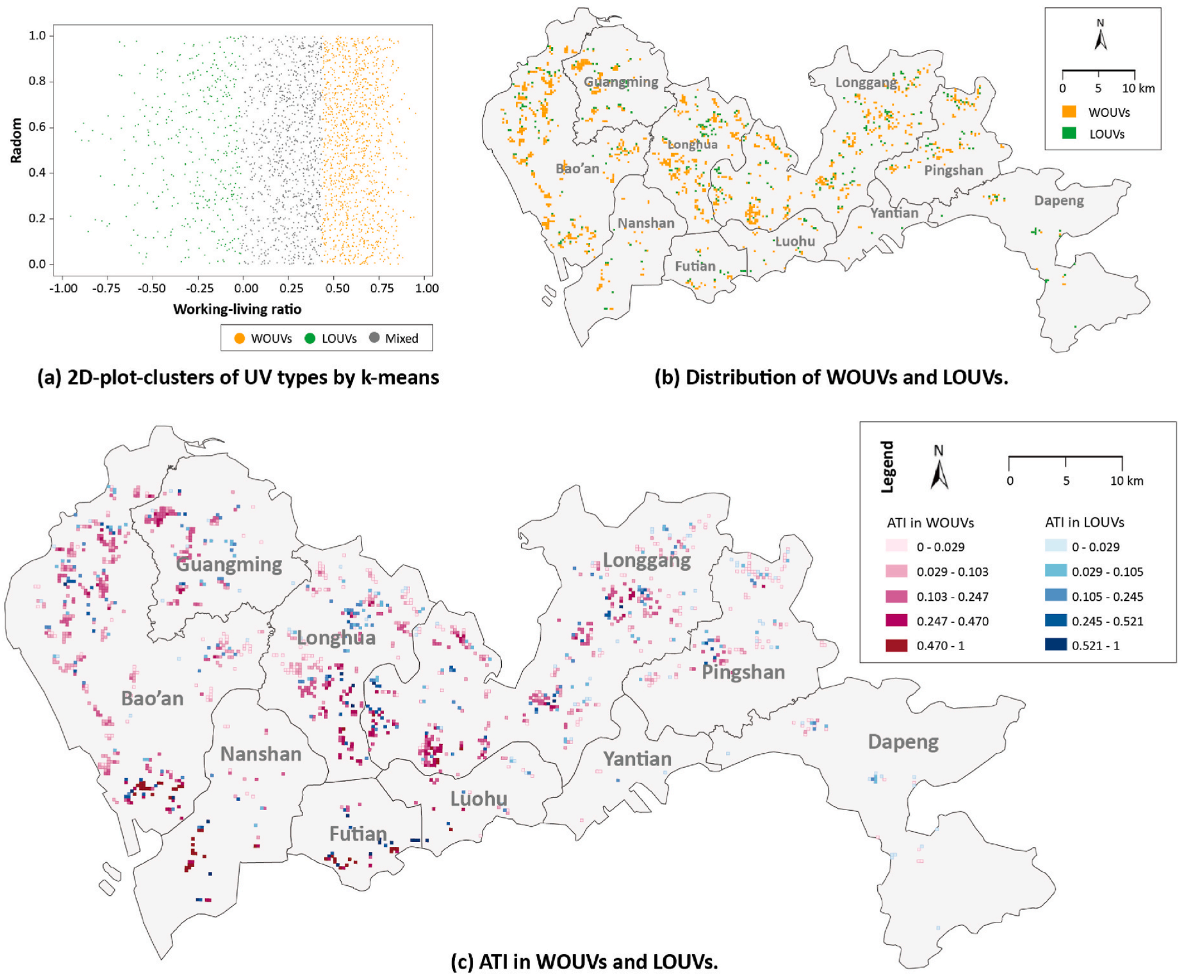

Fig. 4. Clustering results and the ATI of urban villages.

multicollinearity problems (Table A 4). The adjusted R^2 of Cluster I, II, and III were 0.350, 0.289, and 0.295 in WOUVs, and 0.380, 0.366, and 0.224 in LOUVs, respectively. Cluster I demonstrated better explanatory power for the ATI. Next, we identified variables for further examination in MGWR models.

3.2.2. *Spatial clustering of the ATI*

Fig. 5 b revealed that the Moran's I index of the ATI stood at 0.693, with a z-value of 57.307 ($p < 0.001$), highlighting a significant positive spatial dependence and spatial clustering effect. The LISA cluster map indicated the presence of hot and cold spots regarding the ATI of UVs. In Fig. 5 a, the spatial clustering map displayed the concentrated presence of ATI hot spots (High-High clusters) scattered throughout the core

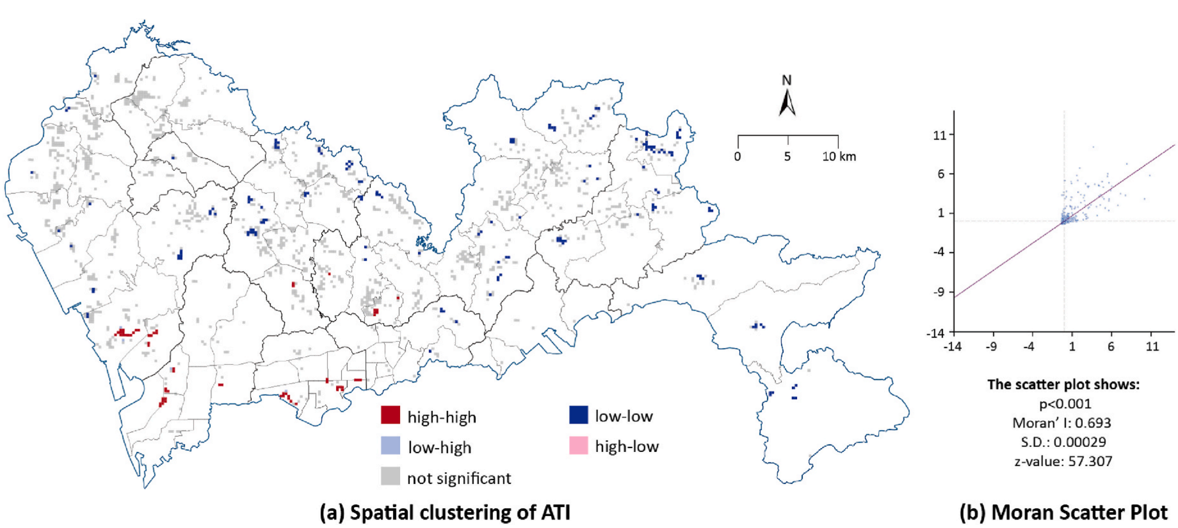

Fig. 5. The LISA cluster map displayed local Moran' I about the ATI.

urban areas, particularly in the southern part of the Nanshan as well as Futian and Luohu districts, suggesting a positive correlation with the proximity to the city center. Conversely, the ATI cold spots (Low-Low clusters) manifest in the suburbs.

3.2.3. MGWR and RF results of the ATI

We examined the global Moran's I for selected variables to assess the presence and strength of spatial autocorrelation. Table A 5 showed that the selected variables exhibited strong spatial autocorrelation. Table A 3 indicates their suitability for the MGWR model. In contrast to the OLS model that incorporates district dummies, MGWR captures continuous and unbounded spatial effects. Table 2 demonstrated that MGWR models performed better than the OLS models, as reflected by the higher R^2. Additionally, AIC and AICc in MGWR models were smaller than the OLS models, emphasizing their enhanced capacity in data explanation.

Fig. 6 presents the variable importance rankings for influencing the ATI. In WOUVs, the importance of variables exhibited a gradual decline. The highest-ranking variables were public service density, distance to the city center, and residential density. However, the results for LOUVs differed, with distance to the city center occupying the top spot at nearly 0.6, surpassing other variables by a considerable margin. This highlighted the critical role of location factors in residents' choice of residential UVs. The following variables included housing price and transportation service density, indicating the consideration of economic and time costs.

3.3. Joint analysis of different independent variable clusters and the ATI

3.3.1. The impact of cluster I on the ATI

Regarding FAR of WOUVs, findings from both one-way PDPs and MGWR indicated a low correlation between FAR and the ATI (Fig. 7 a.1.1 and a.1.3). However, two-way PDPs revealed that the proximity to the city center does not have a significant effect on the relationship between FAR and the ATI of WOUVs (Fig. 7 a.1.2). In conjunction with MGWR, areas displaying a weak correlation were situated in the Futian, Luohu, southern Longhua, and southwestern Longgang (Fig. 7 b.1.3). One-way PDP demonstrated a positive correlation between FAR and the ATI of LOUVs (Fig. 7 a.2.1). MGWR revealed a consistent positive correlation (Fig. 7 a.2.3). FAR contributed to the promotion of the ATI both numerically and geographically. By incorporating two-way PDPs, MGWR, and the spatial distribution of FAR, we observed a positive correlation between FAR and ATI. High FAR not only reduces the urban imageability but also significantly contributes to the stigmatization of UV spaces. However, maintaining FAR in LOUVs ensures efficient land use while providing low-priced living spaces (Hao et al., 2013). It is necessary to strike a balance between inefficient land use and over-densification. In the UV renovation in Guangzhou, comprehensive economic feasibility factors were considered, and the strategy of 'more commerce over residence' was attempted to balance FAR (Wu et al., 2018).

OLS and MGWR revealed a weak negative correlation between commercial density and the ATI in LOUVs (Fig. 7 b.2). One-way PDP indicated a nonlinear correlation between commercial density and the ATI. The relationship was relatively stable when the coefficient was under 0.25, and it showed a nonlinear correlation above 0.25 (Fig. 7 b.1). The distinctive environment of UVs has become a fertile ground for commercial activities and retailers. These intensively commercialized UVs cater not only to their immediate inhabitants but also radiate services to the surrounding neighborhoods (Rui & Li, 2023). The prevailing commercial typology within these UVs leans heavily towards restaurants. Owing to the high cost-effectiveness of these restaurants' operations leading to lower sanitary standards, they are sometimes informally referred to by residents as 'fly-ravaged eateries.' This spatial stigmatization stems not just from concerns over the cleanliness of village space but is also intricately tied to business operational paradigms. Consequently, the functional use of the space and its business models can be leveraged as a potential countermeasure against such stigmatization.

One-way PDP demonstrated a positive correlation between public service density and the ATI in WOUVs (Fig. 7 c.1.1). MGWR revealed a correlation coefficient ranging from −0.214 to 0.358, with positive correlations concentrated in the core areas and the southern part of Bao'an (Fig. 7 c.1.2). This implied that WOUVs in the urban center have developed diverse public service facilities to meet the needs of the migrant population (Hao et al., 2012). One-way PDP exhibited a positive correlation between public service density and the ATI in LOUVs (Fig. 7c.2.1). The coefficient for MGWR ranged from −0.051 to 0.660, demonstrating a relatively positive relationship (Fig. 7c.2.2). In contrast to WOUVs, the strong positive correlation in LOUVs was identified in Longgang, followed by the core areas. This might be due to the different types of public facilities. To delve deeper into the underlying causes of this phenomenon, we extracted information on life services within public services. Table 3 illustrates that the quantity of life services in Longgang was higher than in the core areas, although the per capita availability was lower (except for Luohu). This suggested that the proximity and quantity of surrounding living amenities play a pivotal role in determining the appeal of UVs as residential choices. Therefore, augmenting the number of living amenities in suburban areas can substantially enhance the spatial attractiveness of LOUVs.

3.3.2. The impact of cluster II on the ATI

The proximity to the city center was positively correlated with the ATI of WOUVs (Fig. 7 d.1.1). One-way PDP unveiled a threshold effect, indicating that the relationship with the ATI was negative when the coefficient of distance to the city center was below 0.2. Considering the dispersed distribution of UVs with the high ATI (Fig. 4 c) and a higher proportion of commuting within the district, as well as the "tolerance limit" for commuting time (Du et al., 2020; Tian et al., 2022), it becomes evident that residents who are far from the city center lack the inclination and necessity to seek employment in the core areas, but choose to be employed within proximity to their living places. This further corroborates the pivotal role of UVs in urban living-working balance.

In LOUVs, the correlation coefficient in MGWR was consistent with the distribution of values of distance to the city center, i.e., increasing from the core areas to the suburbs (Fig. 7 d.2 and Figure A 4). PDP

Table 2
Comparison of OLS and MGWR models.

WOUVs	OLS				MGWR			
	R^2	Adj.R^2	AIC	AICc	R^2	Adj.R^2	AIC	AICc
Cluster I	0.354	0.35	2718.498	2720.598	0.658	0.629	2165.126	2180.286
Cluster II	0.292	0.289	2819.030	2821.130	0.683	0.647	2133.363	2160.302
Cluster III	0.297	0.295	2806.213	2808.266	0.595	0.566	2331.715	2343.137
LOUVs	**OLS**				**MGWR**			
	R^2	Adj.R^2	AIC	AICc	R^2	Adj.R^2	AIC	AICc
Cluster I	0.394	0.38	839.799	842.219	0.681	0.649	622.141	669.057
Cluster II	0.377	0.366	842.647	844.820	0.510	0.491	775.433	776.616
Cluster III	0.235	0.224	917.163	919.406	0.452	0.409	839.922	844.322

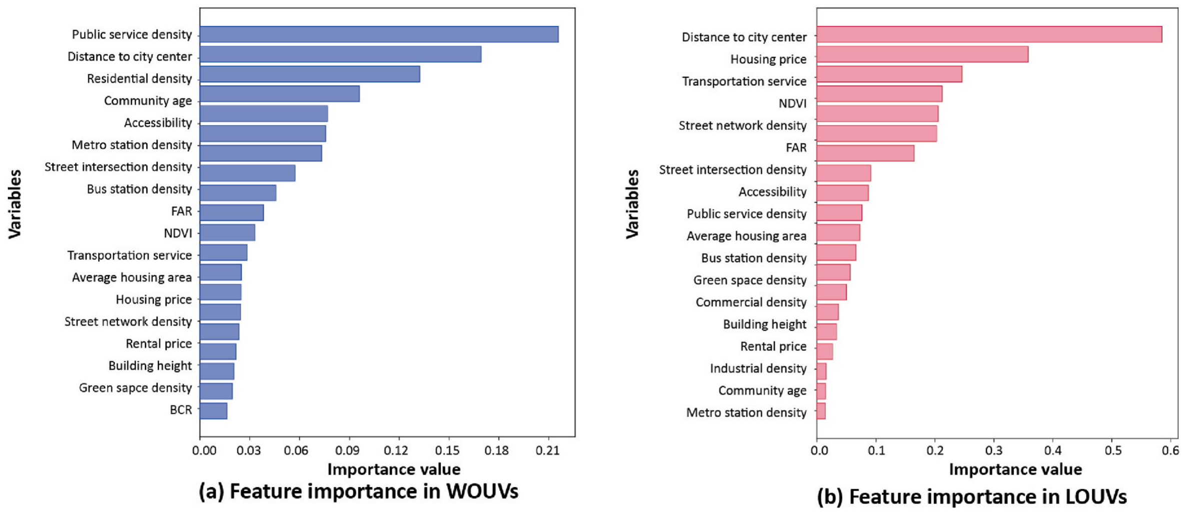

Fig. 6. Feature importance to determine the ATI in WOUVs and LOUVs.

regarding housing price and rental price offered interesting insights (Fig. 7 d.2.2, d.2.3): Residents exhibited a preference for acquiring residential properties in LOUVs with lower housing prices (coefficients ranging from 0.25 to 0.5) located at a medium distance from the city center (coefficients ranging from 0.25 to 0.40) (Fig. 7 d.2.3). Conversely, when choosing UVs as temporary places of residence, residents tend to favor those situated in the core areas. These findings underscored the unique attributes of UVs, offering affordable rental housing and advantageous geographical positioning.

3.3.3. The impact of cluster III on the ATI

The geographic distribution of housing price exhibited a decreasing trend from the city center to the suburbs in all types of UVs (Figure A 4). Both one-way PDP and MGWR demonstrated a negative correlation between housing price and the ATI of WOUVs (Fig. 8 a.1.1, a.1.2, b.1.1 and b.1.2). Areas with negative values were concentrated in the core areas and the southern part of Bao'an. This aligned with the apartment purchasing preferences of the low-income groups. The distance to the city center did not exhibit a positive synergistic contribution to the ATI with housing price, as the higher prices associated with central locations became obstacles (Fig. 8 a.2.2) (Wu et al., 2015).

Similarly, high coefficient values of rental prices are concentrated in core areas and the southern part of Nanshan (Figure A 4). Both one-way PDP and MGWR demonstrated a slight negative correlation (Fig. 8 b.1.1 and b.1.2). Intriguingly, the rental prices in WOUVs showed a geographically opposite distribution with the coefficient in MGWR, where high rents almost align with the low ATI. This observation resonated with the current paradigm where WOUVs served as production spaces. WOUVs that are closely integrated with e-commerce attract investments in suburban areas through lower rents, establishing small economic zones that stimulate regional employment and internal circulation. These findings challenged the previous stigmatized perception of poverty associated with UVs while rectifying and recognizing the positive contributions of UV workers to the urban economy.

LOUVs exhibited a more pronounced negative correlation between rental price and ATI (Fig. 8 b.2.1). Residents become more sensitive to rental costs when choosing UVs as living spaces. In the city centers, higher rents can enhance the spatial appeal of LOUVs (Fig. 8 b.2.2). Urban villagers in the city centers accord importance to proximity to workplaces and convenient amenities. While rents in these LOUVs are relatively high, they present a more favourable cost-benefit ratio when contrasted with formal settlements.

4. Discussion

4.1. The interaction between UV renewal and destigmatization

Promoting the integration of UVs with the city, the renewal of UVs and urban development is a symbiotic process of integrated development. Destigmatization is an essential step for UVs to blend into the city. There is a close interactive relationship between the strategies for UV renewal and the gradual achievement of destigmatization.

Scholars' perspectives on UVs have shifted from negative and passive to recognizing their socio-economic value. Simultaneously, the planning practice towards UVs has changed from a "demolish-modify-retain" approach to a "retain-modify-demolish" strategy (Pan & Du, 2021b). This shift is essentially the destigmatization of UVs at the theoretical level, which has influenced the changes in practical strategies. Theoretical studies at the social and economic levels struggle to reverse the negative impressions of other urban residents towards UVs. Reversing these negative impressions requires optimization and upgrading of physical spaces and service functions. Additionally, the transformation of UVs can improve residents' living standards and potentially stimulate various sectors, including commercial investment, demand, and consumption. For the overall economy, the renewal of UVs can be seen as a significant infrastructural investment. By transforming the built environment and commercial patterns of UVs, their attractiveness can be increased, driving regional consumption demand and revitalizing urban assets. Not only in China, but studies in other regions, such as Zimbabwe, also demonstrate that UV renewal can bring significant economic benefits (Matamanda et al., 2020).

In general, renewal can improve the appearance of UVs. However, it is important to note that inappropriate renewal of UVs may exacerbate their stigmatization. For instance, resettlement housing based on UV construction is often perceived as a "second-class community," with designs clashing with the surrounding environment. This reflects the influence of the pervasive stigmatization of UVs on decision-makers and urban designers, who fail to recognize the significant role of UVs, thereby preserving or even intensifying their negative image. Decision-makers need to consider the opposite situation to stigmatization, where aggressive UV renewal might lead to radical gentrification (Wu & Wang, 2017). This requires government intervention to prevent gentrification, with the "Milieuschutz" areas planned by the government in Berlin serving as a reference (Kadıoğlu, 2022).

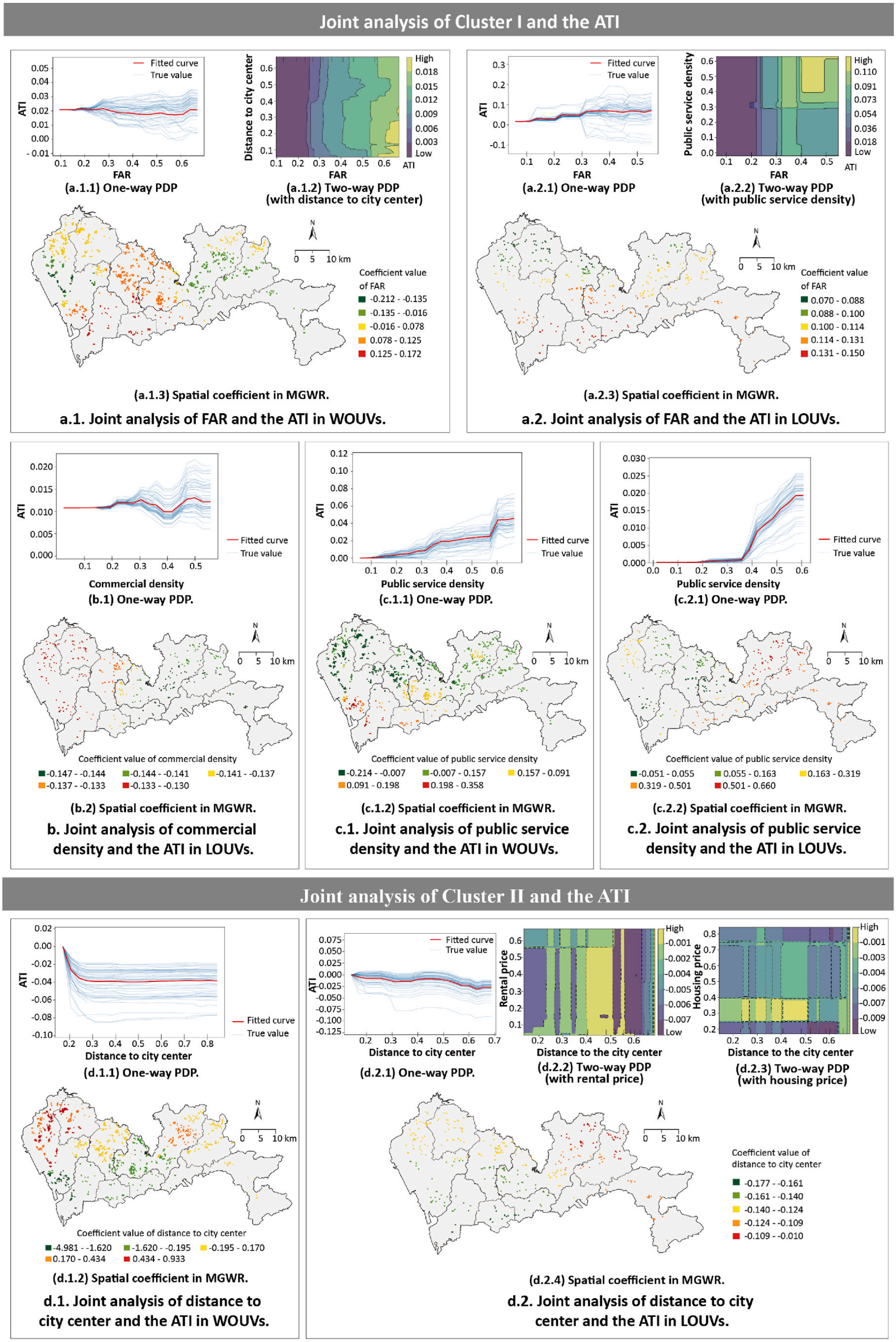

Fig. 7. Joint analysis of cluster I and II on the ATI.

Table 3
Life services in public service facilities (POIs).

ADs	Life services				Population	Ratio
	School	Sports & Leisure	Medical care	Total	Unit: Million	Number of POIs per million people
Futian (core)	378	1205	1706	3289	1.553	2117.836
Nanshan (core)	582	1747	1753	4082	1.795	2274.094
Luohu (core)	294	588	741	1623	1.143	1419.947
Longgang	837	2066	3691	6594	**3.979**	1657.200

Source of population data: Seventh National Census Bulletin (http://tjj.sz.gov.cn/ztzl/zt/szsdqcqgrkpc/szrp/content/post_8772114.html).

4.2. *Supporting the destigmatization of informal settlements by quantitative evidence*

Our study provides evidence for the spatial destigmatization of Chinese informal residential areas—UVs. This research quantitatively demonstrates the positive value created by UVs in promoting job-housing balance and providing diverse services, further offering decision-making suggestions for UV renewal. Quantifying the mobility patterns of residents and the spatial attractiveness of UVs plays a pivotal role in mitigating UV stigmatization. On one hand, by categorizing UVs based on mobility patterns, we gained a deeper understanding of the distinct spatial characteristics of different UVs and the unique roles they play within the city. For example, we discovered that the number of WOUVs is even greater than that of LOUVs, which serves as typical evidence of the multifunctional services provided by UVs within the city. We suggest that the planning department utilize geospatial data to further clarify the types of different UVs, update the public perception of the functions of UVs, and reduce stereotypes.

The transformation of UVs is a crucial component of urban renewal in China's major cities, representing a necessary stage in the shift from large-scale incremental construction to a dual focus on the improvement and transformation of existing stock and the structural adjustment of incremental developments (Wang et al., 2009). Given that spatial regeneration is one of the political levers for destigmatization, and considering that the planning recommendations from macro settlements and slum studies in other countries are not directly applicable to UVs (Rui, 2023a; von Seidlein et al., 2021), we have undertaken a quantitative assessment of the UV environment. Specifically, aspects such as the village environment, spatial structure and layout, and housing economic indicators have been quantified, offering insights into destigmatization from the perspectives of spatial design, structural layout, and

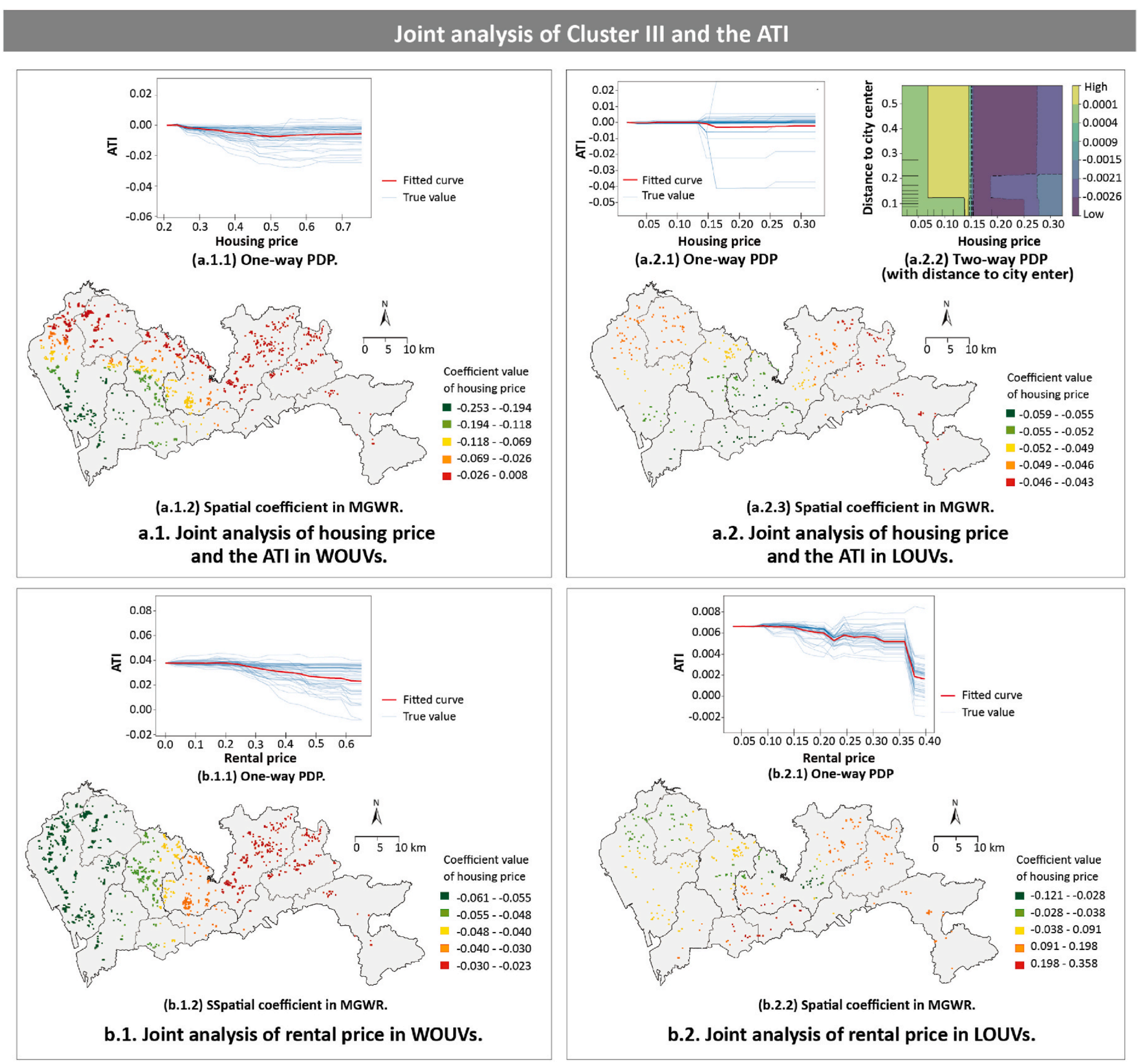

Fig. 8. Joint analysis of cluster III on the ATI.

governmental affordable housing policies. Additionally, by focusing on the mobility of marginalized urban villagers, we underscored the significance of attending to low-income groups and transient populations in fostering social inclusivity, economic vitality, and urban diversity (Hao et al., 2012; Pan & Du, 2021a; Rui, 2023a). For instance, the “mini-economic zones” formed by the integration of suburban WOUVs with e-commerce act as catalysts for development in urban peripheries.

The phenomenon of informal settlements has garnered extensive attention and discussion at the international level. Over the past two decades, projects aimed at improving the quality of life in these areas have been incorporated into numerous global initiatives. As Average (2020) points out, informality has become a significant aspect of urbanization. Scholars from various disciplines, including social sciences, economics, and public health, have extensively studied informal settlements, offering a range of perspectives. In this context, researchers worldwide have noted that urban informal settlements serve functions well beyond providing affordable housing (Brown-Luthango et al., 2017; Celhay & Gil, 2020). However, the stigmatization of informal settlements is a pervasive issue globally, particularly in developing countries (Brown-Luthango et al., 2017; Dovey et al., 2021; Gouverneur, 2014). In China, this stigmatization primarily stems from the dual urban-rural socio-economic management system and the poor quality of the physical environment. In other countries, destigmatizing informal settlements may involve addressing more complex issues such as apartheid planning (Lochner & Ntema, 2013), tenure rights (Patel, 2013), illegal demolition (Roberts & Okanya, 2022), and even political issues (Björkman, 2014). Studies on the destigmatization of informal settlements often focus on social and political aspects, with less attention given to the material and spatial dimensions. Indeed, as early as 2003, Abbott (2003) highlighted the need to integrate spatial and social data in the renewal of informal areas, linking the physical environment with socio-economic data related to the inhabitants for effective management and flexible decision-making structures. Tjia and Coetzee (2022) also emphasized the importance of accurate geospatial information in decision-making for the renewal of informal settlements. In this respect, we provide a replicable methodological framework based on geospatial and socio-economic data for the study of informal settlements in other countries and regions.

4.3. Renovation and policy recommendations for further destigmatizing UVs

4.3.1. Directions for UV regeneration orientation

Informal settlements, exemplified by “UVs,” have been assigned a variety of labels, with the vast majority being negative. Such spatial discourses have contributed to the stigmatization of UVs. As a result, enhancing the spatial quality of UVs is a critical approach to addressing their stigmatization. Through collaborative multi-model analysis, we suggest that future UV transformations can be approached from the following directions.

4.3.1.1. Refinement of urban structure and layout. Based on mobile phone data, this research elucidates the mobility trajectories of residents within UVs. The results indicate that UVs play a pivotal role in achieving a job-housing balance in Shenzhen. The majority of the residents prefer shorter daily commutes, primarily within a 20-km radius. This preference accentuates the contribution of UVs in proffering both residential and employment opportunities to the neighboring communities. UVs in the Futian and Nanshan districts exhibit the highest ATI. UVs in specific areas like Guangming, Bao’an, and Longgang districts also demonstrate strong appeal, suggesting that Shenzhen is evolving into a polycentric urban structure. Additionally, location and transportation convenience are critical considerations for the migrant population when choosing UVs as living places. Thus, enhancing transportation accessibility can effectively boost the attractiveness of UVs.

4.3.1.2. Optimization of built environment and public service facilities in UVs. Through correlation analysis, we found that the FAR of both WOUVs and LOUVs shows a positive correlation with the ATI. This indicates that UVs with higher FARs, characterized by their relatively lower rents, moderate living costs, and comprehensive functionalities, attract a significant number of rural migrants. However, it’s notable that when the FAR of LOUVs exceeds a certain threshold, the ATI correspondingly decreases (Fig. 7 a.1.1). This phenomenon might be attributed to some UVs having excessively high FARs (between 4 and 6), where widespread illegal constructions not only deteriorate the urban landscape but also contribute significantly to the spatial stigmatization of UVs. Therefore, addressing the issues of areas with high FAR and setting a reasonable upper limit for FAR, considering economic feasibility, is crucial. Moreover, there is an urgent need to strengthen the infrastructure development in UVs. During urbanization, the supply of public service facilities within government jurisdictions is insufficient (Table 3), coupled with the restrictions of the household registration system, making it difficult for the migrant population to access these services. To cater to the diverse needs of residents in UVs, the promotion of a variety of public facilities and diversified land use patterns is recommended (Hao et al., 2012). Addressing shortcomings in education, healthcare, elderly care, and domestic services in UVs can not only resolve the issues of unbalanced and insufficient development in large cities but also integrate more people into urban life, ensuring equal access to public services.

Furthermore, it is important to note that some UVs preserve cultural and social value in the form of historical relics and buildings. During the transformation process, these heritages should be thoroughly excavated and displayed. This not only enhances the sense of place and cultural identity within the region but also helps to reduce societal stigmatization towards UVs. This strategy has been validated in studies conducted in countries like South Africa and India (Auerbach, 2018; Brown-Luthango et al., 2017; Lochner & Ntema, 2013; Patel, 2013).

4.3.1.3. Revitalization of commercial vitality in UVs. There exists a synergistic relationship between commercial development and the destigmatization of UVs. On one hand, in the increasingly competitive business environment, if UVs fail to reverse their stigmatized image, it could lead to a decline in commercial activities and an increase in vacant storefronts. On the other hand, our research in LOUVs reveals that an increase in commercial facilities significantly enhances the attractiveness of these areas (Fig. 7 b.1). We hypothesize that commercial facilities within UVs can generate a radiating effect. A distinctive feature of UV commerce is its adaptability to market demand changes, necessitating a more considered approach in the introduction and cultivation of industries, given its demand-driven elasticity. Community commerce, as a social form that most vividly reflects the progress of urban civilization and the convenience of living services, is closely linked to societal development and can even facilitate social integration among different groups. For instance, the recent rise of community fresh markets and convenience stores not only attracts local consumers but also improves the commercial ecosystem of the community, thereby promoting social integration within UVs.

From a management and regulatory perspective, to achieve a sustainable spatial renewal model, it is imperative to encourage and support private capital investment in UVs. This point is also emphasized in studies of other informal residential areas (Hegazy, 2016). Joint planning by government agencies and village residents not only ensures the feasibility of projects but also maintains consistency from preliminary planning to post-construction operation, thus ensuring the effective implementation of unique commercial initiatives. Moreover, participatory communication can indirectly change participants’ negative perceptions of UVs.

4.3.1.4. Recommendations for the adjustment of housing policy. Firstly, rural migrants face multiple challenges, such as economic difficulties and social stigmatization. Given these circumstances, affordable housing is crucial for their initial adaptation to city life. For LOUVs, the strategic utilization of baseline plot ratios coupled with land price levers is imperative to counteract issues such as densification and escalating rents, thereby facilitating the transformation of UVs. Secondly, there's a need to promote the equitable distribution of facilities, emphasizing essential living services in core areas and strengthening public facility construction in remote regions. Notably, LOUVs located in city centers can serve as pilot areas for economically affordable housing policies, offering cost-effective social housing. Thirdly, the government should cultivate enduring, constructive synergies between urban industrial capital and WOUVs. Multi-stakeholder communication models can help shift negative perceptions of UVs among participants. Within the framework of urban industrial configurations and spatial development paradigms, it is essential to consider the integration of UVs into the city's production landscape, especially the unique economic and employment contributions of peripheral WOUVs.

4.3.2. Differentiated strategies for optimizing WOUVs and LOUVs

Informal settlements play various roles in urban environments. For instance, Zhang (2018) has identified three primary functions of informal settlements in India: providing affordable housing, serving as marketplaces, and acting as "vote banks." Our research, grounded in the context of China, primarily categorizes the UVs of Shenzhen based on their residential and employment functions. The clustering method we propose, which utilizes k-means and common geospatial data, can also be applied to informal settlements in other regions. It is essential to conduct research on different types of informal settlements and develop corresponding strategies, aimed at unearthing the potential and constraints unique to each area based on their current conditions (Brown-Luthango et al., 2017; Matamanda et al., 2020; Zhang, 2023).

Different strategies are expected for UVs. Figure A 2 illustrates the high frequency of short-distance trips in the region. While sensitivity to travel distance may vary among individuals, it is apparent that most residents exhibit a preference for travelling within proximity to one another. For WOUVs, the incorporation of diverse service facilities and the implementation of mixed land use cater to the work and living requirements of the migrant population. Additionally, promoting proximity-based employment opportunities and reducing reliance on the city center can effectively catalyze the optimization of intra-regional circulation.

For LOUVs, increasing and decreasing FAR can potentially exacerbate challenges such as densification and soaring rents. It is necessary for the government to use the benchmark plot ratio and land price leverage to implement dual regulation of UV transformation. Moreover, balanced facilities include desirable accessibility and homogenization of resources. To achieve this goal requires the redistribution of facilities. For instance, it is suggested to add life services around UVs in the core areas. Conversely, prioritizing the bolstering of public service facilities in remote UVs could improve social equity and well-being (Rui & Othengrafen, 2023). Moreover, residents prefer to rent UVs that are close to the city center. Considering its higher rental price, the government can contemplate integrating subsidized housing policies with UVs in the core areas. Currently, the ongoing V-town project (http://www.v-town.com), spearheaded by state-owned enterprises, aligns with our recommendations. Taking the Yuanfen village in Longhua District as a pilot, this groundbreaking initiative represents the first attempt to establish a "high-quality and affordable" rental platform.

5. Conclusion

In general, destigmatization represents a significant social and political issue, involving the elimination of negative stereotypes, discrimination, and prejudice against specific groups or identities (Demirtaş-Madran, 2020; Zhang et al., 2021). Innovatively, we attempted to approach the stigmatization of UVs in Shenzhen from a spatial perspective. The value of spatially related research lies in its ability to provide quantifiable evidence and apply it within specific informal settlement units. We depicted the interactive relationship between behaviours and the village environment to comprehend the complex spatial and social dimensions of informal settlements. Grounded in the "Spatial-Behavioural Interaction Theory," we explored the behavioural patterns of UV residents and the residential and employment attributes of UV spaces. Moreover, by quantifying residents' behavioural destinations as the attractiveness of UVs, we tried to figure out the impact of village environment characteristics, spatial structure and layout, and housing economic indicators on residents' travel behaviours.

Our study has the following limitations. While previous research has relied on OD distances to calculate urban gravity models, we took a different approach by integrating OD points and street networks to identify the shortest paths. Although this method brought us closer to capturing the real travel dynamics, it failed to represent the diversity of path choices. Moreover, our regression models did not combine all types of UVs, which may ignore the geographical interrelationship between WOUVs and LOUVs. Additionally, UVs in China are highly complex socioeconomic composites. Future research could integrate macroeconomic and microeconomic approaches to further investigate how the presence of UVs reduces the operational costs of entire cities, thereby quantifying the positive externalities offered by UVs, providing additional data to support for their destigmatization.

CRediT authorship contribution statement

Jin Rui: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Yuhan Xu:** Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Xiang Li:** Visualization, Investigation.

Declaration of competing interest

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

Appendix A. Tables

Table A 1
POI classification is based on the sectoral classification on the Baidu platform.

Integration Clusters	Counts	POI classifications	Secondary classifications
Residence	30,499	Accommodation services	Communities, apartments, villas, dormitories etc.
Commerce	133,954	Business, commercial, entertainment, and other facilities	Shopping centers, restaurants, hotels, cinemas, banks etc.
Industry	90,755	Factories and industrial buildings	Companies, factories, science and technology parks, industrial parks, etc.
Public services	66,218	Land for cultural, educational, medical treatment and health, and social welfare activities.	Education and culture services, health care services, sport and leisure services, government agencies, public facilities, etc.
Green space	2637	Public places such as parks, green spaces, and squares.	Scenic spots, zoos, botanical gardens, parks, squares, etc.
Transportation services	47,081	All the roads, facilities, and junctions.	Parking lots, transport facility services, car and motorcycle services or repairs, etc. (Bus stations, metro stations and street intersections are excluded.)

Table A 2
Comparison of different machine learning model performance.

WOUVs	MSE		RMSE		MAE		MAPE	
	Train dataset	Test dataset	Train dataset	Test dataset	Train dataset	Test dataset	Train dataset	Test dataset
Decision tree	0	0	0.016	0.017	0.01	0.009	81.264	129.662
Radom forest	0.001	0	0.027	0.013	0.014	0.008	65.515	100.081
K-nearest neighbors	0.004	0	0.065	0.015	0.026	0.008	68.447	119.834
Support vector machines	0.007	0.003	0.086	0.054	0.053	0.043	125.581	153.58
LightGBM	0	0	0.014	0.016	0.006	0.009	104.027	182.322
XGBoost	0	0	0.001	0.013	0.001	0.008	41.024	139.016
LOUVs	Train dataset	Test dataset	Train dataset	Test dataset	Train dataset	Test dataset	Train dataset	Test dataset
Decision tree	0	0.007	0.005	0.082	0.003	0.032	69.487	147.957
Radom forest	0.002	0.001	0.04	0.028	0.016	0.016	56.745	88.921
K-nearest neighbors	0.008	0.001	0.089	0.025	0.031	0.012	86.224	166.31
Support vector machine	0.01	0.003	0.098	0.052	0.057	0.042	156.272	104.232
LightGBM	0.001	0.001	0.025	0.034	0.09	0.019	72.679	387.618
XGBoost	0	0.004	0.001	0.061	0	0.023	31.859	167.458

Table A 3
Descriptive statistics of all variables.

Clusters	Variables (Unit)	Explanation	Min	Max	Mean	S.D.
Dependent variable	ATI in WOUVs (N)	The attractiveness value of different UVs.	0.55	18410799.6	622910.41	1791983.23
	ATI in LOUVs (N)		0.066	8788167.42	294768.93	982555.59
Independent variable I: Village environment characteristics	Building Coverage Ratio (m^2/m^2)	The ratio of ground coverage by buildings within a study unit.	0.101	0.632	0.316	0.092
	Floor Area Ratio (m^2/m^2)	The ratio of buildings' total floor area within a study unit.	0.236	14.667	1.772	1.351
	Residential density (N)	POI density is calculated by kernel density.	2	636.974	186.375	13667.087
	Commercial density (N)		1	821.05	222	23538.799
	1.5 Industrial density (N)		0	447.644	92	2436.239
	Green space density (N)		0	22.9	2	3.699
	Public service density (N)		0	271.712	73	2293.688
	Transportation service density (N)		0	183.523	48	961.443
	NDVI (N)	Normalized Difference Vegetation Index.	0.037	0.309	0.116	0.002
Independent variable II: Spatial structure and layout	Street network density (m/m^2)	The total length of the street network in each study unit.	0	0.085	0.011	0
	Street intersection density (N)	Density is calculated by kernel density.	2.307	456.242	58.36	1885.375
	Bus station density (N)		2.198	142.44	44.146	655.191
	Metro station density (N)		0	6.242	0.826	0.773
	Distance to the city center (m)	The distance of each study unit from the Shenzhen Municipal Government.	1238.46	49119.507	23443.124	95597842.295
	Accessibility (N)	Vehicular accessibility of each study unit.	0	57.633	23.611	67.242
Independent variable III: Housing economic indicators	Housing price (yuan/m^2)	The average housing prices within each study unit.	10076.724	74900.917	23970.359	8858.472
	Building height (m^2)	The average height of the buildings within each study unit.	4.017	56.125	14.106	5.573
	Community age (years)	The Average Age of communities within each study unit.	0.936	31.397	15.364	3.088

(*continued on next page*)

Table A 3 (*continued*)

Clusters	Variables (Unit)	Explanation	Min	Max	Mean	S.D.
	Average housing area (m^2)	The average housing areas within each study unit.	19.582	205.308	73.392	25.316
	Rental price (yuan/m^2)	Average monthly rental price per square meter within each research unit.	28.035	285.478	62.946	29.852

Table A 4
OLS results by variable clusters in WOUVs and LOUVs.

WOUVs								
Cluster I: Village environment characteristics								
Variables	Coefficient	St. Error	t	Probability	VIF	R^2	Adj.R^2	F
Intercept	−0.09	0.012	−7.408	0.000***	–	0.354	0.35	F = 87.359
BCR	0.024	0.017	1.417	0.157	1.68			P = 0.000***
FAR	0.169	0.037	4.617	0.000***	1.502			
Residential density	−0.05	0.024	−2.045	0.041***	3.581			
Green space density	0.137	0.03	4.617	0.000***	1.21			
Public service density	0.315	0.032	9.8	0.000***	5.667			
Transportation service	−0.01	0.027	−0.369	0.712	3.603			
NDVI	0.043	0.02	2.203	0.028**	1.628			
Cluster II: Spatial structure and layout								
Intercept	−0.053	0.012	−4.472	0.000***	–	0.292	0.289	F = 77.126
Street network density	−0.035	0.03	−1.133	0.257	1.312			P = 0.000***
Street intersection density	0.378	0.036	10.379	0.000***	1.793			
Bus station density	0.116	0.018	6.302	0.000***	1.809			
Metro station density	0.046	0.021	2.155	0.031**	1.307			
Distance to the city center	−0.05	0.018	−2.807	0.005***	2.107			
Accessibility	0.068	0.023	2.948	0.003***	1.634			
Cluster III: Neighborhood attributes								
Intercept	−0.063	0.011	−5.588	0.000***	–	0.297	0.295	F = 95.277
Housing price	0.117	0.023	5.104	0.000***	1.562			P = 0.000***
Building height	0.068	0.026	2.68	0.007***	1.164			
Community age	0.015	0.018	0.882	0.378	1.045			
Average housing area	0.02	0.018	1.076	0.282	1.092			
Rental price	0.35	0.027	12.969	0.000***	1.598			
LOUVs								
Cluster I: Village environment characteristics								
Intercept	−0.054	0.012	−4.712	0.000***	–	0.394	0.38	F = 27.29
BCR	0.034	0.018	1.927	0.055*	1.566			P = 0.000***
FAR	0.21	0.033	6.384	0.000***	1.633			
Commercial density	−0.087	0.028	−3.117	0.002***	4.902			
Industrial density	0.027	0.024	1.108	0.269	2.092			
Green space density	−0.041	0.034	−1.211	0.227	1.261			
Public service density	0.128	0.032	3.997	0.000***	6.936			
Transportation service	0.049	0.026	1.913	0.057*	4.427			
NDVI	0.033	0.017	1.93	0.054*	1.573			
Cluster II: Spatial structure and layout								
Intercept	−0.01	0.01	−0.977	0.329	–	0.377	0.366	F = 34.768
Street network density	−0.008	0.026	−0.319	0.750	1.434			P = 0.000***
Street intersection density	0.177	0.033	5.423	0.000***	2.614			
Bus station density	−0.008	0.015	−0.545	0.586	1.596			
Metro station density	0.085	0.018	4.71	0.000***	1.744			
Distance to the city center	−0.035	0.016	−2.147	0.032**	2.55			
Accessibility	0.027	0.019	1.392	0.165	1.876			
Cluster III: Neighborhood attributes								
Intercept	−0.061	0.011	−5.385	0.000***	–	0.235	0.224	F = 21.308
Housing price	0.054	0.021	2.548	0.011**	1.523			P = 0.000***
Building height	0.097	0.024	4.045	0.000***	1.267			
Community age	0.053	0.017	3.035	0.003***	1.072			
Average housing area	0.034	0.021	1.638	0.102	1.145			
Rental price	0.09	0.026	3.492	0.001***	1.56			

Table A 5
Global Moran’s I and z-score and p-value.

WOUVs			
Variables	Moran’s I	Z-score	*P* value
BCR	0.151	18.119	<0.001
FAR	0.187	22.712	<0.001
Residential density	0.512	61.187	<0.001

(*continued on next page*)

Table A 5 (*continued*)

WOUVs			
Variables	Moran's I	Z-score	*P* value
Green space density	0.484	58.393	<0.001
Public service density	0.623	74.375	<0.001
Transportation service density	0.631	75.358	<0.001
NDVI	0.301	36.105	<0.001
Street network density	0.112	13.473	<0.001
Street intersection density	0.397	47.723	<0.001
Bus station density	0.641	76.556	<0.001
Metro station density	0.418	50.022	<0.001
Distance to the city center	0.900	107.347	<0.001
Accessibility	0.577	68.932	<0.001
Housing price	0.330	39.497	<0.001
Building height	0.336	40.260	<0.001
Community age	0.332	39.651	<0.001
Average housing area	0.577	68.899	<0.001
Rental price	0.506	60.598	<0.001
LOUVs			
Variables	Moran' I	Z-score	*P* value
BCR	0.343	16.883	<0.001
FAR	0.183	8.573	<0.001
Commercial density	0.409	18.788	<0.001
Industrial density	0.520	23.876	<0.001
Green space density	0.298	14.076	<0.001
Public service density	0.497	22.750	<0.001
Transportation service density	0.531	24.263	<0.001
NDVI	0.240	11.046	<0.001
Street network density	0.138	16.572	<0.001
Street intersection density	0.485	22.596	<0.001
Bus station density	0.586	26.753	<0.001
Metro station density	0.454	20.926	<0.001
Distance to the city center	0.822	37.442	<0.001
Accessibility	0.674	30.764	<0.001
Housing price	0.296	13.625	<0.001
Building height	0.292	13.516	<0.001
Community age	0.331	15.141	<0.001
Average housing area	0.369	16.911	<0.001
Rental price	0.499	23.117	<0.001

Appendix B. RF parameters

Random Forest was introduced by Cutler et al. (2012). It is an ensemble learning method used for regression and other tasks, which operates by constructing a multitude of decision trees at training time and outputting the mean prediction of the individual trees. For LOUVs, the results of the search revealed that the optimal parameter combination was as follows: The number of trees in the forest is set to 200. The maximum depth of the trees is limited to 40 levels. Minimum number of samples required to split an internal node is 6. The minimum number of samples required to be at a leaf node is 5. The number of features to consider when looking for the best split is determined by the logarithm to base 2 of the feature count.

For WOUVs, the results of the search revealed that the optimal parameter combination was as follows: The number of trees in the forest is set to 200. Maximum depth of the trees is limited to 30 levels. Minimum number of samples required to split an internal node is 3. Minimum number of samples required to be at a leaf node is 3. The number of features to consider when looking for the best split is determined by the logarithm to base 2 of the feature count.

Appendix C. UV-related travel through mobile phone data

Figure A 1. a and b depicted residential mobility from ADs to both ADs and UVs, as well as mobility between UVs. Figure A 1. c illustrates the travel ratio within and outside ADs. We found that Longgang, Bao'an, and Longhua districts exhibit the highest number of travels in both ADs and UVs, while Dapeng and Yantian exhibit the lowest frequency of travel. of except for out-of-district travel in Futian (51.54%), all other districts showed a higher frequency of travel within the district. This high frequency of travel within the district indicated the preliminary formation of a polycentric urban structure in Shenzhen.

In Figure A 1, the top three districts for trips associated with UVs as destinations are Longgang, Bao'an, and Longhua District. These are also ADs with the highest proportions of UVs. The percentages of travel originating from UVs within districts are 30.00%, 30.85%, and 28.26% respectively. Nearly one-third of the population is living and working in these UVs. This underscores the role of UVs in offering indispensable living spaces to a large number of rural migrants.

Figure A 2 indicated a decrease in travel frequency from short to long distances across all periods, except for Yantian. This finding suggests that residents prioritize shorter commuting distances for their daily activities, leading to a decline in the frequency of short-distance trips as travel distance increases. Notably, it was observed that the majority of residents prefer to travel within a 20 km radius. This implies that residents have tolerance limits for daily commuting. The top three inter-district travel with UVs as destinations are: from Bao'an to Nanshan, from Longhua to Longgang, and from Longgang to Longhua District. All of these are adjacent ADs. The residents' tolerance for commuting distances further emphasizes the socio-economic contribution of UVs in offering proximate housing and employment opportunities for the surrounding residents.

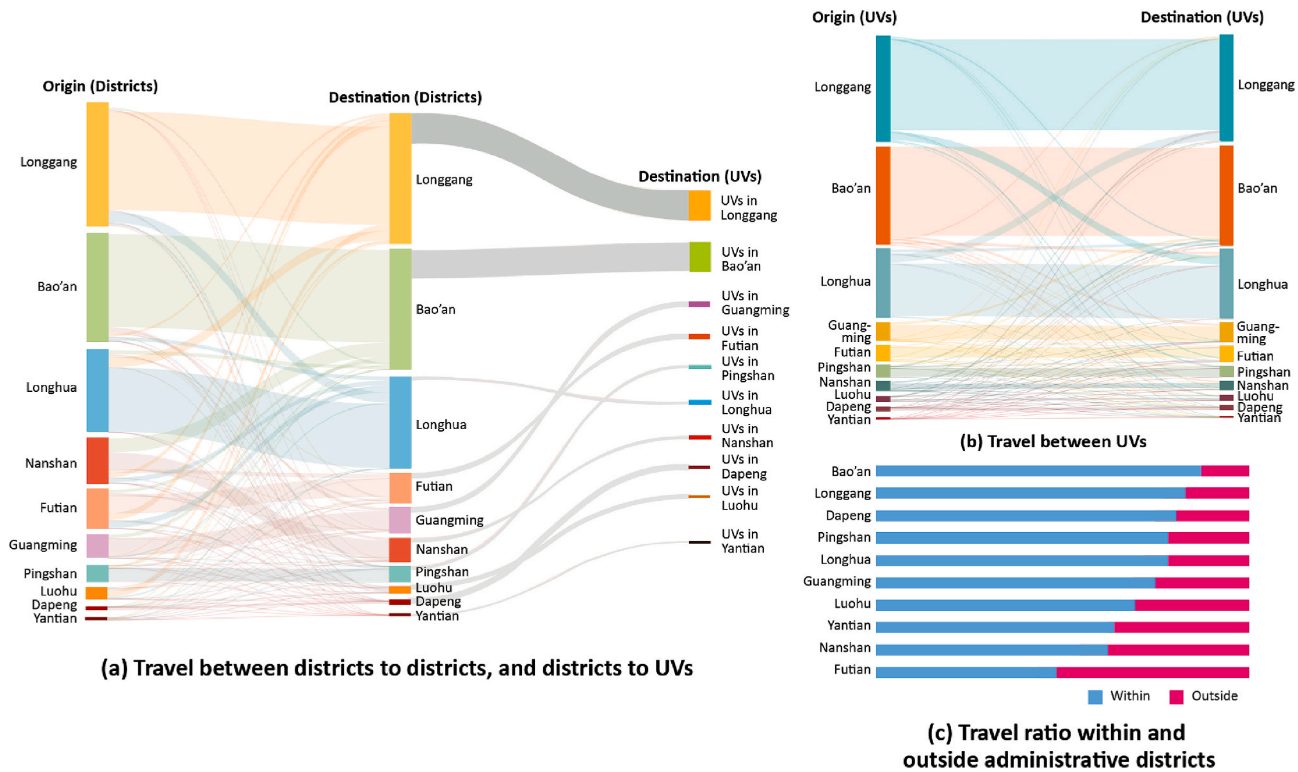

Fig. A 1. Travel patterns between administrative districts and urban villages.

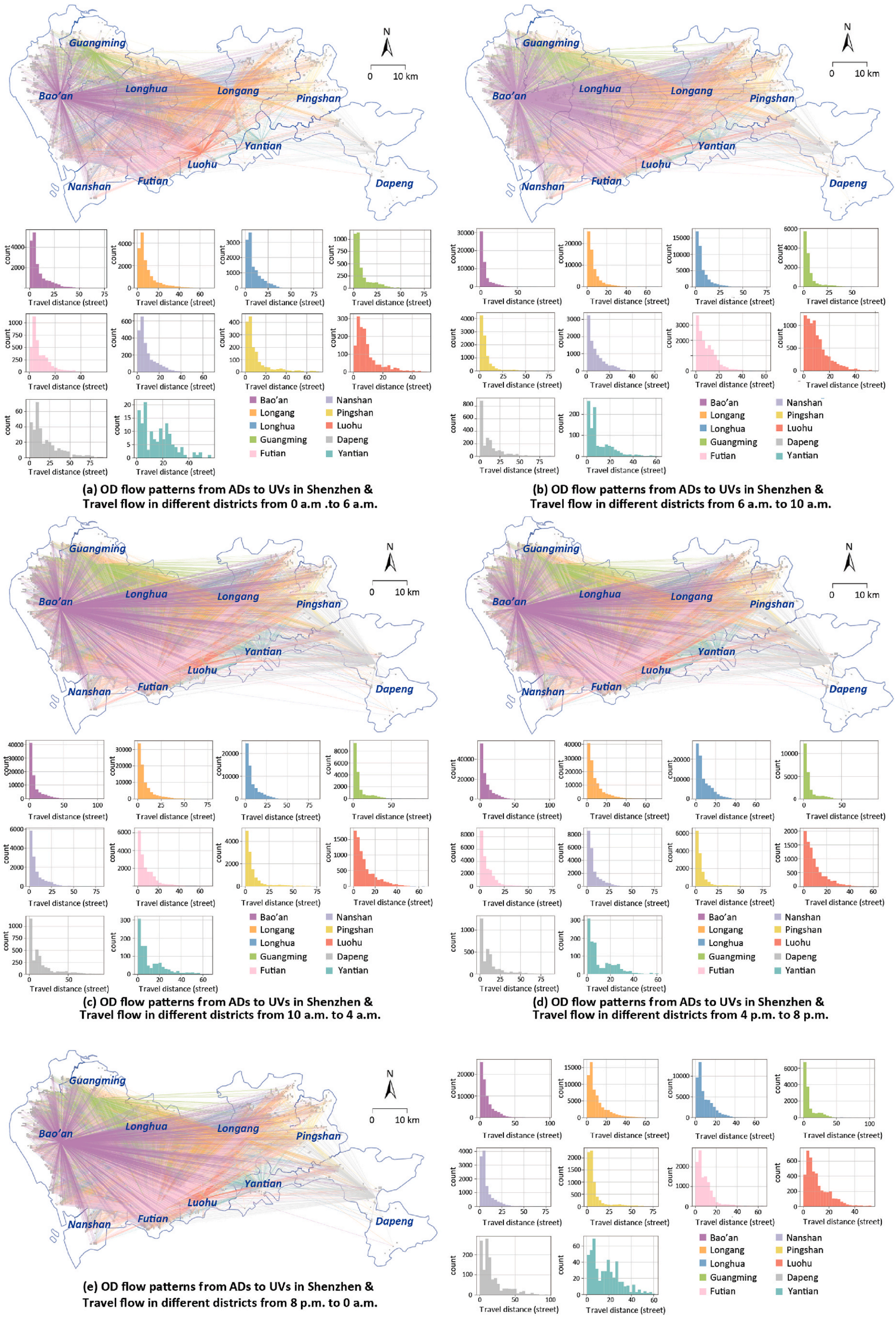

Fig. A 2. Statistics on daily visits to the urban village from the administrative district at different time intervals.

Appendix D. Other figures

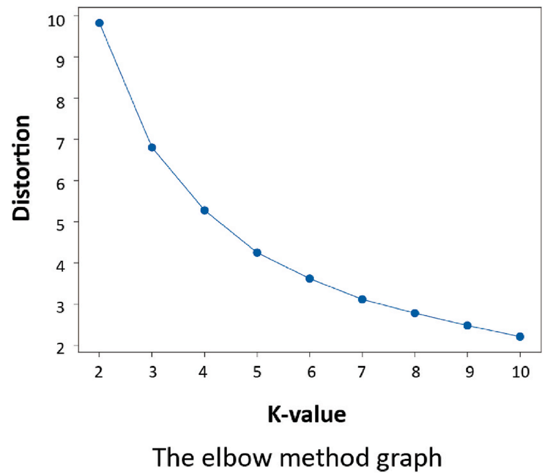

Fig. A 3. The elbow method for finding the optimal number of clusters in k-means.

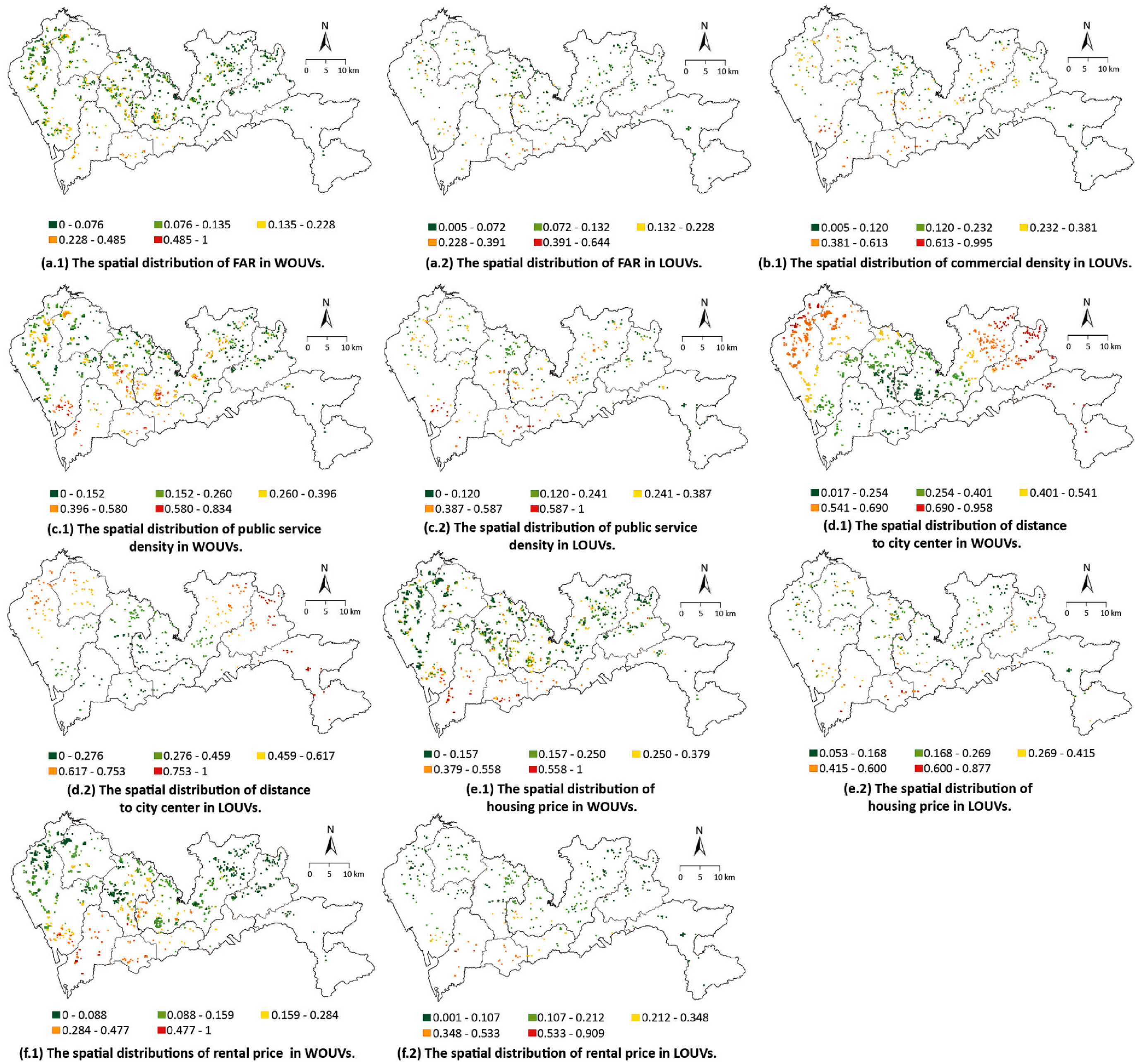

Fig. A 4. The spatial distribution of selected variables in WOUVs and LOUVs.

References

Abbott, J. (2003). The use of GIS in informal settlement upgrading: Its role and impact on the community and on local government. *Habitat International, 27*(4), 575–593. https://doi.org/10.1016/S0197-3975(03)00006-7

Auerbach, A. M. (2018). Informal archives: Historical narratives and the preservation of paper in India's urban slums. *Studies in Comparative International Development, 53*, 343–364. https://doi.org/10.1007/s12116-018-9270-5

Average, C. (2020). Contestations for urban space: Informality and institutions of disenfranchisement in Zimbabwe—the case of Masvingo city. *Geojournal, 85*(5), 1277–1289. https://doi.org/10.1007/s10708-019-10022-4

Björkman, L. (2014). Becoming a slum: From municipal colony to illegal settlement in liberalization-era Mumbai. *International Journal of Urban and Regional Research, 38* (1), 36–59. https://doi.org/10.1111/1468-2427.12041

Brown-Luthango, M., Reyes, E., & Gubevu, M. (2017). Informal settlement upgrading and safety: Experiences from cape town, South Africa. *Journal of Housing and the Built Environment, 32*, 471–493. https://doi.org/10.1007/s10901-016-9523-4

Celhay, P. A., & Gil, D. (2020). The function and credibility of urban slums: Evidence on informal settlements and affordable housing in Chile. *Cities, 99*, Article 102605. https://doi.org/10.1016/j.cities.2020.102605

Chai, Y., Tan, Y., Shen, Y., & Kwan, M. (2017). Space-behavior interaction theory: Basic thinking of general construction. *Geographical Research, 36*(10), 1959–1970. https://doi.org/10.1007/978-981-19-3474-2_2

Cutler, A., Cutler, D. R., & Stevens, J. R. J. (2012). *Random forests*. applications (pp. 157–175).

Demirtaş-Madran, H. A. (2020). Exploring the motivation behind discrimination and stigmatization related to COVID-19: A social psychological discussion based on the main theoretical explanations. *Frontiers in Psychology, 11*, Article 569528. https://doi.org/10.3389/fpsyg.2020.569528

Dong, L., Jiang, H., Li, W., Qiu, B., Wang, H., & Qiu, W. (2023). Assessing impacts of objective features and subjective perceptions of street environment on running amount: A case study of Boston. *Landscape and Urban Planning, 235*, Article 104756. https://doi.org/10.1016/j.landurbplan.2023.104756

Dovey, K., Shafique, T., van Oostrum, M., & Chatterjee, I. (2021). Informal settlement is not a euphemism for'slum': what's at stake beyond the language? *International Development Planning Review, 43*(2). https://doi.org/10.3828/idpr.2020.14

Du, Z., Tang, J., Qi, Y., Wang, Y., Han, C., & Yang, Y. (2020). Identifying critical nodes in metro network considering topological potential: A case study in Shenzhen city—China. *Physica A: Statistical Mechanics and its Applications, 539*, Article 122926. https://doi.org/10.1016/j.physa.2019.122926

Ewing, R., & Cervero, R. (2001). Travel and the built environment: A synthesis. *Transportation Research Record, 1780*(1), 87–114. https://doi.org/10.3141/1780-10

Firey, W. (1945). Sentiment and symbolism as ecological variables. *American Sociological Review, 10*(2), 140–148. https://doi.org/10.2307/2085629

Gehl, J. (2013). *Cities for people*. Island press.

Gong, W., Rui, J., & Li, T. (2024). Deciphering urban bike-sharing patterns: An in-depth analysis of natural environment and visual quality in New York's Citi bike system. *Journal of Transport Geography, 115*, Article 103799. https://doi.org/10.1016/j.jtrangeo.2024.103799

Gouverneur, D. (2014). *Planning and design for future informal settlements: Shaping the self-constructed city*. Routledge.

Greenwell, B. M. (2017). pdp: an R Package for constructing partial dependence plots. *The R Journal, 9*(1), 421. https://doi.org/10.32614/RJ-2017-016

Gu, Z., & Zhang, X. (2021). Framing social sustainability and justice claims in urban regeneration: A comparative analysis of two cases in Guangzhou. *Land Use Policy, 102*, Article 105224. https://doi.org/10.1016/j.landusepol.2020.105224

Han, Y., Zhong, T., Yeh, A. G., Zhong, X., Chen, M., & Lü, G. (2023). Mapping seasonal changes of street greenery using multi-temporal street-view images. *Sustainable Cities and Society, 92*, Article 104498. https://doi.org/10.1016/j.scs.2023.104498

Hao, P., Geertman, S., Hooimeijer, P., & Sliuzas, R. (2012). The land-use diversity in urban villages in Shenzhen. *Environment and Planning A: Economy and Space, 44*(11), 2742–2764. https://doi.org/10.1068/a44696

Hao, P., Geertman, S., Hooimeijer, P., & Sliuzas, R. (2013). Spatial analyses of the urban village development process in Shenzhen, China. *International Journal of Urban and Regional Research, 37*(6), 2177–2197. https://doi.org/10.1111/j.1468-2427.2012.01109.x

Hegazy, I. R. (2016). Informal settlement upgrading policies in Egypt: Towards improvement in the upgrading process. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability, 9*(3), 254–275. https://doi.org/10.1080/17549175.2015.1009845

Jiang, Y., Han, Y., Liu, M., & Ye, Y. (2022). Street vitality and built environment features: A data-informed approach from fourteen Chinese cities. *Sustainable Cities and Society, 79*, Article 103724.

Kadıoğlu, D. (2022). *Producing identifiable neighborhoods: Race, stigma and struggle in Berlin-Neukölln* (pp. 1–23). Housing Studies. https://doi.org/10.1080/02673037.2022.2042494

Li, X., Li, Y., Jia, T., Zhou, L., & Hijazi, I. H. (2022). The six dimensions of built environment on urban vitality: Fusion evidence from multi-source data. *Cities, 121*, Article 103482. https://doi.org/10.1016/j.cities.2021.103482

Liu, S., & Zhang, Y. (2020). Cities without slums? China's land regime and dual-track urbanization. *Cities, 101*, Article 102652. https://doi.org/10.1016/j.cities.2020.102652

Lochner, M., & Ntema, J. (2013). The upgrading of an informal settlement in South Africa: Two decades onwards. *Habitat International, 39*(1), 85–95. https://doi.org/10.1080/02673037.2022.2042494

Matamanda, A. R., Mafuku, S. H., & Mangara, F. (2020). Informal settlement upgrading strategies: The Zimbabwean experience. *Sustainable Cities and Communities (Encyclopedia of the UN Sustainable Development Goals)*, 316–327. https://doi.org/10.1007/978-3-319-71061-7_8-1

Meng, X. (2019). Does a different household registration affect migrants' access to basic public health services in China? *International Journal of Environmental Research and Public Health, 16*(23), 4615. https://doi.org/10.3390/ijerph16234615

Niu, S., Hu, A., Shen, Z., Huang, Y., & Mou, Y. (2021). Measuring the built environment of green transit-oriented development: A factor-cluster analysis of rail station areas in Singapore. *Frontiers of Architectural Research, 10*(3), 652–668. https://doi.org/10.1016/j.foar.2021.03.005

Pan, W., & Du, J. (2021a). Towards sustainable urban transition: A critical review of strategies and policies of urban village renewal in Shenzhen, China. *Land Use Policy, 111*, Article 105744. https://doi.org/10.1016/j.landusepol.2021.105744

Pan, W., & Du, J. J. L. U. P. (2021b). *Towards sustainable urban transition: A critical review of strategies and policies of urban village renewal in Shenzhen* (Vol. 111), Article 105744. China.

Patel, K. (2013). The value of secure tenure: Ethnographic accounts of how tenure security is understood and realised by residents of low-income settlements in Durban, South Africa. In *Urban forum* (pp. 269–287). Springer.

Roberts, R. E., & Okanya, O. (2022). Measuring the socio–economic impact of forced evictions and illegal demolition; A comparative study between displaced and existing informal settlements. *The Social Science Journal, 59*(1), 119–138. https://doi.org/10.1016/j.soscij.2018.12.003

Rui, J. (2023a). Exploring the association between the settlement environment and residents' positive sentiments in urban villages and formal settlements in Shenzhen. *Sustainable Cities and Society, 104851*. https://doi.org/10.1016/j.scs.2023.104851

Rui, J. (2023b). Measuring streetscape perceptions from driveways and sidewalks to inform pedestrian-oriented street renewal in Düsseldorf. *Cities, 141*, Article 104472.

Rui, J., & Li, X. (2023). Decoding vibrant neighborhoods: Disparities between formal neighborhoods and urban villages in eye-level perceptions and physical environment. *Sustainable Cities and Society*, Article 105122.

Rui, J., & Othengrafen, F. (2023). Examining the role of innovative streets in enhancing urban mobility and livability for sustainable urban transition: A review. *Sustainability, 15*(7), 5709. https://doi.org/10.3390/su15075709

Shen, T., Yu, H., Zhou, L., Gu, H., & He, H. (2020). On hedonic price of second-hand houses in Beijing based on multi-scale geographically weighted regression: Scale law of spatial heterogeneity. *Journal of Economic Geography, 40*(3), 75–83. https://doi.org/10.4324/9781003033219-6

Shenzhen Planning and Natural Resources Bureau. (2019). *Shenzhen urban village comprehensive improvement plan* (pp. 2019–2025). http://www.sz.gov.cn/attachment/0/51/51731/1344686.pdf.

Tian, M., Yuan, L., Guo, R., Wu, Y., & Liu, X. (2022). Evaluations of spatial accessibility and equity of multi-tiered medical system: A case study of Shenzhen, China. *International Journal of Environmental Research and Public Health, 19*(5), 3017. https://doi.org/10.3390/ijerph19053017

Tjia, D., & Coetzee, S. (2022). Geospatial information needs for informal settlement upgrading–A review. *Habitat International, 122*, Article 102531. https://doi.org/10.1016/j.habitatint.2022.102531

United Nations. (2017). Goal 11: Make cities and human settlements inclusive, safe, resilient, and sustainable. *A New Era in Global Health: Nursing, 2030*, 339–344. https://doi.org/10.5771/9783748902065-419

von Seidlein, L., Alabaster, G., Deen, J., & Knudsen, J. (2021). Crowding has consequences: Prevention and management of COVID-19 in informal urban settlements. *Building and Environment, 188*, Article 107472. https://doi.org/10.1016/j.buildenv.2020.107472

Wang, Y. P., Wang, Y., & Wu, J. (2009). Urbanization and informal development in China: Urban villages in Shenzhen. *International Journal of Urban and Regional Research, 33*(4), 957–973. https://doi.org/10.1111/j.1468-2427.2009.00891.x

Wei, Y. D., Lin, J., & Zhang, L. (2020). E-commerce, taobao villages and regional development in China. *Geographical Review, 110*(3), 380–405. https://doi.org/10.1111/gere.12367

Wu, F., Li, L.-H., & Han, S. Y. (2018). Social sustainability and redevelopment of urban villages in China: A case study of Guangzhou. *Sustainability, 10*(7), 2116. https://doi.org/10.3390/su10072116

Wu, Y., Sun, X., Sun, L., & Choguill, C. L. (2020). Optimizing the governance model of urban villages based on integration of inclusiveness and urban service boundary (USB): A Chinese case study. *Cities, 96*, Article 102427. https://doi.org/10.1016/j.cities.2019.102427

Wu, W., & Wang, J. (2017). Gentrification effects of China's urban village renewals. *Urban Studies, 54*(1), 214–229. https://doi.org/10.1016/j.habitatint.2022.102531

Wu, J., Wang, M., Li, W., Peng, J., & Huang, L. (2015). Impact of urban green space on residential housing prices: Case study in Shenzhen. *Journal of Urban Planning and Development, 141*(4), Article 05014023. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000241

Wu, F., Xu, J., & Yeh, A. G.-O. (2006). *Urban development in post-reform China: State, market, and space*. Routledge.

Wu, F., Zhang, F., & Webster, C. (2013). Informality and the development and demolition of urban villages in the Chinese peri-urban area. *Urban Studies, 50*(10), 1919–1934. https://doi.org/10.1177/0042098012466600

Xia, C., Yeh, A. G.-O., & Zhang, A. (2020). Analyzing spatial relationships between urban land use intensity and urban vitality at street block level: A case study of five Chinese megacities. *Landscape and Urban Planning, 193*, Article 103669. https://doi.org/10.1016/j.landurbplan.2019.103669

Yang, C., & Qian, Z. (2022). The complexity of property rights embedded in the rural-to-urban resettlement of China: A case of Hangzhou. *Land Use Policy, 122*, Article 106394. https://doi.org/10.1016/j.landusepol.2022.106394

Yu, L., Xie, B., & Chan, E. H. (2019). Exploring impacts of the built environment on transit travel: Distance, time and mode choice, for urban villages in Shenzhen, China. *Transportation Research Part E: Logistics and Transportation Review, 132*, 57–71. https://doi.org/10.1016/j.tre.2019.11.004

Zeng, H., Yu, X., & Zhang, J. (2019). Urban village demolition, migrant workers' rental costs and housing choices: Evidence from Hangzhou, China. *Cities, 94*, 70–79. https://doi.org/10.1016/j.cities.2019.05.029

Zhang, Y. (2018). The credibility of slums: Informal housing and urban governance in India. *Land Use Policy, 79*, 876–890. https://doi.org/10.1016/j.landusepol.2017.05.029

Zhang, X. (2023). Informality and rapid urban transformation: A case study of regulating urban villages in Shenzhen. *Geojournal*, 1–15.

Zhang, M., Qiao, S., & Yeh, A. G.-O. (2021). Blemish of place: Territorial stigmatization and the depreciation of displaced villagers' resettlement houses in Chengdu, China. *Cities, 117*, Article 103330. https://doi.org/10.1016/j.cities.2021.103330

Zhang, J., Zhang, L., Qin, Y., Wang, X., & Zheng, Z. (2020). Influence of the built environment on urban residential low-carbon cognition in zhengzhou, China. *Journal of Cleaner Production, 271*, Article 122429. https://doi.org/10.1016/j.jclepro.2020.122429

Zhao, Y., Zhang, G., & Zhao, H. (2021). Spatial network structures of urban agglomeration based on the improved gravity model: A case study in China's two urban agglomerations. *Complexity, 2021*, 1–17. https://doi.org/10.1155/2021/6651444