The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-G-2025
ISPRS Geospatial Week 2025 “Photogrammetry & Remote Sensing for a Better Tomorrow…”, 6–11 April 2025, Dubai, UAE

Identifying Inefficient Urban Residential Land within Shenzhen City: An Approach Using Gaussian Mixture Model and Multi-Source Big Data

Yixin Liu[1], Shihao Liang[1], Renzhong Guo[1], Weixi Wang[1], Ding Ma[1], Ye Zheng[2], Libo Zhang[3], Qionghuan Liu[1], Xiaoming Li[1,*]

1 Research Institute for Smart Cities & MNR Key Laboratory of Urban Land Resources Monitoring and Simulation, School of Architecture and Urban Planning , Shenzhen University, Shenzhen 518060, China
2 Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, 315211, China
3 National Quality Inspection and Testing Center for Surveying and Mapping Products, Beijing 100830, China

Keywords: Inefficient urban land, Intellegent identification, multi-source geographical big data, GPS data, Urban Regeneration.

Abstract

As land resources become increasingly scarce, urban spatial development patterns in Chinese cities are now shifting from incremental expansion to inventory optimization. Accurate identification of inefficient urban residential land is the key for government to make regeneration policy and improve the living environment of more urban residents. With the rapid urbanization process and uneven resource allocation, China currently faces a declining trend in residential land use efficiency, significantly impacting urban residents' quality of life and satisfaction. Previous research mainly analyzed the land use efficiency of individual residential areas, neglecting to discuss the impact of surrounding environment and the utilization of urban dynamic big data. To address these issues, this paper employs the Gaussian Mixture Model (GMM) clustering method and integrates multi-source geographical big data to quantitatively characterize land use efficiency. Additionally, Spearman's coefficient analysis and Principal Component Analysis methods are applied for data dimensionality reduction. This methodology was initially applied in Bao'an District of Shenzhen and then expanded to cover the entire city. The research results validate the effectiveness and robustness of the approach. The study found that the kappa coefficients for inefficient residential communities and inefficient urban village residences are 0.637 and 0.721, respectively. Spatial analysis reveals that inefficient residential communities are dispersed, while inefficient urban village residences are concentrated in specific areas. This outcome provides important guidance for future government strategies on renewing inefficient urban spaces. At the methodological level, the Gaussian Mixture Model (GMM) clustering method, with its objectivity, multi-dimensionality, and precision, offers a new perspective and approach for studying inefficient urban residential land issues.

1. Introduction

In recent years, rapid global economic and population growth has led to large-scale urbanization, in which extensive migration has moved towards metropolitan areas (Xu et al., 2019). To meet the demands of urbanization, governments across the world have been expanding the supply of land for urban construction. As a result, many cities are now experiencing massive spatial expansion (Liu and Long, 2016). This long-standing extensive and outward-oriented urban growth model has given rise to a series of urban issues. These problems not only pose threats to urban residents live experience, but also result in some severe traffic congestion in metropolitan areas (Festus et al., 2020). The disorderly urban sprawl has given rise to a multitude of urban issues, ultimately imposing considerable economic, social, and environmental costs that adversely affect the quality of life for urban residents (Bai et al., 2020).

In response to these urban challenges, many developed cities, including Shenzhen in this research, are now shifting from rapid expansion to urban regeneration model (Liu et al., 2017). In general, urban regeneration involves the revitalization and optimization of land within built-up areas, aiming to enhance the current land-use efficiency, urban functionality and environmental quality (Liu et al., 2024). For example, in Europe the Berlin's adaptive renewal strategies of historic buildings and the renewal of traditional industrial space in Germany are successful urban regeneration cases (Petzet, 2012). In numerous developed cities across the globe, the transformation of old industrial bases and urban functions through urban regeneration has become a significant trend in the contemporary global landscape. This is particularly evident in those developed countries such as the United States, Australia, and France. After two decades of rapid urban expansion, China has significantly slowed its urban growth rate and is now entering a critical phase of urban regeneration (Han et al., 2022).

Since the last decade, the Chinese government has initiated a series of redevelopment programs targeting low-efficiency land use, with the goal of enhancing urban living quality and economic efficiency (Long, 2014). Unlike traditional urban sprawl models that focus on intensively developing unused urban land or farmland, in this stage, the government aims to identify and redevelop existing urban land that require regeneration, addressing the genuine needs of urban residents (Liu, 2020). However, the urban regeneration process is far more complex than merely constructing new urban districts. The regeneration process need to consider the interests of various stake-holders, while considering overall changes in structure and function in larger urban scale. Thus, it is important to accurately identify the inefficient urban land. But, currently, most governments and urban planners make urban regeneration decisions on simple social surveys, expert discussions, and planning experience, rather than on more scientific testing and objective evaluations (Long, 2014). This type of decision-making process may potentially misguide the regeneration objectives and, consequently, compromise the final outcomes (Chen, 2020). Therefore, it is imperative to accurately assess the efficiency of stock urban land and establish scientific methodologies for identifying inefficiency spaces.

According to the transition in governmental strategies toward land use, the identification of inefficient land have gained extensive attention from the urban scholars. Regarding traditional methods for identifying inefficient urban land, two primary approaches are currently employed: one involves the use of remote sensing imagery, leveraging technologies such as multi-scale segmentation; the other approach is based on social survey analysis and specialists' experience, where an evaluation criterion is established (Jin, 2023; Jiao and Huang, 2023). With the advancement of information and big-data technology, the incorporation of multi-source geographical big data has progressively become an integral part of urban research. Numerous big data sources are now employed in evaluating inefficient urban land use, thereby facilitating the regeneration process (Song et al., 2022; Sun et al., 2023).

In terms of the redevelopment of inefficient residential land, as our research target, many researchers has put forward different evaluation dimensions, including building quality, land use intensity, community satisfaction, ecological environments, security, infrastructural provision and financial systems (Wang et al., 2024). These strategies target different types of inefficient residential land, such as old residential areas and living zones where building quality are relatively low and those urban villages with excessive population density (Chang et al., 2024). These ongoing studies emphasize the close connection between inefficient residential lands and their residents' live, reflecting its profound impact on urban living experiences (Sun et al., 2023).

To sum up, this study aims to integrates multi-source geographical big data and hierarchical clustering methods to identify inefficiency urban residential land, mainly urban communities and urban villages respectively, seeking to contribute to traditional identification methods that often lack the incorporation of big data and automatic identification approaches. To test the efficiency of this methodology, we have applied it in the core urban districts of Shenzhen City, China. Our research objectives are: (1) to utilize multi-source geographical data for evaluating the efficiency of current land use; (2) to cluster the existing inefficient residential land based of multi-source big data, with the goal of identifying samples of low-efficiency urban land use; (3) to unveil the spatial characteristic of inefficient residential land. This research can contribute to the process of making urban regeneration policy both in China and countries all around the world.

2. Study area and data source

2.1 Study area

Shenzhen is a representative city of China's rapid urbanization and a good case for the research of identifying inefficient urban land. Located on the eastern bank of the Pearl River Delta and near Hong Kong, Shenzhen has rapidly evolved from a rural area into a globally renowned center for technology, finance, and manufacturing since it became a Special Economic Zone in 1980. In 2023, according to Shenzhen Yearbook 2024, its GDP reached 3.46 trillion yuan (approximately 477.48 billion US dollars). During the rapid urbanization process, Shenzhen has witnessed significant changes in land use. Between 1976 and 2020, the city's impervious surface area increased by 76.2 times. Thanks to Shenzhen's strong economic appeal, its GDP has continued to grow. By the end of 2023, the permanent resident population of Shenzhen had reached 17.79 million people (Shenzhen Yearbook 2024).

After decades of urban expansion, Shenzhen's land resources are nearing depletion. The contradiction between population growth and land scarcity has led to tight land supply and rising housing prices, making it increasingly difficult for people to afford priced housing. In addition, there is a large amount of inefficiently land in Shenzhen's residential areas, characterized by scattered layouts, unreasonable purposes, and dilapidated buildings. Urban villages, as an important source of existing housing in Shenzhen, generally have poor living quality, with problems such as high building density, insufficient supporting facilities, and safety hazards.

Our main experimental district, Bao'an District, located in the western part of Shenzhen core area, is a major industrial and population center of the city. The Bao'an District government has actively promoted the transformation of urban villages and the renewal of old residential areas. Therefore, in the coming years, numerous older urban communities and urban villages will confront the reality of urban regeneration. It is imperative for us to accurately identify inefficient residential land that warrants investment for urban renewal, thereby providing improved living spaces for the ever-growing population of Shenzhen. These are the reasons why we select Bao'an District as our main research area.

2.2 Research data

The data used in this study include vector data related to residential land use and various types of geo-spatial big data. These data consist of the building census data , the WorldPop open population dataset, POI and AOI data from AMap, community and urban village rental reference price data from the Shenzhen Government Data Open Platform, travel navigation flow data from AMap, GPS locating data, and street view image data from Baidu Map. Considering the availability and timeliness of the data, most data utilised in this research ranges from 2019 to 2024. Table 1 summarizes the formats, years, and sources of these datasets.

The building census data, obtained from the Shenzhen Housing and Construction Bureau, cover basic information on various types of buildings across the city, including building use, floor area, number of floors, etc.. This datasets provide detailed spatial distribution and attribute information of buildings in various communities. The World-pop datasets offer high-precision population data for China in 2020 at a 100-meter resolution. The POI data obtained from AMap cover the location information of various facilities and service points in Shenzhen, while the AOI data provide boundary information of different functional areas in the city. In this study, AOI data are used to extract the boundaries of residential land as the unit of analysis. The Shenzhen municipal government updates the community and urban village rental reference price data annually. The travel navigation flow data from AMap and GPS locating data record residents' daily movment and travel information, which are used to capture dynamic human activities.

Purpose	Dataset	Data Source	Year
Building attribute	Building census	Shenzhen Housing and Construction Bureau	2019
Population density	World-pop datasets	WorldPop Mainland China dataset for 2020	2020
Environmental quality	POI		2022
Base map	AOI	Amap (www.amap.com)	2024
Crowd vitality	travel navigation		2020
Regional economic	Rental reference	Shenzhen Government Data Open Platform	2023
Accuracy evaluation	Street view image	Baidu Map Open Platform	2024

Table 1. Description of datasets

3. Research Methods

This study has developed a method for identifying inefficient residential land based on Gaussian Mixture Model (GMM) clustering, which involves the following steps: (1) a set of multi-dimensional indicators of residential land are selected, such as building attributes, economic vitality, and environmental quality. These indicators are derived from various sources of geospatial big data. After standardizing the data indicators, redundant indicators are eliminated using Spearman correlation analysis, and data dimensionality is reduced through Principal Component Analysis (PCA). (2) Subsequently, Gaussian Mixture Model clustering is performed on the reduced-dimensional indicators within Bao'an District. Clusters that exhibit characteristics of inefficient utility are identified as the results of inefficient residential land through statistical analysis. (3) The identified inefficient residential land are then validated through random sampling, combined with field surveys and street view imagery. The accuracy of the method is assessed by calculating the kappa coefficient. (4) Finally, the Random Forest classification model and the TreeSHAP interpreter are employed to reveal the mechanisms influencing inefficient residential land. This provides planners with actionable insights to formulate targeted redevelopment strategies and enhance land use efficiency.

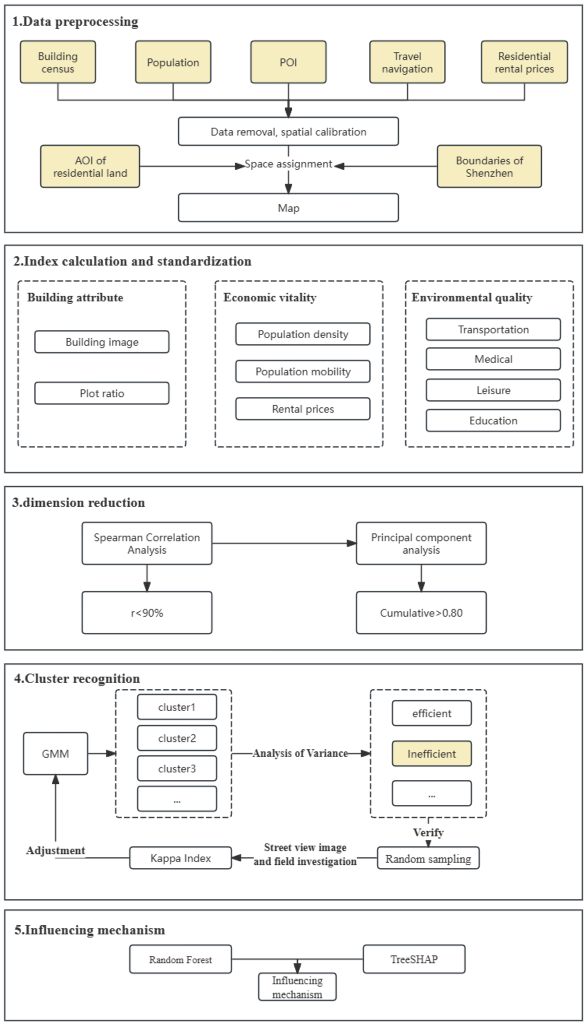

Figure 1. Workflow of this study

3.1 Principal Component Analysis (PCA)

PCA is a technique for data dimensionality reduction. It can retain as much of the original data information as possible using a smaller number of variables, known as principal components. PCA effectively reduces high-dimensional data indicators to lower dimensions while extracting the main features of the data, which aids in identifying the key factors of inefficient residential land. We further reduced the dimensionality of the extracted indicators using PCA to simplify the indicators. We

https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-993-2025 | © Author(s) 2025. CC BY 4.0 License.

retained the cumulative contribution rate of 80%. Subsequently, we employed the varimax rotation method to ensure that each indicator primarily loads on one principal component, seeking to enhance the interpretability of the model.

3.2 Kernel Density Estimation (KDE)

Kernel Density Estimation (KDE) is a method in Geographic Information Systems (GIS) for analyzing the probability density of points in real space. It effectively transforms discrete data points into a continuous density surface, revealing patterns of spatial distribution and trends of spatial decay. In this study, we utilized KDE to process Points of Interest (POI) data to measure the variations in the density of environmental features around residential land. This approach aims to quantify relevant environmental indicators that reflect the efficiency of residential land.

3.3 Gaussian Mixture Model (GMM)

The Gaussian Mixture Model (GMM) is a soft clustering method describing the data distribution through a linear combination of multiple Gaussian distributions. It assigns each data point a probability to each cluster, rather than rigidly categorizing data points into specific land-use feature groups. This probabilistic assignment allows GMM to better capture the uncertainty in the data, making it particularly suitable for situations where the boundaries of inefficient land use are ambiguous. In this study, GMM was implemented using Python to cluster residential areas and urban villages based on the distributions of the principal components. The number of clusters was generated by Bayesian Information Criterion (BIC), which help balances the goodness of the model and avoid over-fitting. The optimal model is selected by minimizing the BIC value.

3.4 Indicator selection

In this study, we processed multi-source data to quantify various indicators for evaluating efficiency of residential land. Population density and floor area ratio were calculated using the WorldPop datasets and building census data in ArcGIS Pro. Travel flow was obtained by constructing a complex network and calculating out-degrees using travel navigation data from AMap. Environmental indicators, including transportation, medical, education, and leisure convenience, were derived by applying kernel density calculations in ArcGIS Pro. Rental data for residential communities and urban villages, which were independent datasets, were cleaned in ArcGIS Pro and then assigned to the corresponding residential plots.

3.5 Accuracy Evaluation

To evaluate the accuracy and reliability of our model, we referred to the regulations on inefficient residential land, issued by the Ministry of Natural Resources and Shenzhen municipal government, and related literature to establish the following criteria for assessment: (1) Older buildings with low plot ratios, high building density, and underutilized residential units. (2)Poor infrastructure conditions, dilapidated buildings, and inadequate amenities. We employed a random sampling method to select plots from the results obtained through our methodology. Efficiency assessments were applied through street view imagery, social survey and field trip, and the kappa coefficient was calculated to measure agreement.

3.6 Impact Mechanism

To explore the impact mechanisms of inefficient residential land, we selected the Random Forest algorithm from machine learning to conduct regression modeling between the identified inefficient residential land and various indicators. We then employed the Tree SHAP method to interpret the regression results and reveal the impact of different indicators on inefficient residential land use. We utilized the classic Random Forest algorithm based on decision trees, which is widely used due to its simplicity, strong generalization ability, and resistance to over-fitting. Tree SHAP is an efficient implementation of the SHAP (Shapley Additive exPlanations) method specifically designed for machine learning models based on Random Forest. It quantifies the contribution of each feature to the model's predictions by calculating the average marginal contribution of each feature across all possible combinations of features.

4. Results

4.1 Indicator dimension reduction

The Spearman correlation coefficient was employed to analyze the relationships among the various indicators, as shown in Figures 2. This analysis revealed the inter-dependencies between the original indicators. Indicators with a correlation coefficient larger than 0.9 were deemed redundant and were thus eliminated. For residential communities, the analysis results indicated that the indicators for parking, medical facilities, training institutions, and school accessibility exhibited high correlation, with coefficients exceeding 0.9. Considering the importance of the medical system to residents, the medical accessibility indicator was retained, while the indicators for parking, training institutions, and school accessibility were removed.

In the case of urban villages, the overall correlation among the indicators decreased compared to residential communities. However, the indicators for parking, medical facilities, training institutions, and school accessibility still showed high correlation, with coefficients greater than 0.9. Therefore, the same approach was applied to the indicators for urban villages as for residential communities. After removing the redundant indicators, the final set of indicators included population density, population mobility, floor area ratio, regional economy, subway accessibility, medical accessibility, bus accessibility, shopping accessibility, and leisure accessibility.

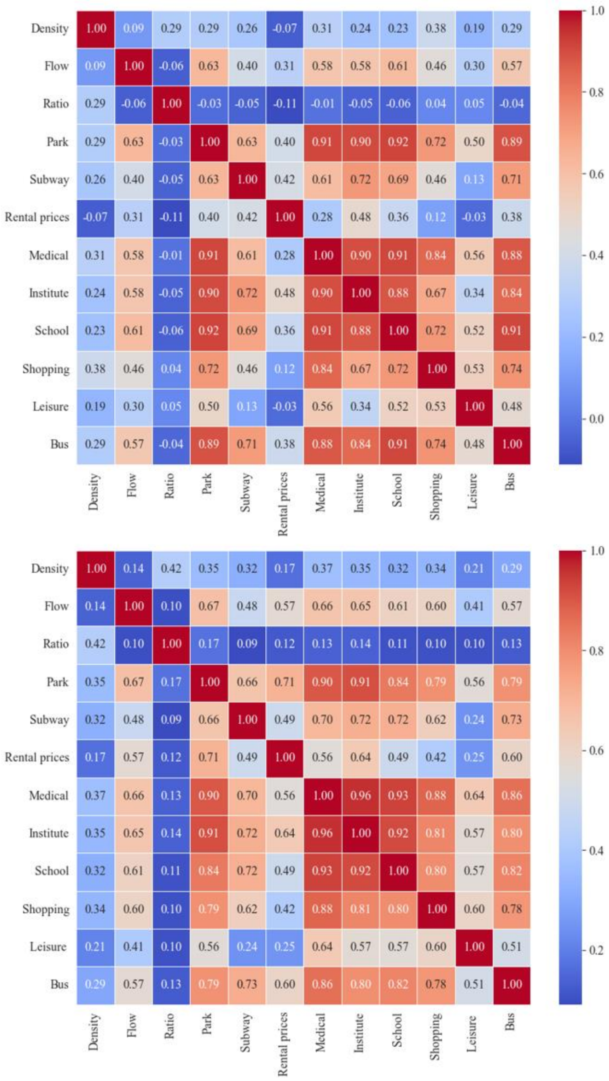

Figure 2. Correlation Coefficient Matrix for Low-Efficiency Urban communities (Upper) and Urban Villages (Lower)

To ensure comparability among the indicators within the indicator system, data transformation and standardization were applied to each indicator. For data exhibiting varying degrees of left or right skewness, transformations such as square root, logarithmic, and Box-Cox were applied to convert non-normally distributed data into normally distributed data. Subsequently, Z-score standardization was performed on the indicators. Z-score standardization transforms the values of different indicators to a standard scale with a mean of 0 and a standard deviation of 1, maintaining a standard normal distribution for different variables. This process enhances the clustering accuracy of Gaussian Mixture Model clustering methods, which are based on the assumption of normal distribution.

After completing the data preprocessing, Principal Component Analysis (PCA) was used to extract key indicators, further reducing the dimensionality of the multi-dimensional indicators. The principal components with a cumulative contribution rate of 80% were retained, resulting in 4 principal components for both residential communities and urban villages, with variance explanation rates of 80.93% and 84.94%, respectively. To enhance the interpretability of the obtained principal components, the varimax rotation method was employed to improve the interpretability of the factor loadings. The factor loading after varimax rotation indicates that each original indicator is highly correlated with one of the resulting principal components, facilitating the understanding of the latent variables represented by each factor in the subsequent clustering analysis.

4.2 Identified inefficient residential land

A Gaussian Mixture Model (GMM) analysis was conducted on residential communities and urban villages in Bao'an District,

with the Bayesian Information Criterion (BIC) used to identify the optimal number of clusters. The BIC results indicated that the optimal number of clusters for residential communities and urban villages was 4 and 3. To identify relatively inefficient and efficient residential land use categories from the clustering results, a one-way ANOVA was performed on the clustering outcomes. The analysis revealed that, for each indicator, there were statistically significant differences among different clusters ($p < 0.01$). This suggests that the different categories obtained through GMM exhibit significant variations in residential land use efficiency.

For urban villages, Cluster 2 exhibited the lowest population density, population mobility, subway accessibility, and regional economy. These urban villages may be located far from subway stations, with poor community vitality and relatively low housing prices. Compared with the other clusters, Cluster 2 displayed obvious inefficiencies, and can be considered as relatively inefficient urban villages. For residential communities, the economic data among various clusters showed little variation; however, the fourth cluster of residential areas is identified as relatively inefficient due to their low plot ratio and low vitality of residents' travel activities, as well as poor accessibility to subway and bus services (See Table 2).

Land	Indicator	F-value	Cluster 1 mean	Cluster 2 mean	Cluster 3 mean	Cluster 4 mean
urban villages	Density	20.815	0.249	-0.292	0.587	
	Flow	123.609	1.254	-0.381	-0.353	
	Ratio	14.698	0.112	-0.231	0.567	
	Subway	82.238	1.113	-0.352	-0.269	
	Rental prices	150.755	1.313	-0.350	-0.519	
	Medical	183.830	1.357	-0.494	-0.133	
	Shopping	106.485	1.176	-0.442	-0.073	
	Leisure	38.302	-0.354	0.381	-0.731	
	Bus	282.251	-1.493	0.487	0.318	
communities	Density	26.768	-0.044	-0.144	0.625	-0.268
	Flow	137.940	-0.577	0.726	0.746	-0.213
	Ratio	4.691	0.139	-0.218	-0.055	-0.004
	Subway	119.046	-0.555	0.420	0.925	-0.131
	Rental prices	105.688	1.139	1.199	1.188	1.199
	Medical	413.974	-0.650	0.602	1.361	-0.443
	Shopping	245.721	-0.450	0.287	1.355	-0.501
	Leisure	80.738	-0.124	0.616	0.502	-0.664
	Bus	638.171	0.741	-0.544	-1.513	0.363

Table 2. ANOVA and descriptive statistics for residential land.

Based on the overall identification results of inefficient residential land in Bao'an District (Figure 3), we could find that inefficient urban communities are mostly concentrated on the fringes of densely built-up areas, particularly near the Guangzhou-Shenzhen-Hong Kong Expressway and Qianhai Development Zone. The spatial distribution of these inefficient urban villages is relatively scattered, but they are overall located to the northern part of Bao’an district.

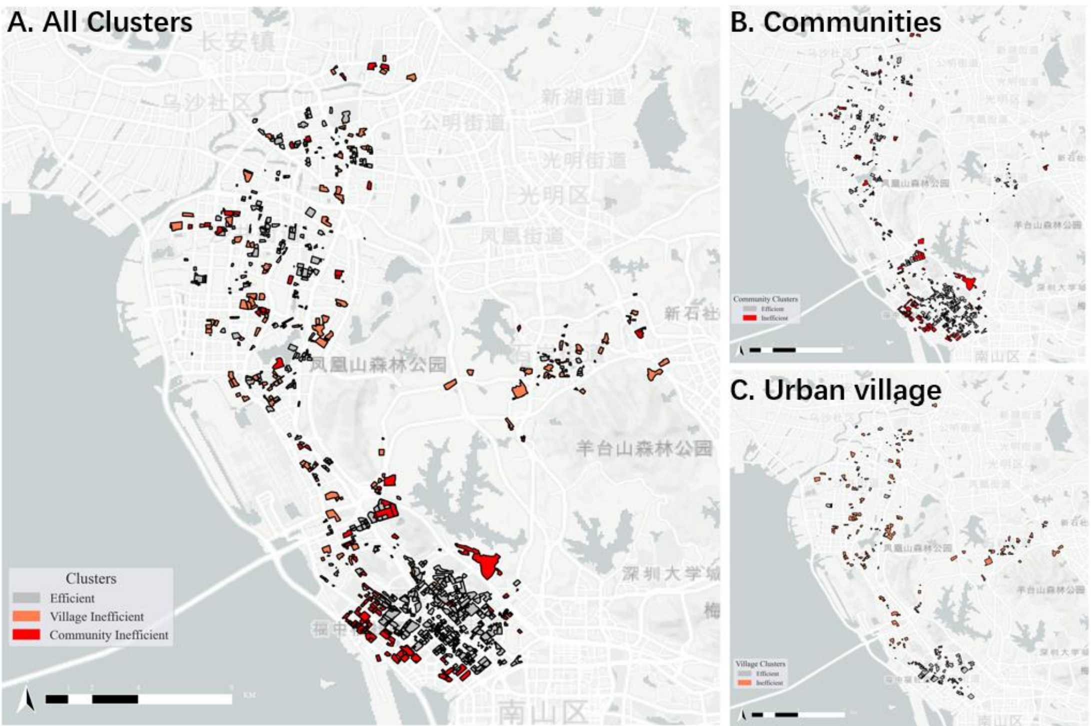

Figure 3. A. Results of identified inefficient residential land，urban communities（B）and urban village（C）respectively

4.3 Accuracy Evaluation

Based on the criteria outlined in Section 3.5, this study validated the reliability of the identification method for inefficient residential communities and urban villages in Bao'an District through random sampling, using a combination of street view images and field surveys of a few selected plots. The number of sampled plots for residential communities and urban villages were 102 and 80, respectively. The validation results, summarized in the confusion matrix in Table 3, indicated that the identification accuracy for inefficient residential land and urban villages was 0.843 and 0.875, respectively, both exceeding 0.8. Additionally, the kappa coefficients for the identified inefficient residential communities and urban villages were 0.637 and 0.721, respectively, both surpassing 0.6. These results demonstrate a high level of consistency between the identification outcomes and the actual conditions, thereby confirming the reliability and accuracy of the identification method employed in this study.

Predicted / Autual	inefficient residential communities		inefficient urban village residences	
	Positive	Negative	Positive	Negative
Positive	24	11	22	7
Negative	5	62	3	48
Accuracy	0.843		0.875	
Kappa	0.637		0.721	

Table 3. Confusion matrix, Accuracy and Kappa.

4.4 Impact Mechanism

To explore the factors influencing the efficiency of inefficient residential land use, we employed the clustering results as the dependent variable and various indicator as an independent variable for the Random Forest classification. Subsequently, we utilized Tree SHAP to provide both global and local interpretations of the model, thereby obtaining the contribution of individual factor to the efficiency of inefficient residential land use. This approach elucidates how each indicator influences the efficiency of residential spaces, as illustrated in Figure 4:

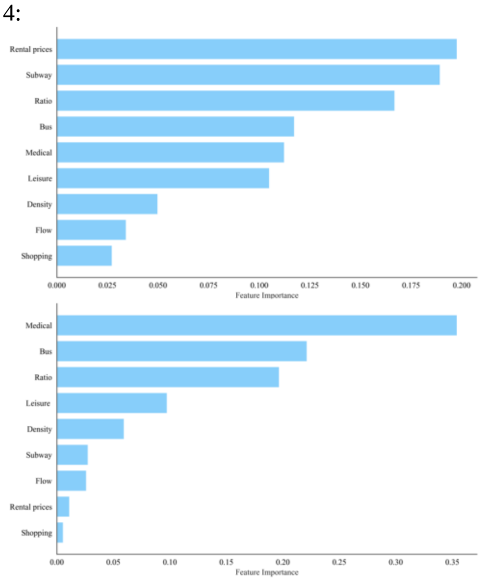

Figure 4. Importance of indicators for inefficient urban communities (Upper) and urban villages (Lower)

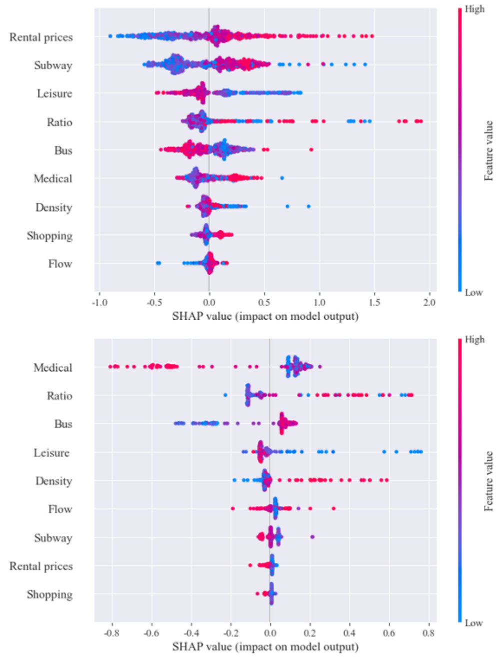

Figure 5. Impact on model output for inefficient residential communities (Upper) and inefficient urban village residences (Lower)

As depicted in Figure 5, for residential communities, the three most significant factors contributing to low land efficiency are rental prices, subway accessibility, and plot ratio that have a notable positive impact. For urban villages, the three main factors are healthcare accessibility, bus accessibility, and plot ratio. Among them, bus accessibility and plot ratio have a positive impact on land use efficiency, while healthcare accessibility has a significant negative impact. The significant distinction is that, for upscale urban communities, housing prices have a particularly crucial impact on land efficiency, whereas in urban villages, medical facilities play a prominent role in safeguarding the quality of life for impoverished residents. The plot ratio and public transportation accessibility are crucial for both types of residential land.

5. Discussion

Our proposed method in evaluating the inefficiency of urban residential has been tested in the Bao'an District of Shenzhen City. The results show that our model is capable of accurately identifying inefficient urban residential spaces, with kappa coefficients of 0.637 for urban communities and 0.721 for urban villages, respectively. However, when our model is applied to the entire area of Shenzhen City, the accuracy of the identification outcomes decreases to approximately 0.5 due to insufficient data collection and a reduction in data variety. This issue should be addressed more effectively in future research. Another shortcomeing of our research is that we can develop clear acitivity model of urban residents from GPS data, but have not sufficiently use in the evaluating the efficiency of urban land. Another limitation of our research is that although we have the capability to develop clear activity models of random-sampled urban residents based on GPS navigation data, we have not fully utilized this potential in evaluating the efficiency of urban land use. This represents an area for further exploration and improvement in our future work.

Acknowledgements

This research was funded by National Key Rsearch and Development Project of China (Grant No. 2023YFC3804805) , the Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources (Grant No. KF-2022-07-008), Tencent "Rhinoceros Birds" - Scientific Research Foundation for Young Teachers of Shenzhen University and Shenzhen Science and Technology Program (No. 20231121100546001).

References

Bai, Y., Zhou, W., Guan, Y., Li, X., Huang, B., Lei, F., Yang, H. and Huo, W., 2020. Evolution of policy concerning the readjustment of inefficient urban land use in China based on a content analysis method. *Sustainability*, 12(3), p.797.

Chang, Y., Li, G., Zhang, P., Liu, Y., Chen, Z., Xing, G. and Li, M., 2024. Relationships among six urban air pollutants and identification of pollution types–A case study of Chinese cities above prefecture level. Atmospheric Pollution Research, 15(7), p.102160.

Chen, K., Long, H., Liao, L., Tu, S. and Li, T., 2020. Land use transitions and urban-rural integrated development: Theoretical framework and China's evidence. *Land use policy*, 92, p.104465.

Festus, I.A., Omoboye, I.F. and Andrew, O.B., 2020. Urban sprawl: environmental consequence of rapid urban expansion. Malaysian Journal of Social Sciences and Humanities (MJSSH), 5(6), pp.110-118.

Han, B., Jin, X., Wang, J., Yin, Y., Liu, C., Sun, R. and Zhou, Y., 2022. Identifying inefficient urban land redevelopment potential for evidence-based decision making in China. *Habitat International*, 128, p.102661.

Jiao, H. and Huang, J., 2023. Study on inefficient land use determination method for cities and towns from a city examination perspective. Applied Mathematics and Nonlinear Sciences, 8(1), pp.2425-2438.

Jin, R., Huang, C., Wang, P., Ma, J. and Wan, Y., 2023. Identification of Inefficient Urban Land for Urban Regeneration Considering Land Use Differentiation. Land, 12(10), p.1957.

Liu, S., Xiao, W., Li, L., Ye, Y. and Song, X., 2020. Urban land use efficiency and improvement potential in China: A stochastic frontier analysis. Land Use Policy, 99, p.105046.

Liu, Y., and Long, H., 2016. Land use transitions and their dynamic mechanism: The case of the Huang-Huai-Hai Plain. *Geogr. Sci*, 26, 515–530.

Long, H., 2014. Land use policy in China: Introduction. Land use policy, 40, pp.1-5.

Petzet, M. and Heilmeyer, F., 2012. Reduce, reuse, recycle. Architecture as resource.

Wang, X., Bao, X., Ge, Z., Xi, J. and Zhao, Y., 2024. Identification and Redevelopment of Inefficient Residential Landuse in Urban Areas: A Case Study of Ring Expressway Area in Harbin City of China. Land, 13(8), p.1238.

Song, Y., Lyu, Y., Qian, S., Zhang, X., Lin, H. and Wang, S., 2022. Identifying urban candidate brownfield sites using multi-source data: The case of Changchun City, China. Land Use Policy, 117, p.106084.

Sun, Y., Xiao, D. and Huang, L., 2023. Exploration of the reconstruction of old residential districts in historical urban areas of Guangzhou under the concept of complete community. South Archit, 9, pp.62-69.

Sun, Y., Hu, H., Han, Y., Wang, Z. and Zheng, X., 2023. Large-Scale Automatic Identification of Industrial Vacant Land. ISPRS International Journal of Geo-Information, 12(10), p.409.

Xu, D., Deng, X., Guo, S. and Liu, S., 2019. Labor migration and farmland abandonment in rural China: Empirical results and policy implications. Journal of environmental management, 232, pp.738-750.