



**Integrating remote sensing and geospatial big data on urban
land use mapping and impact analysis of urban land use
changes on green space distribution**

Jiadi Yin

Student number – 20155716

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

March 2022

I, Jiadi Yin, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Urban land use significantly affects the urban environment, especially urban green space. Urban green spaces, such as parks, farmland, and gardens, can provide extensive benefits. Previous studies have proved that the ecological and socio-economic benefits of urban green spaces are highly dependent on their spatial patterns. The urban green landscape can act as a planning strategy to improve the urban environment and make the city achieve the Sustainable Development Goals (SDGs). However, the impact of urban land use changes on urban green space distributions has largely been ignored due to the lack of high-quality urban land use maps and impact analysis methods. Therefore, this Ph.D. thesis aims to improve the existing urban land use maps for investigating the dynamics of urban land use and assess the impact of urban land use changes on urban green space distribution. It is of great significance to study the urban functional patterns and urban green space distribution under the background of the transition from urban land cover to urban land use. The research uses Hangzhou as a case study, as it is one of the most typical cities in terms of urbanization, population growth, economic development, and land use changes.

Chapter 1 described the background of the study and introduces the research aims, objectives, main definitions, and structures in this thesis. In **Chapter 2**, I summarized the relevant literature, focusing on the urban land use mapping as well as the impact of urban land use on urban green space distribution. **Chapter 3** introduced the methods

mainly used in this thesis. The results in **Chapter 4** (research chapter 1) demonstrated that integrating remote sensing (RS) and geospatial big data (GBD) provides new opportunities for urban land use classification. The integration strategies were categorized into decision-level integration (DI) and feature-level integration (FI) according to the fusion mode and process. **Chapter 5** (research chapter 2) tested the DI and FI methods for mapping urban land use in Hangzhou city and concluded that the differences in data sources, features, classifiers, training samples, and land use types might lead to the different classification results according to the process of integration methods. **Chapter 6** (research chapter 3) investigated the spatial-temporal dynamics of urban land use changes from 2017 to 2021, indicating the increase of institution and residence parcels increased, and the decrease in business and open space parcels. **Chapter 7** (research chapter 4) found that the urban green space distribution was affected significantly by urban land use changes in Hangzhou city. The area of urban green space has increased from 2017 to 2021 totally, which was mainly concentrated in the urban core regions, indicating Hangzhou city has made remarkable achievements in green space planning within the urban center. In addition, the residence parcels have the largest increase in urban green space areas, while the areas of urban green space in the open space parcels have decreased. There are certain differences in the fragmentation, complexity, aggregation, and uniformity of urban green space patches within different urban land use change types. **Chapter 8** synthesized the results

obtained from this thesis with each of the research chapters. An outlook and recommendations for future research are also presented.

Acknowledgments

Firstly, I would like to thank my supervisors, Dr. Ping Fu, Professor Ali Cheshmehzangi, and Professor Jinwei Dong for their incredible support and guidance throughout the Ph.D. They have always been there when I needed them. Their drive and passion for their work have been a constant source of inspiration for me over the past three years. I am extremely grateful for the opportunity they have given me. Their knowledge, advice, and enthusiasm for science will stay with me in future endeavors.

I would like to thank Professor Giles Foody, my internal assessor, for his brilliant and useful comments and suggestions every year. Thanks to Professor Tao Lin at the Institute of Urban Environment, the Chinese Academy of Sciences, my external assessor, and Professor Foody for letting me know that the viva voce can be enjoyable and again for the great discussion we had on it.

I would also like to extend my gratitude to Dr. Nicholas A. S. Hamm and Dr. Zhichao Li for their help and advice on the Ph.D. project. Thanks, in particular, go to my hosts Professor Doreen Boyd and Dr. Geertje van der Heijden at Nottingham University's, School of Geography for their kind help and support.

This Ph.D. would have not been possible without the help of my colleagues at the LUGC group at the Institute of Geographic Sciences and Natural Resources Research, the Chinese Academy of Sciences (IGSNRR), and the geospatial group at the

University of Nottingham, Ningbo, China. I appreciate their patience in answering all of my queries and for their significant intellectual contributions to my study.

This research would not have been possible without my family, mom, dad, and Jialun. Thanks a lot for all the love and support, for encouraging me to insist when I just wanted to give up, and for all the sacrifices you have made on my behalf. My special thanks to my husband, Dr. Xiaoyue Wang. You are my greatest inspiration, and without your unwavering belief in me, this stage of my life would have never been possible.

Table of Contents

Chapter 1: Introduction	19
1.1 Background	19
1.2 Novel aspects of the research	23
1.3 Aims and objectives	23
1.4 Definitions	24
1.5 Thesis structure	25
Chapter 2: Literature review	27
2.1 Introduction	27
2.2 Urban land use.....	27
2.2.1 Evolution from urban land cover to urban land use	27
2.2.2 Integrating RS and GBD on urban land use mapping	30
2.2.3 Urban land use change.....	49
2.2.4 Accuracy assessment	50
2.3 Impact analysis of urban land use	52
2.3.1 Urban green space	53
2.3.2 Impact of urban land use on urban green space.....	57
2.4 Summary	58
Chapter 3: Methodological approach.....	61
3.1 Introduction	61
3.2 DI-based mapping method	62
3.3 FI-based mapping method.....	63
3.4 Accuracy assessment.....	65
3.5 Attribute importance analysis.....	66
3.6 Urban green space mapping method	69
3.7 Impact analysis.....	71

Chapter 4: Integrating remote sensing and geospatial big data for urban land use mapping: A review	74
4.1 Introduction	74
4.2 Evolution from urban land cover to urban land use	76
4.3 RS and GBD features for urban land use mapping	80
4.3.1 RS-based features	81
4.3.2 GBD-based features.....	84
4.4 Integration of RS and GBD for urban land use mapping	88
4.4.1 Feature-level integration.....	88
4.4.2 Decision-level integration.....	90
4.5 Discussion	92
4.5.1 Advantages and disadvantages of FI-based and DI-based classification	92
4.5.2 Limitations and future consideration for RS and GBD integration.....	93
4.5.3 Potential applications of urban land use maps derived from RS and GBD integration.....	94
4.6 Conclusion.....	97
Chapter 5: Decision-level integration and feature-level integration of remote sensing and geospatial big data for urban land use mapping	99
5.1 Introduction	99
5.2 Related Work of RS and GBD Integration Used in Urban Land Use Mapping	101
5.2.1 DI-based Urban Land Use Mapping.....	102
5.2.2 FI-based Urban Land Use Mapping	103
5.3 Case study	104
5.3.1 Study Site.....	104
5.3.2 Data Source and Preprocessing	105

5.3.3 Methods	107
5.4 Results	118
5.4.1 Quantitative Performance of DI-based and FI-based Urban Land Use Mapping.....	118
5.4.2 Qualitative Performance of DI-based and FI-based Urban Land Use Mapping.....	120
5.5 Discussion	122
5.5.1 Summary of the Advantages and Disadvantages of the Two Methods	123
5.5.2 The improved Method for Urban Land Use Mapping.....	125
5.6 Conclusions	126
Chapter 6: Quantifying spatiotemporal patterns of urban land use change: A case study in Hangzhou, China	128
6.1 Introduction	128
6.2 Materials and methods	130
6.2.1 Study site	130
6.2.2 Data.....	131
6.2.3 Methods	133
6.3 Results	139
6.3.1 Urban land use map in 2021	139
6.3.2 Urban land use map in 2017	143
6.3.3 Urban land use changes from 2017 to 2021	145
6.4 Discussion	149
6.5 Conclusion.....	151
Chapter 7: Investigating the urban green space distribution impacted by urban land use changes: A case study in Hangzhou, China	153
7.1 Introduction	153

7.2 Materials and methods	155
7.2.1 Study site	155
7.2.2 Data.....	156
7.2.3 Methods	158
7.3 Results	163
7.3.1 Urban green space distribution.....	163
7.3.2 Spatial distribution of urban green space	164
7.3.3 Spatial distribution of urban green space in different urban land use changes	171
7.4 Discussion	179
7.4.1 Differences in the distribution of urban green space.....	179
7.4.2 The development of urban green space distribution in China.....	182
7.5 Conclusion.....	184
Chapter 8: Synthesis	185
8.1 Research chapter summary.....	185
8.2 Future work	188
8.2.1 The improved method for urban land use mapping.....	188
8.2.2 The mixed urban land use parcels	191
8.2.3 The configuration and composition of urban green space.....	192
8.3 Concluding remarks	194
References:.....	195
Appendix 1: Integrating remote sensing and geospatial big data for urban land use mapping: A review	245
Appendix 2: Decision-level integration and feature-level integration of remote sensing and geospatial big data for urban land use mapping.....	257

List of Figures

Figure 2-1 Urban green space samples.	54
Figure 2-2 Urban green space benefits.	54
Figure 3-1 Methodology workflow.	62
Figure 3-2 DI-based categorization strategies.	62
Figure 3-3 FI-based categorization strategies.	64
Figure 3-4 RF classification code in GEE	64
Table 3-2 Confusion matrix of classification results in the combination of spectral, texture, and frequency.	68
Table 3-3 Confusion matrix of classification results in the combination of spectral, texture, and density.	68
Figure 3-5 Testing samples of urban green space mapping: (a) urban green space samples in 2017; (b) urban green space samples in 2021.	69
Figure 3-6 Operation of statistical analysis	72
Figure 3-7 Operation of hotspot analysis.	73
Figure 4-1 Transformation of the need for urban land products from physical attributes to socioeconomic attributes due to enhanced anthropogenic activities. The three boxes above represent the transformation from urban land cover (e.g., impervious surface and pervious surface) to urban land use (e.g., traffic, institution, urban lake, residential green space, and others). It should be noted that the change of area proportions of different urban land use types is not reflected.	80
Figure 4-2 Summary of the features of RS and GBD. The dotted black box and red box in the middle show the commonly used RS data (e.g., MODIS, Landsat, Sentinel) and GBD (e.g., traffic, social media data, geo-tagged photos), respectively. The upper dotted black boxes represent the features (spectral, textural, temporal, and spatial) extracted from RS, while the dotted red boxes below represent the features (spatial, temporal, semantic, and sequence) extracted from GBD.	81

Figure 4-3 FI-based categorization strategies.	88
Figure 4-4 DI-based categorization strategies.	90
Figure 4-5 Urban land use map integrating RS and GBD for urban management.	95
Figure 5-1 A general framework of the two integration strategies: Decision-level integration (DI), and Feature-level integration (FI).	102
Figure 5-2 The study site (a) Sentinel-2 true-color composition image with OSM river network; (b) Spatial distribution of Point-of-interests (POI); (c) Road network extracted from OpenStreetMap (OSM).	105
Figure 5-3 This is a figure. Schemes follow the same formatting. Research methodological framework of mapping urban land use with decision-level integration (DI) and feature-level integration (FI). (a) Urban parcel generation based on the road network and river data; (b) Collection of testing and training parcels; (c) DI-based urban land use mapping (DI); (d) FI-based urban land use mapping.	108
Figure 5-4 Illustration of the institution and residence parcels in Baidu map and Baidu street view: (a) represents the institution parcel in Baidu map; (b), (c), (d) and (e) represent the four perspectives of institution parcel in Baidu street view; (f) represents the residence parcel in Baidu map; and (g), (h), (i) and (j) represent the four perspectives of residence parcel in Baidu street view. The Chinese labels for (a) and (f) represent the name of POIs.	111
Figure 5-5 Spatial distribution and illustration of training and testing parcels for the four urban land use types. (a) represents the spatial distribution of training and testing parcels; (b), (d), (f), and (h) represents the Baidu map image for institution, residence, business, and open space, respectively; and (c), (e), (g) and (i) represent the Baidu street view for the institution, residence, business, and open space, respectively. The Chinese labels for (b), (d), (f), and (h) represent the name of POIs.	112
Figure 5-6 (a) Distribution of the pervious and impervious classes; (b) Distribution of the built-up and non-built-up parcels.	113

Figure 5-8 Illustration of the (a) DI-based and (b) FI-based urban land use maps in Hangzhou city.	121
Figure 5-9 Classification results of DI and FI methods and testing samples in the test regions. Subareas A, B, C, and D were selected randomly in Hangzhou city, and each region was shown with DI-based classification, FI-based classification, and Baidu map. The Chinese labels for the subfigure of right column represent the name of POIs. ..	122
Figure 6-1 The study site.	131
Figure 6-2 Sentinel-2 images. (a) 2017 Sentinel image; (b) 2021 Sentinel image. ...	132
Figure 6-3 OpenStreetMap road network data. (a) 2017 OSM data; 2021 OSM data.	132
Figure 6-4 Spatial distribution of Points-of-interest (POIs). (a) 2017 POIs; (b) 2021 POIs.....	133
Figure 6-5 Methodology framework. (a) Urban parcel generation; (b) Accuracy assessment; (c) DI-based classification; (d) FI-based classification.....	134
Figure 6-6 Spatial distribution of training and testing parcels in 2017. (a) Spatial distribution of parcels selected randomly; (b) Testing parcels; (c) Training parcels.	137
Figure 6-7 Spatial distribution of training and testing parcels in 2021. (a) Spatial distribution of parcels selected randomly; (b) Testing parcels; (c) Training parcels.	138
Figure 6-8 Illustration of the (a) DI-based and (b) FI-based urban land use maps in Hangzhou city in 2021.....	143
Figure 6-9 Illustration of the FI-based urban land use map in Hangzhou city in 2017.	145
Figure 6-10 Post-processing of the urban land use map in 2017 (a) and 2021 (b). ...	146
Figure 6-11 Urban land use changes from 2017 to 2021. I:Institution; R: Residence; B:Business; O:Open space.	148
Figure 7-1 Study area.....	156
Figure 7-2 Gaofen-2 images	157

Figure 7-3 Urban land use maps in 2017 and 2021	158
Figure 7-4 Methodology framework.....	159
Figure 7-5 Urban green space distribution in 2017 (a) and 2021 (b).....	163
Figure 7-6 Hotspot analysis of urban green space distribution at ring-level in 2017 (a) and 2021 (b)	165
Figure 7-7 Hotspot analysis of urban green space distribution at ring-level from 2017 to 2021	166
Figure 7-8 Regression analysis of the relationship between urban green space areas and patch areas of different urban land use types in 2017.....	168
Figure 7-9 Regression analysis of the relationship between urban green space areas and patch areas of different urban land use types in 2021.....	169
Table 7-4 Regression analysis of the relationship between urban green space areas and patch areas of different urban land use types in 2021.....	169
Table 7-5 Regression analysis of the relationship between urban green space areas and patch areas of different urban land use types in 2021.....	170
Figure 7-10 Landscape analysis of urban green space distribution at ring-level in 2017 and 2021	171
Figure 7-11 Areas and proportion analysis of urban green space in the institution parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space).....	172
Figure 7-12 Areas and proportion analysis of urban green space in the residence parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space).....	173
Figure 7-13 Areas and proportion analysis of urban green space in the business parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space).....	174

Figure 7-14 Areas and proportion analysis of urban green space in the open space parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space)	175
---	-----

List of Tables

Table 2-1 Key terms for urban land use	32
Table 2-2 Integration publications for urban land use classification	33
Table 2-3 Contingency table for accuracy assessment (referring to <u>Card (1982)</u>)	51
Table 3-1 Summary of RS and GBD features used in parcel-level urban land use mapping.....	67
Table 3-4 Classification accuracy of each urban land use type via different feature combinations of features. a: Spectral, texture, and frequency, b: Spectral, texture, and density.	68
Table 3-5 Confusion matrix of the urban green space mapping results in 2017. UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map (N: Non-green space; G: Green space). ..	70
Table 3-6 Confusion matrix of the urban green space mapping results in 2021. UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map (N: Non-green space; G: Green space). ..	70
Table 4-1 Comparison of existing global urban land cover products.	78
Table 4-2 Comparison of concepts related to GBD in previous studies.	85
Table 4-3 Summary of the GBD and relevant features used in urban land use mapping.	86
Table 5-1 The framework used for aggregating initial Gaode POI types.	107

Table 5-2 The road and river levels and their corresponding buffer sizes used in this study.	109
Table 5-4 Spectral and textural features from RS, and density features from GBD used in FI-based classification.	117
Table 5-5 Confusion matrix of the impervious and pervious surface extraction results. UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map.	119
Table 5-6 Confusion matrix of DI-based classification results (I: Institution; R: Residence; B: Business; O: Open Space). UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map. ..	119
Table 5-7 Confusion matrix of FI-based classification results (I: Institution; R: Residence; B: Business; O: Open Space). UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map. ..	120
Table 6-1 Summary of RS and GBD features used in parcel-level urban land use mapping.....	136
Table 6-2 Contingency table for accuracy assessment.....	139
Table 6-3 Confusion matrix of the impervious extraction results in 2021. UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map.....	140
Table 6-4 Confusion matrix of DI-based classification results in 2021(I: Institution; R: Residence; B: Business; O: Open Space). UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map. ..	141
Table 6-5 Confusion matrix of FI-based classification results in 2021 (I: Institution; R: Residence; B: Business; O: Open Space). UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map. ..	142

Table 6-6 Confusion matrix of FI-based classification results in 2017 (I: Institution; R: Residence; B: Business; O: Open Space). UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map. ..	144
Table 6-7 Quantitative in urban land use types in 2017 and 2021.....	147
Table 6-8 Quantitative analysis of different urban land use changes from 2017 to 2021	149
Table 7-1 Gaofen-2 Satellite Sensor Specifications.....	157
Table 7-2 Landscape metrics	161
Table 7-3 Analysis of urban green space area and proportion at ring-level in 2017 and 2021.....	167
Table 7-6 Landscape analysis of urban green space in the institution parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space)	176
Table 7-7 Landscape analysis of urban green space in the residence parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space)	177
Table 7-8 Landscape analysis of urban green space in the business parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space)	178
Table 7-9 Landscape analysis of urban green space in the open space parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space)	179

Chapter 1: Introduction

1.1 Background

With the advent of Anthropocene ([Ellis and Ramankutty 2008](#); [Steffen et al. 2011](#)), urbanization is expected to accelerate, with the urban population expected to rise from 4.2 billion (57.5%) in 2018 to over 6.7 billion (69.1%) in 2050 ([Seto et al. 2011](#)). Rapid urbanization is exerting pressure on the urban environment, especially urban green spaces ([Huang et al. 2018b](#); [Hersperger et al., 2018](#); [Seto and Kaufmann, 2003](#)). Urban green space, such as parks, greenways, woods, squares, farmland, green roofs, and gardens, is important not only to the physical and psychological health of urban dwellers, it most essentially reduces the urban heat island, mitigates the effects of climate change, and improves air conditions ([Aram et al. 2019](#); [Braubach et al. 2017](#); [Kondo et al. 2018](#)). The ecological and socio-economic functions of urban green spaces are considered to be highly dependent on their spatial patterns ([Li et al. 2015](#); [Song et al. 2021](#)). For instance, the increasing area of urban green space is highly correlated with the reduction of urban heat island (UHI) ([Meng et al. 2018](#)). Also, there is a strong correlation between urban green space quantity and gross domestic product ([Chen et al. 2017](#)). According to the target 11.7 of Sustainable Development Goals (SDGs), cities should provide universal access to safe, inclusive, and accessible, green and public

spaces, particularly for women and children, older persons, and persons with disabilities by 2030 ([Koch and Krellenberg 2018](#)). The urban green landscape can act as a planning strategy to improve the urban environment and make the city achieve the SDGs.

Several studies have demonstrated the importance of urban green space distribution in response to the urban land cover process ([Wang et al. 2019c](#); [Zhou and Wang 2011](#)). However, the urban land use process also proved to be related to urban green landscapes ([Woldesemayat and Genovese 2021](#)), which are still poorly understood. In addition, many of the related studies used municipal cadastral data or urban land use data from the urban planning bureau to understand the urban land use dynamics ([Woldesemayat and Genovese 2021](#)), which might introduce large errors due to the low quality and uncertainty from the urban land use data. Understanding the impact of urban land use changes on urban green space distribution can help guide green space planning and management from a more socio-economic perspective. It is thus critical to look at how urban green space distribution changes in response to urban land use changes (e.g., from residential land to business land). To achieve this, two issues need to be addressed. (1) To improve the existing urban land use maps to understand the urban land use dynamics. (2) To investigate the impact of urban land use changes on urban green space distribution by using certain methods.

Previous studies have demonstrated that the inner-urban functions cannot be recognized by using remote sensing (RS) data only due to the lack of socio-economic

information. The emergence of geospatial big data (GBD) ([Li et al. 2016](#)) such as mobile phone positioning data ([Ratti et al. 2016](#); [Wu et al. 2020](#)), social media data ([Ilieva and McPhearson 2018](#); [Yammine et al. 2018](#); [Ye et al. 2016](#)), traffic trajectory data ([Niu et al. 2017](#)), and geotagged photographs ([Cadavid Restrepo et al. 2017](#); [Krylov et al. 2018](#); [Srivastava et al. 2018](#)) provide the new opportunity for classifying urban land use by capturing socio-economic characteristics ([Liu et al. 2015](#)). Thus, the integration of RS data and GBD could improve the existing urban land use maps by providing both physical and socio-economic information ([Dong et al. 2019](#); [Sarmin and Ismail 2016](#); [Yin et al. 2021a](#)). Promising progress has been made in the applications of integrated RS and GBD on urban land use mapping at different scales and regions ([Shi et al. 2019](#); [Zhang et al. 2020c](#); [Zhang et al. 2019](#)). Despite the significant potential of combining RS and GBD to provide improved insights into urban land use, the discrepancies in spatial data quality (e.g., semantic, timestamp, and scale), technological format, and data structure make it difficult to combine them ([Cao et al. 2020](#); [Yin et al. 2021b](#)).

Several approaches have been utilized to investigate the patterns of urban green spaces at different regions and scales ([Chan and Vu 2017](#); [Gavrilidis et al. 2019](#); [Yang et al. 2018](#)). These efforts can be categorized into statistical analysis, landscape analysis, and geospatial analysis. Among them, the statistical analysis could reveal the spatial patterns of urban green space at the landscape level ([Feng and Astell-Burt 2018](#);

[Wüstemann et al. 2017](#)), which has the potential to understand the distribution of urban green space. Meanwhile, the landscape analysis could relate the landscape patterns of urban green space to urbanization and determine their influence on the ecological attributes of the urban environment ([Daz et al. 2020](#); [Grafius et al. 2018](#)). Moreover, geospatial analysis such as hotspot analysis could reveal the spatial heterogeneity of urban green space, and further, elucidate the effect of urban development on urban green space ([Liu et al. 2020c](#)). Few studies, on the other hand, have sought to combine these methodologies to investigate the spatial-temporal patterns of urban green space and its response to urban land use changes.

The urban green space planning and management of most undeveloped regions in China are still at the initial stage and have ample space for further improvement. While large and medium-sized cities, such as Beijing, Shanghai, and Hangzhou, have an advantage in terms of governmental support, economic investment, and practical research for green space planning and management. The understanding of urban green space distributions in response to urban land use changes in these cities is valuable for guiding small cities to improve their urban green space patterns and system. Therefore, this Ph.D. research takes Hangzhou as the study site, and aims to improve the existing urban land use maps by integrating RS and GBD and assess the urban green space distributions in response to urban land use changes.

1.2 Novel aspects of the research

This work for the first time reviewed the existing efforts on urban land use mapping by integrating RS and GBD and categorized the integration methods into decision-level integration (DI) and feature-level integration (FI).

This work for the first time investigated the spatial and temporal characteristics of urban green space within different urban land use changes (e.g., from open space to residence) and evaluated the impact of urban land use changes on green space in Hangzhou city.

1.3 Aims and objectives

This study aims to improve the urban land use maps by integrating RS and GBD and then utilize the obtained maps to investigate the impact of urban land use changes on urban green space distributions in Hangzhou city. In this study, four individual objectives have been defined:

(1) To review the existing literature on the nature of RS and GBD, as well as their integration strategies in urban land use classification.

(2) To propose a methodology framework based on the integration strategy and evaluate the framework in Hangzhou city in 2019.

(3) To classify urban land use maps in Hangzhou for 2017 and 2021, and understand the urban land use dynamics from 2017 to 2021.

(4) To investigate the impact of urban land use changes on urban green space variations in Hangzhou from 2017 to 2021.

1.4 Definitions

When reading this thesis, it is critical to be familiar with how I defined a few terms and concepts in the context of this Ph.D. study.

Urban land use refers to the land in urban areas which used for several purposes, e.g. sports centers, airports, hospitals, and houses. The distinction between urban land cover and urban land use is significant because land cover relates to the physical aspect within urban areas (e.g. forests, grasslands, water bodies, and bare land), while land use relates to human activities.

Urban land use mapping is a fundamental method for recognizing and locating urban land uses, such as industrial, residential, institutional, and commercial zones. Methods for updating urban land use maps depend on supervised classification and unsupervised classification.

Geospatial big data (GBD) is generated every day mostly by fixed and mobile sensors such as environmental sensors, cameras, webcams, social media, or even residents' daily activities. These data such as mobile phone data, traffic trajectories, geo-tagged photos, Points of interest (POIs), and social media data provide an alternative approach to uncover how cities function.

Remote sensing (RS) is the process of detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted radiation at a distance (typically from satellite or aircraft). RS images can be used to extract urban land use information.

Urban green space is defined as all urban land covered by vegetation of any kind, e.g. community gardens, parks, residential green spaces, and greenways. This covers vegetation on private and public grounds, irrespective of size and function, and can also include small water bodies such as ponds, lakes, or streams.

1.5 Thesis structure

This thesis is divided into eight chapters:

Chapter 1 describes the background of the study and introduces the research aims, objectives, and main definitions in this thesis.

Chapter 2 summarizes the relevant literature, focusing on the urban land use mapping as well as the impact of urban land use on urban green space distribution.

Chapter 3 introduces the methods mainly used in this thesis.

Chapter 4 reviews the existing literature and focuses on the state-of-the-art perspective of the urban land use categorization by integrating RS and GBD. The most often utilized RS and GBD features are identified and the integration strategies were categorized.

Chapter 5 proposes a methodology framework based on the integration strategies proposed in **Chapter 4** and tests the methodology framework in Hangzhou city, China. The advantages and disadvantages of the framework are discussed via bibliographic evidence and quantitative analysis.

Chapter 6 improves the methodology framework proposed in **Chapter 5** and classifies the urban land use in 2017 and 2021 in Hangzhou city. In addition, the physical process of urban land use change from 2017 to 2021, and the underlying socioeconomic process are analyzed.

In Chapter 7, this research investigates the impacts of urban land use changes (e.g., the transformation from open space to residence) on urban green space distribution by integrating statistical analysis, landscape analysis, and geospatial analysis methods.

Chapter 8 synthesizes the results obtained from this thesis with each of the four research papers. An outlook and recommendations for future research are also presented.

Chapter 2: Literature review

2.1 Introduction

This chapter will provide a general overview of urban land use, urban land use changes, accuracy assessment, urban green space, and the impact of urban land use changes on urban green space distribution.

2.2 Urban land use

2.2.1 Evolution from urban land cover to urban land use

Urban land cover and land use play an important role in modern urban planning and management ([Ratnasari et al. 2017](#); [Song et al. 2016](#); [Taubenböck et al. 2012](#); [Zhang and Xu 2018](#)). The distinction between urban land cover and urban land use is significant because land cover relates to the physical aspect within urban areas, while land use relates to human activities ([Cao et al. 2018](#); [Xing et al. 2018](#)). To be more specific, the urban land cover represents spatial information on different types (classes) of natural coverage of the urban environment, e.g. forests, grasslands, impervious surfaces, water bodies, and bare land ([MacLachlan et al. 2017](#); [Myint et al. 2011](#); [Schneider et al. 2001](#)). While urban land use refers to the spatial information on different urban land use types of the socio-economic functions ([Zhang et al. 2017a](#)).

Urban land cover mapping using RS data has a long history ([Howarth and Boasson 1983](#); [Patino and Duque 2013](#); [Reba and Seto 2020](#); [Yang et al. 2003a](#)). With the substantial progress of RS techniques, the spatial resolution of urban land cover maps has gradually improved from coarse spatial resolutions (e.g., MODIS, AVHRR) to moderate and high resolutions (e.g., SPOT, Landsat) ([Friedl et al., 2002](#); [Schneider et al., 2010](#)[Deng et al. 2019](#)). The time period of these urban land cover products has transformed from a single period to repeated observations, which could provide time series urban land cover information ([Gong et al. 2019](#); [Li and Chen 2018](#); [Momeni et al. 2016](#); [Reda and Kedzierski 2020](#)). Overall, these urban RS studies have provided opportunities for a better understanding of physical urban attributes (e.g., impervious surface, built-up areas, artificial surfaces, and urban extent) ([Defries and Townshend 2007](#); [Li et al. 2020b](#); [Schneider et al. 2009](#); [Zhu et al. 2019](#)).

However, more specific information on inner-urban structures (e.g., urban land use) cannot be retrieved by using RS only ([Li et al. 2020a](#); [Liu et al. 2015](#)). This is likely due to the fact that RS data could only offer physical aspects of the urban environment, such as texture, spectra, and contexture ([Wetherley et al. 2017](#); [Wu and Murray 2003](#)). While these physical characteristics may not be able to map urban land use due to the great correlation of texture, contexture, and spectra features among urban land use categories ([Hagenauer and Helbich 2012](#); [Jepsen and Levin 2013](#); [Lou et al. 2019](#)). Recently, the demands for urban land use products with socio-economic characteristics

have increased, emphasizing a transformation from urban land cover (e.g., physical aspects) to urban land use (e.g., socio-economic aspects). However, it is difficult to map urban land use – the human activities on land, particularly in modern cities where buildings often host multiple purposes and can be renovated and repurposed ([Andrade et al. 2020](#); [Ye et al. 2020](#)).

Urban land use has widespread effects on urban heat islands, biodiversity, air condition, and public health ([Li et al. 2018a](#); [Meng et al. 2018](#); [Van Zanten et al. 2016](#)). With increasing urbanization and population growth, more than half of the global population will live in cities by 2050 ([Seto and Shepherd 2009](#)). The accelerated population from rural areas to urban areas has led to rapidly changing urban land use patterns ([Chen et al. 2021](#)). Several efforts have been made for identifying urban land use structures ([Huang et al. 2018a](#); [Shi et al. 2019](#); [Su et al. 2020](#); [Zhang et al. 2018](#)). For example, [Chang et al. 2015](#) chose the residential type and divided it into three groups based on social media data and home transaction records: gated communities, regular communities, and urban slums. Furthermore, [Frias-Martinez and Frias-Martinez \(2014\)](#) presented a method for automatically determining land uses (business, leisure, nightlife, and industry) in metropolitan locations using unsupervised learning by grouping geographical zones with comparable tweeting activity patterns. These urban land use maps only focused on certain types of urban land use rather than a complete classification system. With the development of classification techniques (e.g.,

machine learning, deep learning), urban land use maps with more types of urban land use were generated. [Xu et al. \(2020\)](#) suggested a paradigm for functional recognition that employs a deep multi-scale neural network (e.g., industrial, residence, administration, green space, transportation, commercial). [Gong et al. \(2020\)](#) published a new urban land use map for the whole Chinese mainland that incorporates 10-m satellite photos, OpenStreetMap, nighttime lights, POI, and Tencent social big data as input characteristics, and employed a two-level categorization algorithm. It can be noted that the classification system of urban land use is gradually diversifying with different mapping targets. Furthermore, the classification techniques and data sources used for the urban land use categorization have been developed.

2.2.2 Integrating RS and GBD on urban land use mapping

The emergence of GBD (e.g., social media data, POIs, OSM data, mobile phone data) provides a new opportunity for urban land use classification due to the abundant anthropogenic information, which also compensates for the lack of socio-economic attributes of RS data ([Batty 2013](#); [Campbell 2009](#); [Fischer 2012](#); [Harvey 2013](#); [Huang and Wong 2016](#); [Liu et al. 2016](#); [Liu et al. 2018b](#); [Munoz et al. 2016](#); [See et al. 2016](#); [Wu et al. 2015](#); [Zong et al. 2020](#)). The integration of space-based RS data and time-based GBD can improve the existing urban land use maps by providing both physical and socio-economic information ([Dong et al. 2019](#); [Sarmin and Ismail 2016](#)). Promising progress has been made in the applications of integrated RS and GBD on

urban land use mapping at different scales and regions ([Goffi et al. 2020](#); [Grippa et al. 2018](#); [Shi et al. 2019](#); [Taubenböck et al. 2009](#); [Zhang et al. 2020c](#); [Zhang et al. 2019](#)). For example, [Liu et al. \(2017\)](#) used six variables to determine the urban land use type inside each land use parcel (spatial, texture, spatial envelope, and rotation-invariant from RS photos, Tencent real-time user density, and POIs from social media data). [Jia et al. \(2018\)](#) categorized RS data and cell phone location data independently before fusing the two findings using a decision fusion approach for urban land use classification.

Here this research reviewed the literature on the nature of RS and GBD and their integration strategies to address urban land use mapping, aiming to identify the opportunities and challenges that synthesizing RS and GBD offers for urban land use studies. The Web of Science database (Science Citation Index-Expanded) was used to search for articles integrating RS and GBD for urban land use mapping published before December 31, 2021. Only articles and conference papers published in English were considered. Specifically, the search extracted articles with “Term 1 AND Term 2 AND Term 3” in the title, keywords, or abstracts (Table 2-1). Finally, 38 primary research publications were included (Table 2-2).

Table 2-1 Key terms for urban land use

Term1	Term2	Term3
RS, remotely sensed imagery, VHR imageries, satellite imagery	Mobile phone, traffic, social media, geotagged photos, maps, search engines, smart card	Urban land use mapping, urban land use recognition, urban land use classification, urban functional zone mapping, urban functional region mapping, urban scene recognition, urban functional pattern, urban land use categories

Table 2-2 Integration publications for urban land use classification

ID	Year	Title	Study area	Spatial unit	RS	GBD	Integration process	Land use types
1	2016	Mapping urban land use by using Landsat images and open social data	Beijing	Parcel-level	Landsat OLI	OpenStreetMap, SINA POIs	Similarity index (normalized feature distance)	Water, cropland, orchard, forest, grassland, shrubland, undeveloped, cottage, community, retail, service, industrial, medical, educational, administrative, public
2	2017	The combined use of RS and social sensing data in fine-grained urban land use mapping: A case study in Beijing, China	Haidian District, Beijing, China	Parcel-level	Landscape Metrics, Spectral and Texture Attributes from Gaofen image	Density and Spatial Pattern of Baidu POIs, Density of Geotagged Weibo Posts	Random forests (RF) model	Open space, institutional, residential, business

3	2017	Classifying urban land use by integrating RS and social media data	Haizhu District, Guangzhou	Parcel-level	Worldview-2 image	OpenStreetMap, Gaode POIs, real-time Tencent user density (RTUD)	Support vector machine (SVM) classifier	Public management-service land, industrial land, green land, commercial land, residential land, parkland, urban village
4	2018	Integrating Aerial and Street View Images for Urban Land Use Classification	New York City	Pixel-level	Aerial images (Bing map)	Google Street View images	Deep neural network	Family building, walk-up building, mixed residential and commercial building, industrial and manufacturing, transportation, public facilities, open space, parking, vacant land, unknown
5	2019	Model Fusion for Building Type Classification from Aerial and Street View Images	49 US states	Building-level	Aerial images (Bing map)	OpenStreetMap, Google Street View images	Two-stream end-to-end fusion network (i.e., a geometric-	Commercial, industrial, public, residential

							level model fusion), and decision-level model fusion	
6	2019	Urban Land Use and Land Cover Classification Using Multisource RS Images and Social Media Data	Guangzhou	Object-level	ZY-3 high-resolution image, Landsat 8 OLI multispectral image, Sentinel-1A SAR image	WeChat user density data	RF classification	Water, urban village, road, residential building, industrial building, greenhouse, vegetation, educational, commercial, bare land

7	2019	Functional urban land use recognition integrating multi-source geospatial data and cross-correlations	Shenzhen	Grid-level	SPOT-5 images	Baidu POIs, real-time Tencent users	RF classification	Residential land, industrial land, commercial land, public management, and service land, green and forest land, waterbody
8	2020	DFCNN-Based Semantic Recognition of Urban Functional Zones by Integrating RS Data and POI Data	Shenzhen	Object-level	worldview-3 satellite	OpenStreetMap, Amap POI	Deeper-Feature Convolutional Neural Network (DFCNN),	Urban green, industrial districts, public services, commercial districts, hospitals, schools, shanty towns
9	2020	Deep learning-based remote and social sensing data fusion for urban region function recognition	N/A	Pixel-level	URFC-A and URFC-B	user visit data	End-to-end deep learning	Residential, school, industrial park, railway station, airport, park, shopping area, administrative district, hospital

10	2020	Mapping the Essential Urban Land Use in Changchun by Applying Random Forest and Multi-Source Geospatial Data	Changchun	Parcel-level	LuoJia-1 satellite, Sentinel-2 satellites	OpenStreetMap, Baidu Street View Images, Baidu POI	RF classification	Residential, commercial, industrial, public, green space
11	2020	Exploring Impact of Spatial Unit on Urban Land Use Mapping with Multisource Data	Futian District, Shenzhen	block, grids, objects	GaoFen2 images	OpenStreetMap road network data, Gaode POI	RF classification	Residential, commercial, public facilities, educational, urban village, natural land
12	2020	Large-scale urban functional zone mapping by integrating RS images and open social data	Beijing	objects	ZiYuan-3 (ZY-3) image	Gaode POI data	SVM	Commercial, residential, institutional, industrial, shantytown, open space

13	2020	Mapping essential urban land use categories in China (EULUC-China): preliminary results for 2018	27 cities in China	Parcel-level	10-m Sentinel-2A/B images, 130-m LuoJia-1 nighttime light images	Tencent mobile phone locating-request (MPL) data, Gaode POI data	RF classification	Residential, business office, commercial, industrial, road, transportation, airport, administrative, educational, medical, sports, park
14	2020	An Ensemble Learning Approach for Urban Land Use Mapping Based on RS Imagery and Social Sensing Data	Beijing	street block	Google Earth images, Tencent street-view images	Baidu POIs, Weibo social media check-ins	RF and Xgboost	Commercial, residential, educational, natural, civic, transport, industrial, agricultural, other
15	2020	Comparison of Machine-Learning Methods for Urban	Hangzhou	Parcel-level	Sentinel-2A satellite	Gaode POIs data	RF, SVM, ANN	Residential, business office, commercial, industrial, road, transportation, airport,

		Land-Use Mapping in Hangzhou City, China			images, land surface temperature (LST), LuoJia-1 nighttime lights (NTLs), Building Height			administrative, educational, medical, sports, park
16	2020	Sampling Strategy for Detailed Urban Land Use Classification: A Systematic Analysis in Shenzhen	Shenzhen	Parcel-level	Sentinel 2A/B, LuoJia 1	OpenStreetMap, Gaode POI, Tencent social big data	RF classification	Residential, business office, commercial, industrial, road, transportation, airport, administrative, educational, medical, sports, park
17	2020	Mapping Essential Urban Land Use Categories in Nanjing	Nanjing	Parcel-level	10-m Sentinel-2A/B	Tencent mobile phone locating-	RF classification	Residential, business office, commercial, industrial, road, transportation, airport,

		by Integrating Multi-Source Big Data			images,130-m LuoJia-1 nighttime light images, building footprint	request (MPL) data, Gaode POI data		administrative, educational, medical, sports, park
18	2020	Regional Mapping of Essential Urban Land Use Categories in China: A Segmentation-Based Approach	Ningbo	Parcel-level	10-m Sentinel-2A/B images,130-m LuoJia-1 nighttime light images, WorldPop	Gaode POI data	RF classification	Residential, business office, commercial, industrial, road, transportation, airport, administrative, educational, medical, sports, park

19	2020	Heuristic sample learning for complex urban scenes: Application to urban functional-zone mapping with VHR images and POI data	Beijing	Parcel-level	WorldView-II image	POIs	Deep forest	Commercial, residential, industrial, urban villages, campuses, park
20	2020	Detailed Mapping of Urban Land Use Based on Multi-Source Data: A Case Study of Lanzhou	Lanzhou	Parcel-level	Sentinel 12A, LuoJia-1 nighttime light images	OpenStreetMap and Gaode road networks, Gaode POI data, Tencent Easygo data	RF classification	Residential, business office, commercial, industrial, road, transportation, airport, administrative, educational, medical, sports, park
21	2015	An integrative method for mapping urban land use change using "geo-sensor" data (conference paper)	Kunming, Yunnan	Pixel-level, parcel-level	Landsat TM, ETM, ASTER GDEM	Tencent map POI (Point of interest), SINA Weibo geotagged social media	Overlay (OSM for built-up parcel generation and POIs for parcel categorization)	Water, commercial, office, recreational, other, manufacturing, transportation, residential (gated communities, ordinary communities, urban slums)

						data, OpenStreetMap road network, Fang.com (Real estate communities, House sale records)		
22	2017	Hierarchical semantic cognition for urban functional zones with VHR satellite images and POI data	Beijing	Parcel-level	QuickBird image	POIs	Hierarchical semantic cognition (HSC)	Commercial, residential, campuses, parks and greenbelt, industrial zones, shanty towns
23	2018	Social functional mapping of urban green space using RS and social sensing data	Beijing	Parcel-level (POI level)	GaoFen-2 remotely sensed imagery	POIs, OpenStreetMap road network	Near-convex-hull analysis (NCHA) and text-concave-	municipal park, theme park, community park, roadside green space, residential green space, other subsidiaries green space

							hull analysis (TCHA)	
24	2018	Mapping Urban Land Use at Street Block Level Using OpenStreetMap, RS Data, and Spatial Metrics	Western Africa	Street Block Level	WorldView-3 land-cover (LC) maps	OpenStreetMap	Supervised random forest classifier	Vegetation, bare soil, nonresidential built-up, planned residential, unplanned residential
25	2018	Urban Land Use Mapping by Combining RS Imagery and Mobile Phone Positioning Data	Beijing	Pixel-level	Gaofen 1 (GF1) wide-field view (WFV) scene	mobile phone positioning data (MPPD)	Decision fusion strategy	business, open, residential, entertainment, other
26	2018	Mapping urban functional zones by integrating very high spatial resolution RS imagery and points of	Xiamen	Object-based (parcel based)	GaoFen-2 RS Imagery	Baidu POIs	The Weights of POIs and Segmentation	Residence, transportation zones, convenience shops, shopping centers, factories, companies, public services

		interest: A case study of Xiamen, China						
27	2018	Portraying Urban Functional Zones by Coupling RS Imagery and Human Sensing Data	Shenzhen	Object-based	SPOT-5 images	Mobile phone positioning data	Hierarchical clustering	urban center, sub-center, suburbs, transit region, urban buffer, ecological area
28	2019	Exploring semantic elements for urban scene recognition: Deep integration of high-resolution imagery and OpenStreetMap (OSM)	Beijing	object-based	Worldview-2	OpenStreetMap data, Baidu POIs	POI label	commercial, entertainment, public service, educational, residential
29	2020	Social Sensing for Urban Land Use Identification	New York City	Object-based	Sentinel-2A remote sensed imagery	Bike data and taxi data, OpenStreetMap	Decision tree and random forest	Residential, entertainment, office, open space, industrial, water

30	2020	Recognizing urban functional zones by a hierarchical fusion method considering landscape features and human activities	Wuhan	Parcel-level	Landsat-8 OLI images,	POI data, taxi GPS trajectory data, OpenStreetMap	Fusion-strategy based recognition	Administrative and public, commercial, industrial, transportation, residential, green space, and square
31	2020	A New RS Images and Point-of-Interest Fused (RPF) Model for Sensing Urban Functional Regions	N/A	Pixel-based	Google earth satellite imagery	Gaode POIs data	RS images and POI Fused (RPF) module.	Green space and square, Industrial area, Administration and public service, Commercial, and business facility, Transportation zone, Residence
32	2020	Land use classification from social media data and satellite imagery	Wuhan	Parcel-level	Google images	POI data, road networks from the OpenStreetMap	Hierarchical Determination method based on Kernel Density Classification (KDC-HDM)	Farmland, undeveloped land, lake, and green land, residential, industrial, commercial, public service, transportation, park

33	2020	Open-source data-driven urban land-use mapping integrating point-line-polygon semantic objects: A case study of Chinese cities	Six areas of cities in China	Object-based	Google Earth images	OpenStreetMap data, POIs	The rule-based category mapping (RCM) model	residential, administration, cultural facilities, education and science, sports, health care, commercial, industrial, logistics and warehouses roads, transportation, municipal, park, green buffer, square, specially designated land, water, farmland, construction land
34	2021	Mapping essential urban land use categories with open big data: Results for five metropolitan areas in the United States of America	the United States	Object-based	VHR, Sentinel, nighttime light data	Twitters	Automatic ensemble learning	Residential, entertainment, transportation, industrial, office
35	2021	Mapping Essential Urban Land Use Categories in Beijing	Beijing	Parcel-based	Sentinel-2, Luojia-1	Baidu POIs, Tencent easygo	RF model	Residential, business, commercial, industrial, administrative, medical,

		with a Fast Area of Interest (AOI)-Based Method			nighttime data	crowdedness data		cultural, greenspace, educational, village
36	2021	Identifying Urban Building Function by Integrating Remote Sensing Imagery and POI Data	Wuhan	Building-level	VHR images	POIs	KDE, spatial similarity	Residential, administration, commercial, municipal utilities, warehouse, transportation, industrial
37	2021	Uncovering the Nature of Urban Land Use Composition Using Multi-Source Open Big Data with Ensemble Learning	Ningbo	Parcel-based	Multi-source RS data	POIs, OSM	Simple non-iterative clustering (SNIC) algorithm	Residential, village, business, commercial, industrial, transportation, administrative, educational, medical, sport, park, undeveloped
38	2021	Decision-Level and Feature-Level Integration of Remote Sensing and geospatial Big Data	Hangzhou	Parcel-based	Sentinel-2 images	Gaode POIs, OSM road network data	RF model	Residence, business, institution, and open space

		for Urban Land Use Mapping							
--	--	-------------------------------	--	--	--	--	--	--	--

2.2.3 Urban land use change

Urban land cover/use change is one of the most important factors profoundly affecting the earth's ecological systems ([Angel et al. 2011](#); [Johnson et al. 2017](#); [Rogan and Chen 2004](#); [Seto and Fragkias 2005](#); [Xing et al. 2017](#); [Yan et al. 2015](#); [Yang and Lo 2010](#); [Zha et al. 2003](#)). Previous researchers have studied urban land cover change for several years ([Chen et al. 2020a](#); [Esch et al. 2012](#); [Goldewijk 2001](#); [Meng et al. 2017](#); [Muller et al. 2010](#); [Zhang and Seto 2011](#)). Among these studies, urban areas (e.g., impervious surfaces, urban land, urban boundary) have been extracted from regional scales to global scales ranging from high resolution to coarse resolution RS images ([Li et al. 2020c](#); [Wu et al. 2017](#); [Yang et al. 2003b](#)). However, few studies have investigated the spatial-temporal patterns of urban land use. Existing efforts on urban land use changes and related analysis are referring to the urban planning maps. For example, [Wang et al. \(2019b\)](#) used a series of mathematical methods and parcel-level urban land use maps from Xi'an municipal planning bureau for a quantification study in the Xi'an city wall area. [Li et al. \(2015\)](#) investigated the spatial variations of green space among different land use categories from the municipal cadastral data within the city of Shanghai at the city, inner-outer ring road, and district scales. While the urban land use data used in these studies have a very low resolution or inaccurate units ([Deng et al. 2008](#)).

Urban land use change will affect both natural (e.g., earthquakes, floods) and human systems (e.g., the changes in urban land use types) in the urban environment ([Eckert 2014](#); [Huang and Wang 2019](#); [Seto and Kaufmann 2003](#)). In particular, it has a widespread effect on urban green space in terms of distribution, composition, configuration, and function ([Hao et al. 2015](#)). Urban land use change can be influenced

by several aspects including policy, economic and social aspects ([Zhang et al. 2021](#); [Zheng et al. 2018](#); [Zhong et al. 2014](#)). Therefore, the spatial-temporal analysis of urban land use changes will, in turn, provide a powerful reference for urban management and policymaking, as well as urban green space planning. It is thus necessary to analyze the spatial-temporal patterns of urban land use changes to understand the urban ecological system and anthropogenic activities ([Qian et al. 2020](#)).

2.2.4 Accuracy assessment

Urban land use greatly influences human activities as well as human health, wealth, and well-being. It is thus necessary to produce high-quality and up-to-date urban land use information. The accuracy assessment refers to the comparison of the predicted map with the actual map that is assumed to be true, which is an unavoidable step of urban land use mapping ([Donner et al. 2000](#); [Olofsson et al. 2013](#)). Different sampling designs and assessment methods during the accuracy assessment process could be used to produce different urban land use maps ([Olofsson et al. 2014](#)). Previous studies have been promoted to improve evaluation methods to enhance their usefulness and help produce high-quality urban land use products ([Foody 2004](#); [McKenzie et al. 1996](#)).

The spatial units used for urban land use mapping can be categorized into pixel-based, grid-based, or object-based ([Stehman and Wickham 2011](#)). Among them, the spatial units of most urban land use maps are pixel-based. While recent studies classify urban land use types based on parcel-level units because of the heterogeneity of urban functions among different land use parcels. It is thus essential to overcome the problem that how to evaluate parcel-based urban land use maps. According to [Stehman and Foody \(2019\)](#)'s work, previous evaluation methods mainly depend on a confusion matrix (e.g., overall, user's, and producer's accuracy), which is easy to implement. In

the confusion matrix, the proportional areal representation of the study site needs to be reflected. And the variability of the accuracy and data estimates should be quantified by standard errors or confidence intervals.

In this research, it might not be appropriate to use the area proportion to estimate the accuracy of parcel-based urban land use maps. The reason is that the size of urban parcels varies, and the larger urban parcel cannot stand for a larger proportion of urban land use. Therefore, the area proportion of urban land use parcels should not be utilized for the estimation. To solve this problem, the quantity proportion of urban parcels can be used to evaluate the confusion matrix instead of the area proportion. To be more specific, the area in category *i* of the contingency table is replaced by the proportion of the number of category *i* to the number of all categories. This research evaluated the urban land use classification results by computing the overall accuracy (OA), user's accuracy (UA), and producer's accuracy (PA) based on the confusion matrix ([Foody 2002](#)). Also, approximate 95% confidence intervals of OA, UA, and PA were calculated based on the equations proposed by [Card \(1982\)](#). The map category is in the column and the true category is in the row referring to the contingency table (Table 2-3).

Table 2-3 Contingency table for accuracy assessment (referring to [Card \(1982\)](#))

		Map category(j)						Total
		1	2	.	.	.	r	
True category(i)	1	n11	n12				n1r	n1.
	2	n21	n22				n2r	n2.
	.							
	.							
	.							
	r	nr1	nr2				nrr	nr.
Total	n.1	n.2				n.r	n	

2.3 Impact analysis of urban land use

Urbanization significantly affects the urban environment, especially urban green space ([Byomkesh et al. 2011](#); [Hernández-Moreno and Reyes-Paecke 2018](#)). Urban green space can help improve microclimate, mitigate urban heat islands ([Xiao et al. 2018](#)), control air pollution ([Heidt and Neef 2008](#)), maintain human welfare, and provide entertainment opportunities ([Li et al. 2019a](#); [Villeneuve et al. 2012](#)). It is thus important to understand urban green space variations influenced by rapid urbanization. Several studies have investigated the spatial-temporal dynamics of urban green space to help optimize green space patterns and develop related policies ([Zhao et al. 2013](#)). For example, [Liu et al. \(2021\)](#) employed integrated approaches to characterize the changing patterns and intensities of urban green space influenced by urbanization in Shanghai from 1990 to 2015. [Wang et al. \(2020\)](#) investigated the importance of a multi-scale perspective in understanding the spatial distribution of urban green space and its change, and its response to urban development. [Jiao et al. \(2017\)](#) used the improved gradient partitioning method and six landscape metrics to characterize urban expansion and green space fragmentation and investigate the relationship between urban green fragmentation and urban expansion using correlation analysis and regression modeling. [Li et al. \(2015\)](#) investigated the spatial variations of green space among different land use categories within the city of Shanghai at the city, inner-outer ring road, and district scales. These efforts provide an improved understanding of the spatial-temporal variations of urban green space in response to urban land cover changes (e.g., physical aspects of urbanization). However, few studies have examined the effects of urban land

use changes (e.g., socio-economic aspects of urbanization) on urban green space distribution.

2.3.1 Urban green space

Urban green spaces refer to urban areas that are covered by greenness, which can not only provide ecological and environmental benefits but also physical and psychological benefits for urban dwellers ([Shahtahmassebi et al. 2021](#)). Urban areas generally have fewer green spaces in comparison to villages or suburbs, resulting in weaker ecosystem services in the urban environment ([Reyes-Riveros et al. 2021](#)). Therefore, urban areas should improve the urban green space patterns to optimize the value of urban green spaces ([Semeraro et al. 2021](#)).

Urban green space is diverse, varying in size, vegetation cover, species richness, environmental quality, proximity to public transport, facilities, and services ([Daniels et al. 2018](#)). The World Health Organization (WHO) defined urban green space as "all urban land covered by vegetation of any kind" ([Wu et al. 2019](#)). It mainly includes two aspects: public green space and private green space (Figure 2-1). Public urban green space includes parks and reserves, sporting fields, riparian areas like the stream and river banks, greenways and trails, community gardens, street trees, and nature conservation areas, as well as less conventional spaces such as green walls, green alleyways, and cemeteries ([Ludwig et al. 2021](#)). Private green space includes private backyards, communal grounds of apartment buildings, and corporate campuses ([Hussainzad et al. 2021](#)). Furthermore, some researchers utilize "urban open space" (e.g., a natural and cultural resource, "unused land" or "park and recreation areas) to describe a broader range of urban green spaces.



Figure 2-1 Urban green space samples.

Urban green space provides many functions and benefits that contribute to the quality of residents' daily life (Figure 2-2) ([Lee and Maheswaran 2011](#)). These benefits can be categorized as follows:

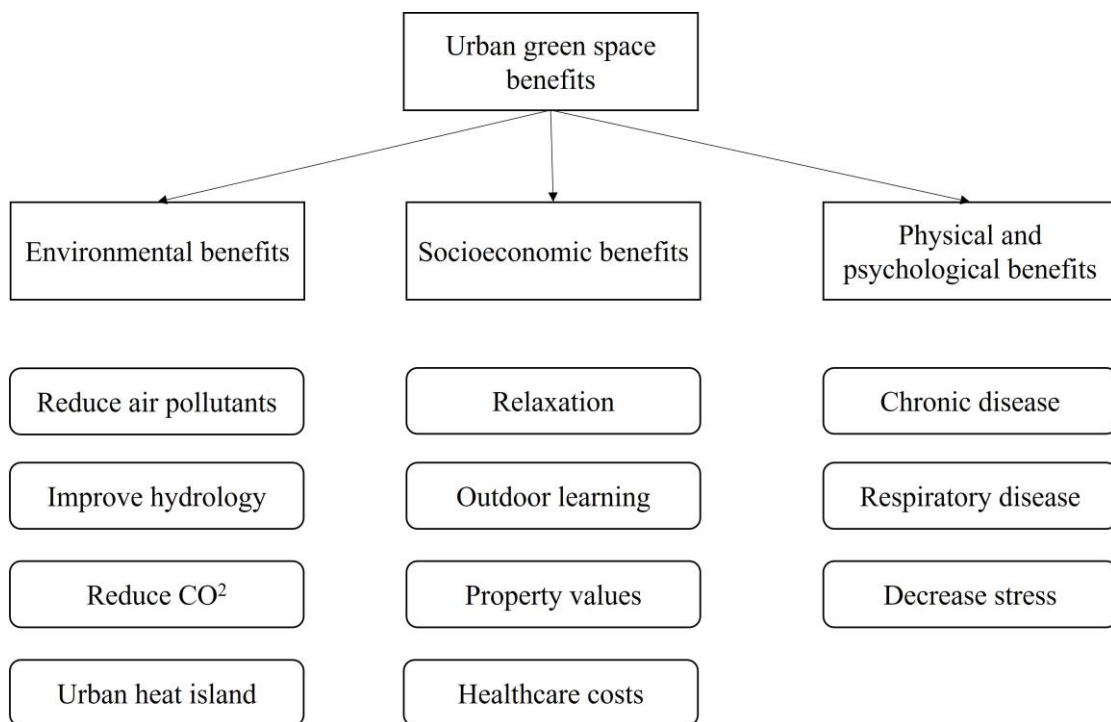


Figure 2-2 Urban green space benefits.

(1) Environmental benefits

Urban green space provides many direct environmental benefits ([Wang 2009](#)). For example, it helps improve air conditions by preventing the distribution of air pollutants (e.g., O₃, PM_{2.5}, NO₂, SO₂, and CO) or by reducing the air pollutants from traveling to other places ([Eisenman et al. 2019](#); [Matos et al. 2019](#)). Urban green space can also help to improve urban hydrology by increasing rainwater infiltration, intercepting rainfall, and increasing the water storage capacity ([Yang et al. 2019](#)). In addition, it can reduce surface runoff by preserving surface water. Urban green space can help reduce the increased temperatures caused by urban heat islands ([Schipperijn et al. 2013](#)). Specifically, shadings and evapotranspiration produced by urban green spaces can contribute to the lower temperature ([Bao et al. 2016](#); [Masoudi and Tan 2019](#)). Urban green space also contributes to the reduction of CO₂ by direct sequestration ([Amoatey et al. 2018](#)).

(2) Socio-economic benefits

Urban green space can provide social benefits such as relaxation or recreation for urban residents. Research shows that people are more likely to engage in physical activity when the urban green spaces nearby are accessible. Furthermore, urban green space can serve as an outdoor educational place for students ([Flouri et al. 2019](#)). The ingenuity and imagination of students have been promoted by exposure to the urban green space.

Urban green space can not only avoid the costs of the establishment of more rainwater retention basins but also reduce healthcare costs due to a reduction in air pollution ([Song et al. 2018b](#)). Besides, the accessibility of urban green space in most urban areas has been recognized as an important variable for property values ([Liebelt et al. 2018](#)). Urban areas close to urban green spaces are aesthetically pleasing and

attractive to urban dwellers. In addition, urban green space plays an important role in cooling buildings.

(3) Physical and psychological benefits

Urban green spaces have a positive influence on people's physical health. ([Kothencz et al. 2017](#); [Teixeira 2021](#)). According to previous studies, urban green spaces can help reduce cardiovascular disease symptoms and improve respiratory health ([de Jalón et al. 2021](#)). Access to urban green space not only reduces the risk of chronic diseases but also avoids the risk of obesity. Increasing access to green space and increased levels of physical exercise have been linked. When access to an urban park was enhanced, the percentage of inactive and moderately active people who used it rose.

In addition, urban green spaces can provide psychological benefits for urban residents ([Wang et al. 2019a](#)). In today's fast-paced culture, mental illness becomes a huge source of concern when rest is devalued. Every year, eight million people die as a result of mental illness throughout the world. Limited access to urban green space in metropolitan settings, as well as the low quality of green space that is accessible, may lead to poor mental health outcomes ([Callaghan et al. 2021](#)). Individual life distance from a green space or park, as well as the fraction of land designated as open space/parks, are negatively connected to communal anxiety/mood disorder treatment numbers ([Gascon et al. 2018](#)). Even when physical activity rates do not rise as a result of increased access to green space, it has been proven that having more green space reduces stress and improves social cohesiveness ([Jennings and Bamkole 2019](#)).

2.3.2 Impact of urban land use on urban green space

Studies have shown that the impacts of urbanization on urban green space are mainly manifested in the distributions, compositions, and configurations ([Ricci et al. 2022](#); [Woldesemayat and Genovese 2021](#)). Distribution presents the patterns and structures of urban green space patches, such as the fragmentation, connectivity, and equity of the patches. Configuration refers to the shape, size, and structure of urban green spaces ([Coban et al. 2021](#)), while composition represents different urban green space types (e.g., grass, trees, shrubs, flowers). Among them, the analysis of urban green space distribution is of great significance since it is directly impacted by urbanization.

Previous studies have focused on the impact of urbanization on green space distribution by utilizing certain methods ([Schetke et al. 2010](#)). For example, [Masoudi et al. 2021](#) compared the interrelationships among land use, spatial pattern, and cooling effect of urban green space between 2005 and 2015 in Singapore. The urban green space patterns were analyzed by using the landscape metrics. [Liu et al. \(2021\)](#) analyzed the landscape patterns of urban green space in Shanghai from 1990 to 2015 through integrated approaches (e.g., green space ratio, transition metrics, dynamic change degree, and landscape metrics). [Wang et al. \(2020\)](#) proposed an analytical framework including rank-size analysis, gradient analysis, and hotspot analysis to investigate the spatial and temporal dynamics of urban green space. [Liang et al. \(2017\)](#) evaluated the landscape patterns of urban green space by utilizing landscape metrics, correlation analysis, and factor analysis. [Jiao et al. \(2017\)](#) characterized the green space fragmentation in 1989, 2001, and 2013 through an improved gradient partitioning

method and six landscape metrics, which further quantified the relationships between green space fragmentation and urbanization.

Urban land use exert a complicated influence on urban green space distribution by changing the urban land use patterns ([Caspersen et al. 2006](#); [Kong and Nakagoshi 2006](#)). The analysis of urban green space distribution in response to urban land use changes will help improve the urban green space system ([Wang et al. 2020](#)). Several methods have been utilized to investigate the spatial and temporal patterns of urban green space dynamics. For example, quantifying the proportion of urban green space areas could reveal the distribution of green spaces ([Li et al. 2015](#)). Meanwhile, the landscape analysis could relate the urban green space distribution to urban land use and explore the ecological characteristics of the green spaces ([Liang et al. 2017](#)). Moreover, identifying the hotspots of urban green space change could reveal the spatial heterogeneity of urban green space, and investigate the effect of urban land use on urban green space ([Wang et al. 2020](#)). However, few studies have sought to combine these methodologies to investigate the spatial-temporal patterns of urban green space and its response to urban land use changes.

Furthermore, the resolutions of the urban land use maps and urban green space maps used in related studies are relatively low, and will thus result in untrustworthy analysis results ([Kong and Nakagoshi 2006](#); [Li et al. 2017a](#); [Teixeira 2021](#); [Wang 2009](#); [Wu et al. 2019](#)). Very-high-resolution (VHR) RS images provide new opportunities for urban green space mapping with detailed and updated information.

2.4 Summary

In this chapter, the lessons learned from the literature are briefly summarized:

The above examination of the relevant literature highlighted several areas worthy of further research. Chief among these is the need to generate high-performance urban land use maps, largely because high-quality land use maps can provide socio-economic information of inner-city rather than urban land cover maps. The examination result of this is that integrating RS and GBD provides new opportunities for delineating urban land use patterns. However, the integration strategies are still poorly understood and require significant research to explore.

Urbanization is one of the most important factors profoundly affecting the urban ecological systems, especially urban green spaces. The impact of urban land cover change (e.g., physical aspects of urbanization) on urban green space variations already has a long history. However, existing efforts on urban land use change (e.g., socio-economic aspects of urbanization) and related impact analysis are still based on urban planning maps from local governments, which have a very low resolution or inaccurate units. The urban land use change is highly related to the green space distributions according to related studies. It is thus necessary to provide urban land use change analysis with finer resolution so that more detailed information on urban green space in response to urban land use changes can be explored.

Urban land use exert a complicated influence on urban green space distribution by changing the urban land use patterns. The analysis of urban green space distribution in response to urban land use changes will help improve the urban green space system. Several methods have been utilized to investigate the spatial and temporal patterns of urban green space dynamics. However, few studies have sought to combine these methodologies to investigate the spatial-temporal patterns of urban green space and its response to urban land use changes.

The research in this thesis seeks to address the challenges and opportunities uncovered by the literature review. Research in all of these areas will contribute to the growing field of urban land use mapping, and the impact of urban land use changes on urban green space distribution, investigating analysis methodology to better map the urban land use, and investigate the influence of urban land use changes on urban green space.

Chapter 3: Methodological approach

3.1 Introduction

This chapter introduces the methodology workflow in this Ph.D. research (Figure 3-1). The workflow mainly includes four parts related to four research chapters. In the first part (**Chapter 4**), the integration strategies of urban land use mapping were categorized into decision-level integration (DI) and feature-level integration (FI) methods. The second part (**Chapter 5**) proposed a methodology framework based on DI and FI methods. The impervious surface mapping method was used for extracting impervious surfaces in the DI method, and the Points of Interest (POIs) calculation procedure was utilized to categorize urban parcels. The random forest (RF) model adopted this research for classifying urban land use types, as well as evaluating the importance of classification features. The optimization analysis was used for improving classification results in the third part (**Chapter 6**). The fourth part (**Chapter 7**) classified urban green space by using the Normalized Difference Vegetation Index (NDVI) calculation method. The impact analysis methods were used for investigating the distribution of urban green space in response to urban land use changes.

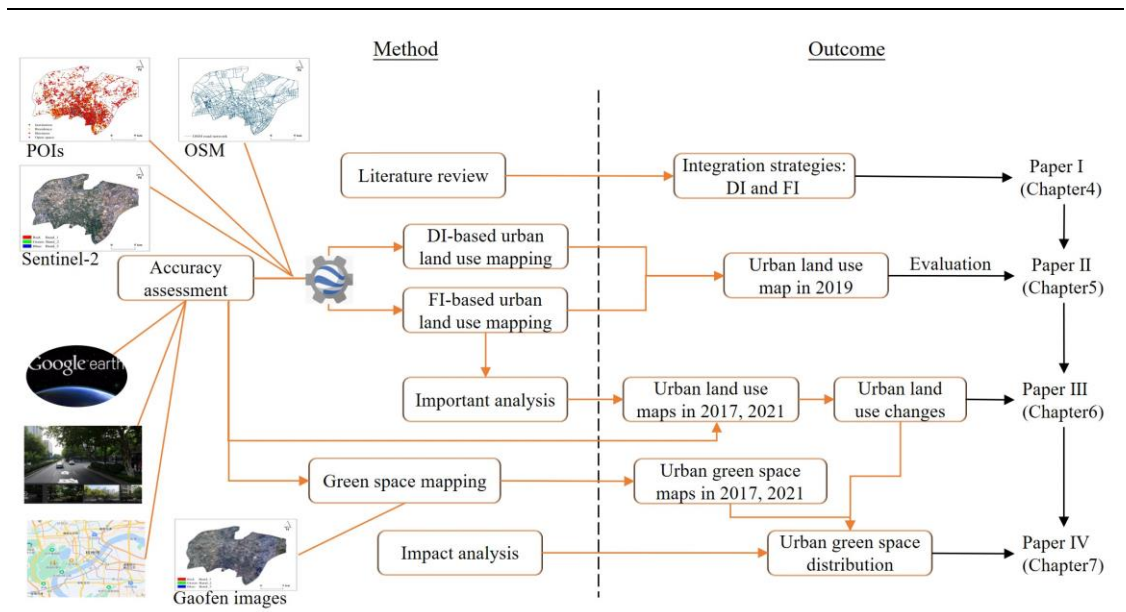


Figure 3-1 Methodology workflow.

3.2 DI-based mapping method

For categorizing urban land use, DI-based classification combines RS-based classification results with GBD-based classification results. To be more specific, RS and GBD features are evaluated independently before being integrated for urban land use mapping utilizing various modes and methods (Figure 3-2).

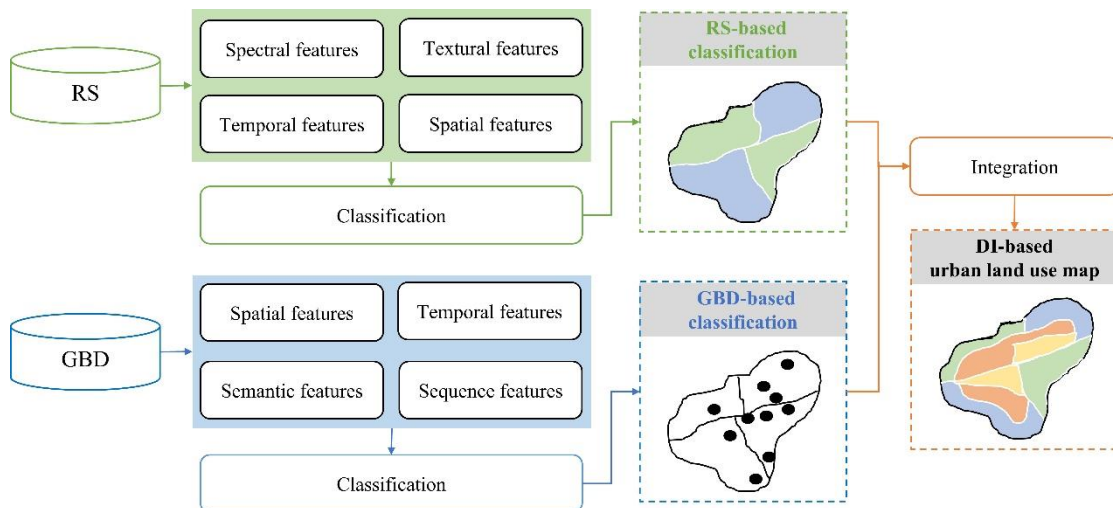


Figure 3-2 DI-based categorization strategies.

The DI-based classification method in **Chapter 5** was proposed based on the review of the integration strategy in **Chapter 4**. The DI-based classification method integrated the results of the RS-based classification (e.g., Sentinel-2) and GBD-based classification (e.g., POIs) by certain rules.

During the DI-based classification process, the quality of POIs was manually checked in terms of the data quantity and data location in the study site. The Baidu Map was used to check the POIs after reclassification. Around 95% of POIs were relatively accurate. Some of the wrong POIs were also manually corrected based on ArcGIS 10.2 Software.

In the DI-based method, the Sentinel-2 image was classified for producing the impervious and pervious classes by using the RF classifier based on the Google Earth Engine (GEE) platform. GEE combines a multi-petabyte catalog of satellite imagery and geospatial datasets with planetary-scale analysis capabilities. The training samples identified by visual interpretation through Google Earth and a field survey were put into the RF classifier for the classification of the impervious and pervious surfaces.

Finally, the type of each built-up parcel (e.g., urban parcels with impervious surfaces) was determined according to the POI frequency density calculation, and the weight calculation.

3.3 FI-based mapping method

In FI-based classification, RS features (e.g., spectral, textural, temporal, and spatial features), and GBD features (e.g., spatial, temporal, semantic, sequence features) are first extracted. These features are fused into integrated feature sets for urban land use classification (Figure 3-3).

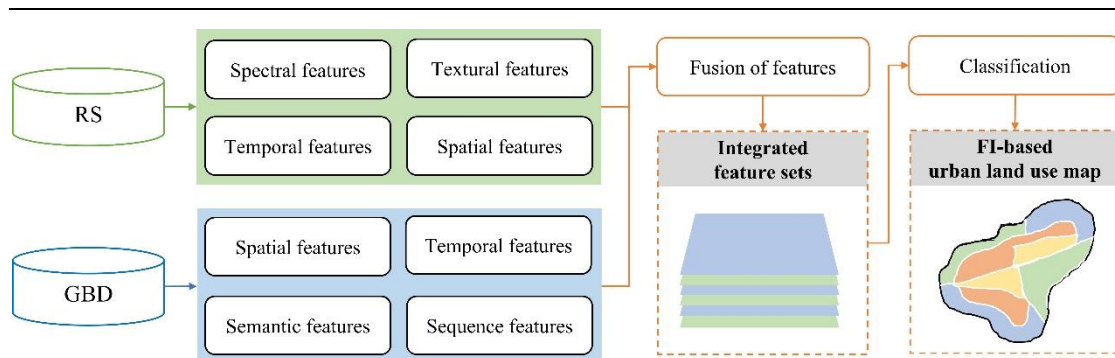


Figure 3-3 FI-based categorization strategies.

The FI-based classification in **Chapter 5** extracted 18 features from the Sentinel-2 images and POIs. The RF model was trained to categorize urban land use parcels by using all the features on the GEE platform. The code I used is attached below (Figure 3-4).

```

79
80 // Classifier
81 var classifier = ee.Classifier.smileRandomForest(10).train(training,'class',bandNames);
82 var classified = merged.classify(classifier);
83 Map.addLayer(classified.selfMask(),{min:0,max:1,palette:['red']},'classified')
84
85 Export.image.toDrive({
86   image : classified,
87   description: 'classified',
88   folder: 'hangzhou',
89   region: boundary.geometry(),
90   scale: 10,
91   crs: 'EPSG:4326',
92   maxPixels: 1e13,
93 })
94
95 // test
96 var test = testingPartition.classify(classifier);
97 var confusionMatrix = test.errorMatrix('class', 'classification');
98 print( confusionMatrix,'Confusion Matrix' );
99
100 // accuracy of the classifier
101 var trainAccuracy = classifier.confusionMatrix();
102 var overallAccuracy = trainAccuracy.accuracy();
103 var consumersAccuracy = trainAccuracy.consumersAccuracy();
104

```

Figure 3-4 RF classification code in GEE

3.4 Accuracy assessment

The accuracy assessment method is used in several chapters of this Ph.D. thesis. In **Chapter 5**, it is used for evaluating urban land use classification performance, as well as the impervious surface classification result. **Chapter 6** utilized the accuracy assessment to evaluate the urban land use classification results. **Chapter 7** evaluated the performance of urban green space maps by using the accuracy assessment.

In this research, it might not be appropriate to use the area proportion to estimate the accuracy of parcel-based urban land use maps. The reason is that the size of urban parcels varies, and the larger urban parcel cannot stand for a larger proportion of urban land use. Therefore, the area proportion of urban land use parcels should not be utilized for the estimation. To solve this problem, the quantity proportion of urban parcels can be used to evaluate the confusion matrix instead of the area proportion. This research evaluated the urban land use classification results by computing the overall accuracy (OA), user's accuracy (UA), and producer's accuracy (PA) based on the confusion matrix ([Foody 2002](#)). Also, approximate 95% confidence intervals of OA, UA, and PA were calculated based on the equations proposed by [Card \(1982\)](#). It should be noted that the training parcels used for FI-based classification in 2019 were selected by using stratified random sampling. While other training parcels and testing parcels were selected randomly.

As for the impervious surface mapping and urban green space mapping, I utilized the area proportion to estimate the classification accuracy. The confusion matrix including OA, UA, and PA was computed and approximate 95% confidence intervals

of OA, UA, and PA were calculated. In addition, the testing and training samples were selected randomly.

The training and testing parcels were identified by using the Baidu map, Baidu street view, and a field survey. The hybrid satellite map of Baidu map presents a top view of the urban landscape with high-resolution satellite imagery and names of ground features, which were further inspected or confirmed using the Baidu street view as it contains high-resolution photos and allows multiple perspectives ([Hoffmann et al. 2019](#); [Kang et al. 2018](#)). I determined the parcel type while different types of street views were observed for the same parcel, and there is no predominant characteristic referring to the rules in essential urban land use categories in China (EULUC-China) ([Gong et al. 2020](#)). To be more specific, I recorded a list of items, including the geolocations, classification types, landmark buildings and facilities, and mixed land use situation and their estimated proportions for each urban parcel. The training and testing parcels were selected by using simple random sampling from the list of parcels, except the training parcels used in the 2019 urban land use mapping. I used the stratified sampling strategy to select training parcels in 2019. For each parcel, the land use class was determined by visual interpretation based on the Baidu map. The training parcels were used for training the RF classifiers in FI-based classification, while testing parcels were used for accuracy assessment in both DI-based and FI-based classification, respectively.

3.5 Attribute importance analysis

In **Chapter 6**, the FI-based classification proposed in **Chapter 5** was improved using the “Variable Selection Using Random Forests for interpretation” (VSURF)

approach proposed by [Genuer et al. \(2015\)](#). RF models were built using different combinations of categories, resulting in two models using three categories (Table 3-1).

Table 3-1 Summary of RS and GBD features used in parcel-level urban land use mapping.

Data source	Features	Variables
Sentinel-2A/B	Spectral features	Mean, standard deviation, kurtosis, skewness of the near-infrared band EVI, NDBI, NDVI, NDWI
	Textural features	Angular second moment, contrast, dissimilarity, and entropy based on GLCM of the near-infrared band
POIs	Density features	Minimum, maximum, range, sum, mean and standard deviation
	Frequency features	Total number of all POIs Total number of each type of POIs The proportion of each type of POIs

Table 3-2, Table 3-3, and Table 3-4 showed the classification accuracy of each urban land use type via different feature combinations of features.

Table 3-2 Confusion matrix of classification results in the combination of spectral, texture, and frequency

a	I	R	B	O
I	0.896	0.418	0.499	0.390
R	0.620	0.837	0.611	0.478
B	0.522	0.636	0.854	0.406
O	0.332	0.510	0.540	0.736

Table 3-3 Confusion matrix of classification results in the combination of spectral, texture, and density.

b	I	R	B	O
I	0.683	0.455	0.350	0.446
R	0.478	0.621	0.435	0.340
B	0.373	0.438	0.674	0.378
O	0.368	0.391	0.378	0.718

Table 3-4 Classification accuracy of each urban land use type via different feature combinations of features. a: Spectral, texture, and frequency, b: Spectral, texture, and density.

Land use type	Accuracy	Feature combination	
		a	b
Institution	UA	0.826	0.725
	PA	0.697	0.557
Residence	UA	0.760	0.684
	PA	0.974	0.826
Business	UA	0.960	0.783
	PA	0.729	0.685
Open space	UA	0.698	0.784
	PA	0.421	0.351

3.6 Urban green space mapping method

Chapter 7 classified Gaofen images by calculating NDVI to obtain urban green space information. This research selected the testing samples by visual interpretation based on Google Earth. Following the approach discussed by [Stehman and Foody \(2019\)](#), this research randomly selected 1000 samples in 2017 and 1000 samples in 2021 for testing from the list of parcels (Figure 3-5). Specifically, 1000 testing samples in 2017 were identified including 295 non-green space samples, and 705 green space samples, while 1000 testing samples in 2021 were identified including 243 non-green space samples and 757 green space samples. Table 3-5 and 3-6 showed the confusion matrix of the urban green space mapping results in 2017 and 2021.

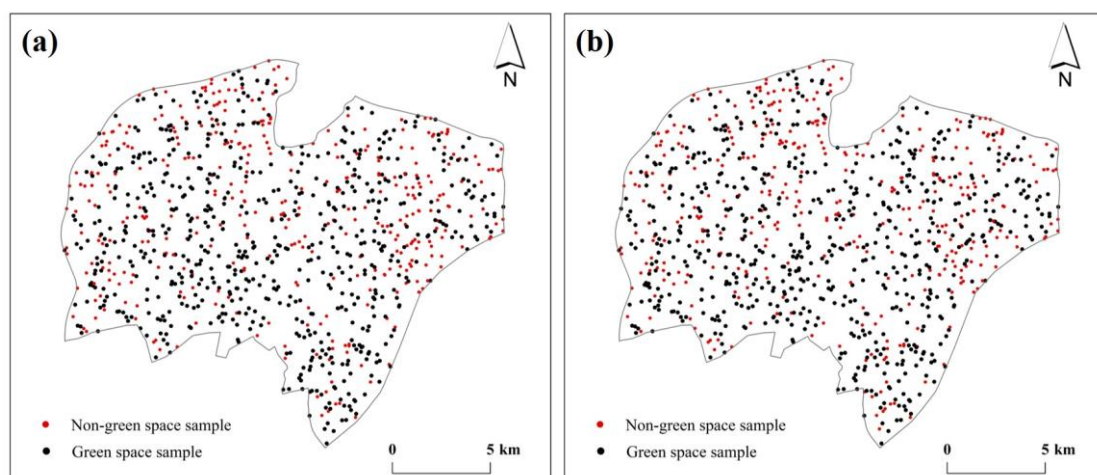


Figure 3-5 Testing samples of urban green space mapping: (a) urban green space samples in 2017; (b) urban green space samples in 2021.

Table 3-5 Confusion matrix of the urban green space mapping results in 2017. UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map (N: Non-green space; G: Green space).

		Map Category				
		Class	N	G	Total PA	OA
True Category	N	0.272	0.017	0.289	0.941±0.026	0.959±0.012
	G	0.023	0.688	0.711	0.967±0.012	
	Total(π_j)	0.295	0.705	1.000		
	UA	0.921±0.031	0.975±0.011			

Table 3-6 Confusion matrix of the urban green space mapping results in 2021. UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map (N: Non-green space; G: Green space).

		Map Category				
		Class	N	G	Total PA	OA
True Category	N	0.226	0.031	0.257	0.878±0.037	0.952±0.013
	G	0.017	0.726	0.743	0.977±0.010	
	Total(π_j)	0.243	0.757	1.000		
	UA	0.930±0.032	0.959±0.014			

The estimated areas of urban green space in 2017 and 2021 were analyzed based on the method proposed by [Olofsson et al. \(2014\)](#). Urban green space area estimated from the collected samples was 59.52 km² with a SE of 4.73 km² in 2017 (Table 3-7). While the estimation area of green spaces in 2021 was 81.29 km² with a SE of 5.49 km². The green space stratum in 2017 covered 22% of the total estimated soybean area. Its contribution to the total SE was also comparatively low. While the green space stratum in 2021 covered 30% of the total estimated soybean area.

Table 3-7 Urban green space area (km²) and uncertainty estimates based on sample data obtained for a random sampling design (N: Non-green space; G: Green space).

Stratum	2017		2021	
	Area	SE	Area	SE
G	59.52	4.73	81.29	5.49
N	205.67	16.15	183.91	14.63
Total	265.2	20.88	265.2	20.12

3.7 Impact analysis

In **Chapter 7**, an integrated method was proposed to investigate the urban green landscape in response to urban land use changes. The integrated method mainly includes statistical analysis, geospatial analysis, and landscape analysis.

Statistics (or statistical analysis) is the process of collecting and analyzing data to identify patterns and trends. In this research, I calculated the area and proportions of urban green space within different urban land use changes based on the Tabulate area Tool in ArcGIS 10.2, which has the potential to understand the distribution of urban green space (Figure 3-6).

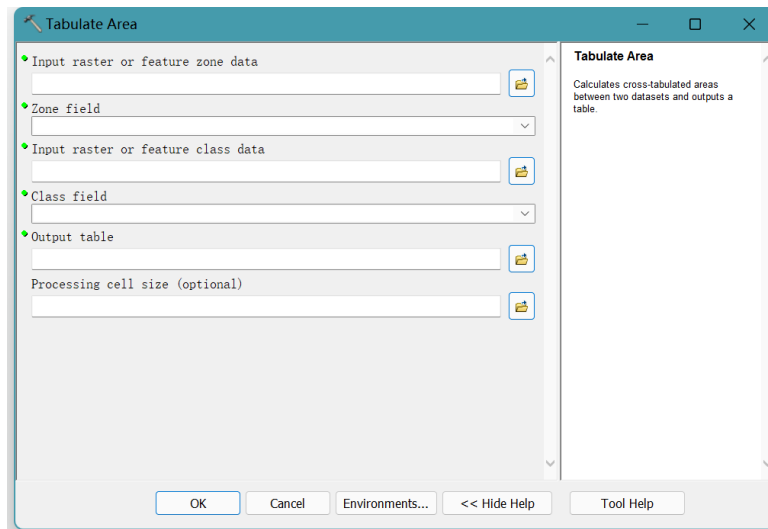


Figure 3-6 Operation of statistical analysis

The landscape analysis could relate the landscape patterns of urban green space to urbanization and determine their influence on the ecological attributes of the urban environment ([Daz et al. 2020](#); [Grafius et al. 2018](#)). The FRAGSTATS 4.2 was used to calculate the landscape metrics of urban green space patches. FRAGSTATS is a spatial pattern analysis program for quantifying the structure (i.e. composition and configuration) of landscapes. The rasters of urban green space within different urban land changes were extracted for the calculation. To understand the fragmentation, complexity, and aggregation of urban green space patches, I selected 6 landscape metrics including patch density (PD), largest patch index (LPI), Mean shape index (SHAPE_MN), and patch cohesion index (COHESION). The three landscape-level metrics include Contagion index (CONTAG) and Shannon's Evenness Index (SHEI). Among these, PD, LPI, and SHAPE_MN can represent the fragmentation and complexity of urban green space patches ([Bosch et al. 2020](#); [Zhang et al. 2020b](#)); COHESION could reflect the connectivity ([Inkoom et al. 2018](#)); CONTAG presents the connectivity of the total landscape ([Miller et al. 2020](#)); SHEI presents the distribution of landscapes ([Pindral et al. 2020](#)).

Moreover, the geospatial analysis could reveal the spatial heterogeneity of urban green space, and further elucidate the effect of urban development on urban green space (Liu et al. 2020c). In this research, the hotspot analysis was adopted to investigate the increasing and decreasing regions of urban green space. The Getis-Ord G_i^* Tool in ArcGIS 10.2 was used to produce the hotspot maps (Figure 3-7).

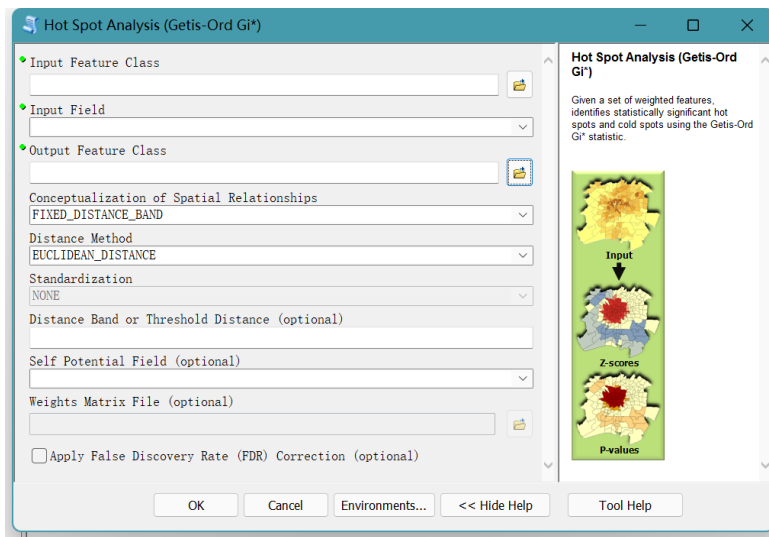


Figure 3-7 Operation of hotspot analysis

Chapter 4: Integrating remote sensing and geospatial big data for urban land use mapping:

A review

Stage of publication - Accepted in International Journal of Applied Earth Observations and Geoinformation (20 August 2021).

4.1 Introduction

With the advent of the Anthropocene ([Ellis and Ramankutty 2008](#); [Steffen et al. 2011](#)), urbanization is accelerating and the urban population is predicted to grow from 4.2 billion (57.5% of the world population) in 2018 to about 6.7 billion (69.1%) in 2050 ([Seto et al. 2011](#)). Such increasing human-induced influences are changing urban land in different dimensions from physical aspects (urban land cover) to socioeconomic aspects (urban land use) ([Elmqvist et al. 2019](#); [Hersperger et al. 2018](#)). A large number of high-accuracy urban land cover products (mainly physical characteristics) at the annual level with relatively high spatial resolution have been developed worldwide ([Li et al. 2020b](#); [Liu et al. 2018a](#); [Zhou et al. 2018b](#)). However, urban land governance and planning need more information on urban land use, which is particularly complex and includes both physical aspects and socioeconomic aspects. Unfortunately, high-quality urban land use products with timely and accurate information related to human activities are still limited ([Gong et al. 2020](#)). Understanding the start-of-the-art of

existing urban land use mapping efforts, considering both physical and socioeconomic functions, would enable better urban land management and monitoring ([Martí et al. 2019](#); [Yammine et al. 2018](#)).

A wide range of satellite remote sensing (RS) data (e.g., Moderate-resolution Imaging Spectroradiometer (MODIS), Landsat TM/ETM/OLS, Defense Meteorological Satellite Program Operation Linescan System (DMSP-OLS)) have been used to study the structures, boundaries, and areas of cities ([Gong et al. 2019](#); [Huang et al. 2021](#); [Schneider et al. 2010](#)). Nevertheless, the complexity and diversity of functional patterns in urban areas cannot be captured well by using RS only due to limited information (e.g. spectral, textural, and temporal information) from RS techniques ([Cao et al. 2020](#)). Advances in information and communication technologies make it possible to get access to geospatial big data (GBD) ([Li et al. 2016](#); [Li et al. 2021](#); [Yin et al. 2021b](#)). Fixed and mobile sensors such as environmental sensors, cameras, webcams, social media, or even urban residents through their regular activities ([Wu et al. 2015](#)) create tremendous GBD every day. These data such as mobile phone data ([Gong et al. 2020](#)), traffic trajectories ([Yu et al. 2019](#)), geo-tagged photos ([Cadavid Restrepo et al. 2017](#); [Krylov et al. 2018](#)), Points of interest (POIs), and social media data ([Huang et al. 2018c](#)) provide an alternative approach to uncover how cities function ([Ye et al. 2016](#)). It is possible for examining the physical and socioeconomic characteristics of the urban land system by taking both the advantages of RS and GBD ([Qi et al. 2019](#); [Song et al. 2021](#); [Xiong et al. 2021](#)).

Despite the great potential of integrating RS and GBD for providing better insights into urban land use, it is challenging to combine them due to the differences in the spatial data quality (e.g., semantic, timestamp, and scale), technical format, and data

structure ([Liu and Long 2015](#)). Summarizing the features of RS and GBD and integration strategies in the literature are needed for guiding future studies and help to understand more detailed urban functional patterns.

In this context, this study examined the literature on the nature of RS and GBD, as well as their integration strategies in urban land use classification, and identified the opportunities and challenges for synthesizing RS and GBD (Table S1). The primary objective of this paper is to review the state-of-the-art in this field by considering (1) the key characteristics of RS and GBD and (2) the methods for integrating RS and GBD. We consider only satellite-based RS and do not consider RS data obtained from airborne platforms. This review is organized into six sections. Section 2 summarized the transformation from urban land cover to urban land use. In section 3, we summarized the commonly used RS and GBD features for urban land use categorization. In Section 4, the integration strategies were analyzed systematically. In Section 5, we discussed the challenges and potential applications of the integration of RS and GBD on urban land use maps. Section 6 concluded the main findings and implications.

4.2 Evolution from urban land cover to urban land use

Using satellite data to map urban land cover has a long history ([Howarth and Boasson 1983](#); [Patino and Duque 2013](#); [Reba and Seto 2020](#); [Yang et al. 2003a](#)). Examples of existing efforts for global urban land cover products that have been derived from RS are shown in Table 4-1. The data source, the nomenclature of urban land, spatial resolution, time period, and reference for each global urban land cover map were summarized. Most urban land cover maps were obtained from coarse spatial resolution images (100m-10km), such as MODIS and Advanced Very High-Resolution

Radiometer (AVHRR) ([Friedl et al., 2002](#); [Schneider et al., 2010](#)). With the substantial progress of RS techniques, recent maps were derived from moderate spatial resolution images (10-100m), including Satellite Pour l' Observation de la Terre (SPOT) and Landsat images ([Deng et al. 2019](#)). The time period of these global urban land cover products has transformed from a single period to repeated observations, which could provide better quality and time series urban land cover information ([Gong et al. 2019](#); [Li and Chen 2018](#); [Momeni et al. 2016](#); [Reda and Kedzierski 2020](#)). Overall, these global urban RS studies focused on the identification of physical urban attributes (e.g., impervious surface, built-up areas, artificial surfaces, and urban extent) that have provided opportunities for a better understanding of global urbanization's effects on human civilization and the environment ([Defries and Townshend 2007](#); [Li et al. 2020b](#); [Schneider et al. 2009](#); [Zhu et al. 2019](#)). Despite the aforementioned extensive applications of RS data for mapping urban land cover, more specific information of inner-urban functions cannot be retrieved by using RS only ([Li et al., 2020a](#); [Liu et al., 2015](#)).

Table 4-1 Comparison of existing global urban land cover products.

Products	Data	Nomenclature of urban land	Spatial resolution	Time period	Reference
GLCC	AVHRR	Built-up areas	1km	1992,1993	Loveland et al., 2000
UMD1km	AVHRR	Urban and built	1km	1992,1993	Hansen et al., 2000
GRUMP	VMAP, Census data, DMSP-OLS, Maps	Urban extent	1km	1995	CIESIN et al., 2011
GLC2000	SPOT-Vegetation, DMSP-OLS	Artificial surfaces and associated areas	1km	2000	Bartholome and Belward, 2005
IMPSA	DMSP-OLS	Impervious surface	1km	2000	Elvidge et al., 2007
NTL-Urban	DMSP-OLS	Urban extent	1km	1992-2013	Zhou et al., 2018
MOD500	MODIS	Non-vegetated, human-constructed elements	500m	2001-2017	Friedl et al., 2002
GHSL	Fine-scale satellite imagery,	Built-up areas	500m	1975, 1990, 2000, 2015	Pesaresi et al., 2013

	census data, and OSM				
GlobCover	SPOT- Vegetation	Artificial surfaces and associated areas	300m	2005, 2009	Arino et al., 2007
CCI-LC	SPOT- Vegetation	Urban extent	300m	1992- 2015	Defourny et al., 2018
GlobalLand30	Landsat TM/ETM+	Artificial surfaces	30m	2000, 2010	Chen et al., 2015
HBASE	Landsat TM/ETM+	Built-up and settlement extent	30m	2010	Wang et al., 2017
GMIS	Landsat TM/ETM+	Impervious surface	30m	2010	Colstoun et al., 2017
FROM-GLC	Landsat TM/ETM+/OLI	Impervious surface	30m	2010, 2015, 2017	Gong et al., 2013
Global Urban Land	Landsat TM/ETM+	Impervious surface	30m	1990, 1995, 2000, 2005, 2010	Liu et al., 2018
GUF	TerraSAR-X, TanDEM-X	Built-up areas	12m	2011	Esch et al., 2012
FROM-GLC10	Sentinel	Impervious surface	10m	2017	Gong et al., 2019b

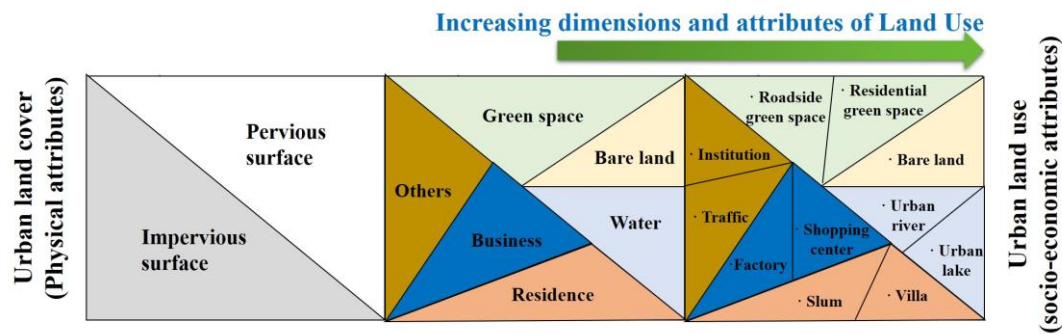


Figure 4-1 Transformation of the need for urban land products from physical attributes to socioeconomic attributes due to enhanced anthropogenic activities. The three boxes above represent the transformation from urban land cover (e.g., impervious surface and pervious surface) to urban land use (e.g., traffic, institution, urban lake, residential green space, and others). It should be noted that the change of area proportions of different urban land use types is not reflected.

The demands for urban land products have changed gradually, with increasing information needs on socioeconomic properties, emphasizing a transformation from urban land cover to urban land use (Figure 4-1). The multi-sourced GBD can contribute to the understanding of socioeconomic characteristics of urban land use, and identify how people use lands ([Cadavid Restrepo et al. 2017](#); [Krylov et al. 2018](#); [Srivastava et al. 2018](#)). Recently, GBD has been used in conjunction with RS data to extract urban land use information ([Liu et al. 2015](#)). Therefore, understanding characteristics derived from RS and GBD and their integration methods are necessary for urban land use mapping ([Li et al. 2017b](#)).

4.3 RS and GBD features for urban land use mapping

The extraction of RS and GBD features is the most essential procedure for urban land use recognition because the performance of urban land use maps relies heavily on

these features. The purpose of this section is thus to introduce several commonly used RS and GBD features for categorizing urban land use types (Figure 4-2).

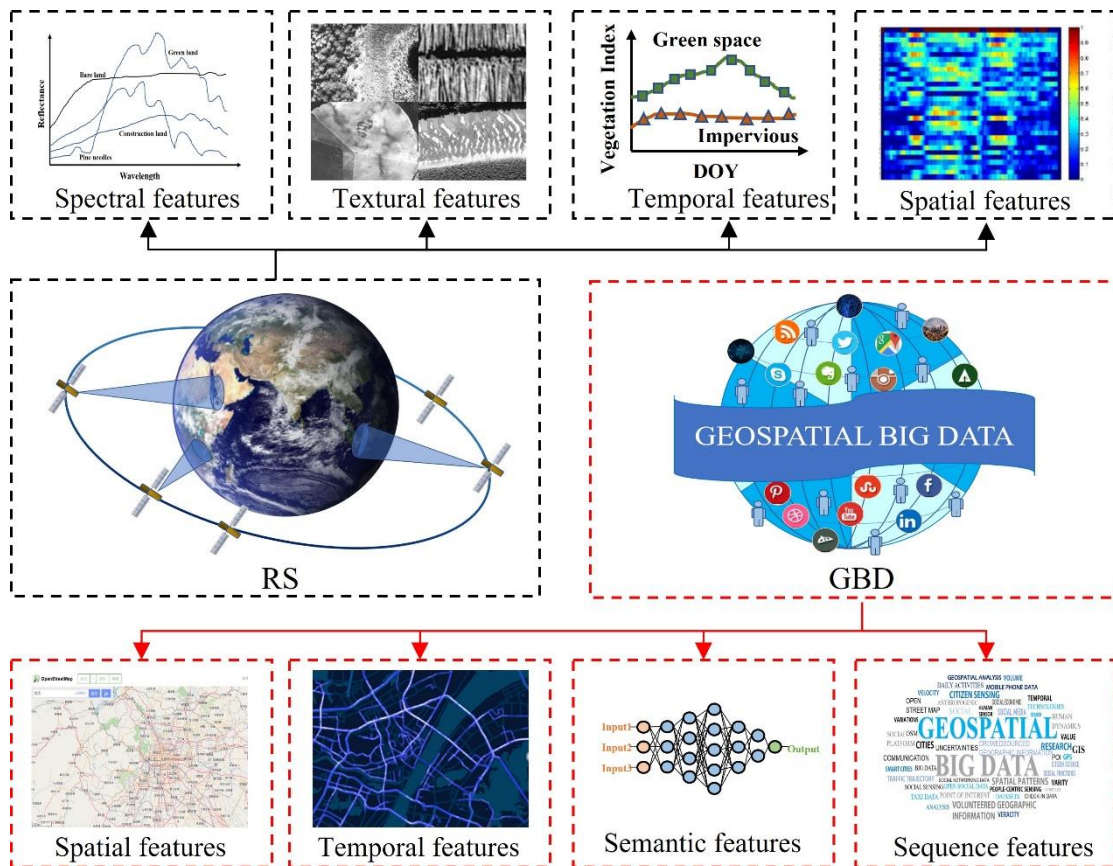


Figure 4-2 Summary of the features of RS and GBD. The dotted black box and red box in the middle show the commonly used RS data (e.g., MODIS, Landsat, Sentinel) and GBD (e.g., traffic, social media data, geo-tagged photos), respectively. The upper dotted black boxes represent the features (spectral, textural, temporal, and spatial) extracted from RS, while the dotted red boxes below represent the features (spatial, temporal, semantic, and sequence) extracted from GBD.

4.3.1 RS-based features

The features derived from RS images used in urban land use classification could be categorized into spectral, textural, temporal, and spatial features ([Gong and Howarth 1989](#); [Zhu et al. 2017](#)). Among them, spectral and textural features are common characteristics of RS data to extract urban land use information because different

textures and spectra could reflect different urban land use types ([Zhu et al. 2019](#)). The temporal features have proven beneficial for improving urban land use maps by providing valuable information (e.g., time series information) on urban land use types ([Zhu et al. 2012](#)). Recently, deep learning techniques provide new possibilities for extracting spatial features automatically from the very-high-resolution (VHR) satellite imageries such as WorldView-3, Gaofen-2, and SPOT-5 ([Zhang et al. 2020a](#)). Spatial features could help classify urban land use at a very detailed level ([Zhao et al. 2019](#)). The details of the RS features are specified as follows.

1) Spectral features: Generally, the spectral features of urban land show lower reflectance in the near-infrared region (NIR), comparing to vegetation, which has higher reflectance in NIR ([Herold et al., 2003](#); [Herold et al., 2004](#)). Additionally, the spectra for the visible, short-wave infrared region (SWIR), and microwave regions were also found to be suitable for characterizing urban objects ([Heiden et al. 2007](#)). Recently, spectral features from the increased number of bands (e.g., from Landsat to Sentinel-2) provide an opportunity for the acquisition of detailed information on the physical attributes of urban land use, but it also leads to data redundancy due to the high correlation between adjacent bands ([Okujeni et al. 2018](#)). [Liu et al. \(2020b\)](#) and [Zhang et al. \(2017b\)](#) extracted spectral features from RS imagery for classifying urban land use. Both the methods calculated the mean and standard deviation of each band by using a certain window.

2) Textural features: Textural features contain rich information of the spatial distribution of tonal changes, as well as the structural arrangement of surfaces and their relationships to the surrounding environment ([Gong and Howarth 1989](#); [Haralick et al. 1973](#)). Different textures (e.g., coarse, smooth, rippled, irregular, and lineated) show

different image characteristics, such as homogeneity, linear structure, and contrast ([Kuffer et al., 2016](#); [Wurm et al., 2017](#)). Therefore, textural features could help to increase the accuracy of land use categorization in heterogeneous landscapes ([Jin et al. 2014](#); [Pacifiçi et al. 2009](#)), where ground objects with different sizes, patterns, structures, and shapes co-exist ([Lu and Weng 2006](#)).

3) Temporal features: Temporal features refer to the differences caused by the changes in the spectral and textural features of urban surfaces over time. Due to the seasonality of vegetation growth, it has proven to be effective in improving vegetation and other land cover mapping accuracy ([Dong et al., 2019](#); [Zurita-Milla et al., 2013](#)). The extraction of urban land usually tends to be less accurate in the autumn and winter due to more bare land ([Weng et al. 2009](#)). However, it is still a challenge to distinguish the variety of processes that generate different time series, for example, due to climate, topography, and terrestrial vegetation ([Pflugmacher et al. 2019](#)).

4) Spatial features: Along with spectral, textural, and temporal features, the most commonly used feature extracted from RS data is the spatial feature. Recent studies used deep learning techniques such as the supervised convolutional neural network (CNN) models and unsupervised autoencoders (AE) models to extract spatial information automatically from RS images ([Reichstein et al. 2019](#)). Deep learning algorithms, which extract high-level spatial information provided by hierarchical structures, demonstrate remarkable capacity in image representation and understanding in these studies. Traditional approaches such as Random forest models (RF), Support vector machine (SVM), and Decision tree (DT) can only process basic features (e.g., spectral, textural, and temporal features) from RS images. Due to the fine structural information (i.e., spatial details) of urban land use in VHR RS images, VHR RS images

were found to be commonly used by deep learning techniques for obtaining urban land use information ([Ma et al. 2019](#)).

4.3.2 GBD-based features

The development of mobile positioning, wireless communication, and the Internet of Things (IoT) provides opportunities for the rapid growth of big data ([Kitchin, 2013](#)). According to [Kitchin and McArdle \(2016\)](#), big data is defined in part by its large size and in part by its characteristics, such as volume, variety, and velocity. [Liu et al. \(2016\)](#) further defined the characteristics of big data as exhaustivity, relationality, veracity, value, and variability. In the IBM Annual Report, 2.5 terabytes of data are generated every day, with 80 % of these data (pictures, texts, and videos) being geo-referenced or capable of being geo-referenced ([Munoz et al. 2016](#)). Therefore, a large proportion of big data is likely to be the GBD (i.e., big data with geographical reference). GBD is generated every day mostly by fixed and mobile sensors such as environmental sensors, cameras, webcams, social media, or even residents' daily activities ([Brovelli et al., 2015](#)).

The most commonly used GBD for land use mapping are social sensing (SS) ([Liu et al. 2015](#)), citizen sensing (CS) ([Jiang et al. 2016](#)), social media data (SMD) ([Ilieva and McPhearson 2018](#)), and volunteered geographic information(VGI) ([Goodchild and Glennon 2010](#)). The descriptions of SS, CS, SMD, and VGI are provided in Table 4-2. The contents and concepts of SS and CS are much broader than GBD. SS data is used in a variety of applications besides urban land use mapping. CS does not contain the data produced by companies and institutions and VGI focuses on user-generated data. However, the term GBD encompasses all of these above-mentioned geospatial data.

Table 4-2 Comparison of concepts related to GBD in previous studies.

Concepts	Refs.	Description
Social Sensing (SS)	Liu et al., 2015	A series of data sources with spatiotemporal information which record human activities, as well as the methods and applications based on such data source.
Citizen Sensing (CS)	Jiang et al., 2016	Datasets contributed by citizens provide benefits for themselves and policymakers.
Social Media Data (SMD)	Ilieva and McPhearson, 2018	Information of ‘big data’ from social media such as Facebook, Flickr, etc.
Volunteered Geographic Information (VGI)	Goodchild and Glennon, 2010	Geographic data provided voluntarily by people use technologies to generate, assemble, and disseminate information.

Each record of GBD contains spatial, temporal, semantic, and/or sequence information associated with individuals reflecting human behavior, although the quality of this information may vary in space and time (Figure 4-2 and Table 4-3). Due to the high correlation between human spatiotemporal activities and urban dynamic socioeconomic attributes, these emerging GBD can help to capture the growing complexity of urban functional patterns. Therefore, in order to better classify and understand urban land use, we add GBD to the identification of socioeconomic and human activities.

Table 4-3 Summary of the GBD and relevant features used in urban land use mapping.

Data sources	Spatial features	Temporal features	Semantic features	Sequence features
Mobile phone data	√	√		
Traffic data	√	√		
Social media data	√	√	√	√
Geo-tagged photos	√		√	
Maps	√			√
Search engine data	√		√	√
Smart card data	√	√		

1) Spatial features: Almost all GBD can provide spatial information (Table 4-3). For example, the OSM has proven to be a useful spatial data source including land use information like buildings, roads, and parks ([Helbich et al. 2012](#)), which could be used to extract training pixels from RS images for classifying urban land use types ([Johnson and Iizuka 2016](#); [Wan et al. 2017](#)). Google maps, Gaode maps, and other maps have also been successfully applied to urban land use mapping ([Xu et al. 2020](#); Zhang et al., 2020). Furthermore, social media data can provide indirect spatial location data ([Long et al. 2018](#); [Yao et al. 2018](#)), for example, [Zhan et al. \(2014\)](#) inferred the urban land use types in New York city by using check-in social media data on Twitter.

2) Temporal features: There are many examples of data sources (e.g., mobile phone data, traffic data, social media data, and smart card data) with temporal features that

could reveal the mobility patterns of human activities ([Pan et al. 2013](#)). For instance, [Gong et al. \(2015\)](#) analyzed the spatiotemporal characteristics of nine daily activity types referring to the trip purposes of taxi trajectory data. [Shi et al. \(2019\)](#) extracted the temporal variation in WeChat (i.e., WeChat is China's most popular messaging app) user density from different land use categories for urban land use classification combined with RS data. The activities of human beings with temporal features can be determined to indicate the social function and patterns of urban land use ([Chen et al. 2017](#); [Frias-Martinez and Frias-Martinez 2014](#); [Pei et al. 2014](#)).

3) Semantic features: Photographs are an important element of GBD. Examples include street view photographs, crowd-sourced geo-tagged photos, and social media photos ([Kang et al., 2018](#); [Xu et al., 2017](#)). Semantic features obtained from photographs have much in common with those obtained from RS data. However, there are also important distinctions that present challenges for analysis. RS is usually undertaken by national and international organizations following established scientific and engineering principles, including regular acquisition cycles. While crowdsourced photographs may be acquired by a range of organizations (e.g., Google Street View) and private individuals. They provide fine-resolution data, but the spatial and temporal sampling may be ad hoc, and the quality can be highly variable ([Hu et al. 2015](#)). Recently, with the development of image recognition and deep learning technology, extracting semantic features from photographs and applying them to the perception of places have become possible ([Xu et al. 2017](#)).

4) Sequence features: Social media data and search engine data have become an important source of GBD in current research ([Zheng et al. 2019](#)). Studies utilizing sequence features mainly include the following three aspects ([Cheng et al. 2018](#); [Long](#)

and Liu 2017): (a) to obtain the evaluation index or topic of a place; (b) to obtain information on emotions related to the place, such as happiness or depression (Yang and Mu 2015; Zheng et al. 2019); (c) to identify public attention to hot events, such as disasters, diseases, and accidents. Specific examples of sequence information include sentiment, opinions, locations, time, and places. For example, Mitchell et al. (2013) generated a method for analyzing the correlations between human being’s real-time expressions and others like emotional, geographic, and demographic characteristics.

4.4 Integration of RS and GBD for urban land use mapping

The RS and GBD features can be then combined using different approaches for urban land use categorization. According to the fusion mode between RS and GBD features, RS and GBD integration techniques in the literature could be divided into feature-level integration (FI) and decision-level integration (DI).

4.4.1 Feature-level integration

In FI-based classification, RS features (e.g., spectral, textural, temporal, and spatial features), and GBD features (e.g., spatial, temporal, semantic, sequence features) are first extracted. These features are fused into integrated feature sets for urban land use classification (Figure 4-3).

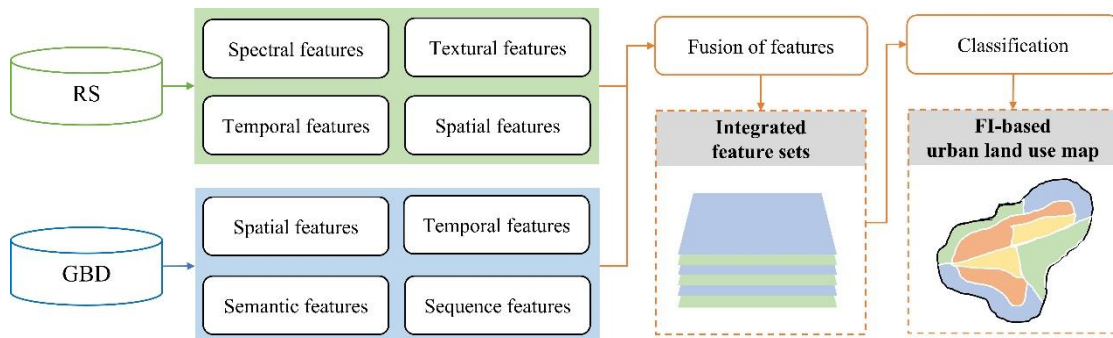


Figure 4-3 FI-based categorization strategies.

Several efforts have been made for extracting urban land use information using the FI method ([Bao et al. 2020](#); [Chang et al. 2020](#); [Hu et al. 2016](#)). For the study unit of the FI-based classification, most studies utilized parcel-level or object-level as the study unit since object-level or parcel-level units are compatible with both RS features and GBD features ([Liu et al. 2017](#)). To be more specific, the format of the RS features is usually grid, while for the GBD features is various. It is thus necessary to unify the two kinds of features. Parcel-level classification units can be generated by using the OSM road network or other road data ([Huang et al. 2020](#)). The urban parcels are obtained by removing the road buffers from the study site. Sometimes, the study site should also exclude the rivers according to the actual situation. Furthermore, the elevated road in cities would interfere with the segmentation results and need to be considered. Pixel-based and grid-based units are also used for FI-based urban land use classification ([Dong et al. 2020](#)).

For the feature extraction stage, the FI-based classification method extracts RS and GBD features respectively. For example, [Zhang et al. \(2019\)](#) delineated physical features including spectral features (e.g., mean and standard deviation for each band), textural features (e.g., contrast, entropy, correlation, and homogeneity for each band) from RS for urban land use categorization by integrating GBD features (e.g., POI words and real-time Tencent users words). [Sun et al. \(2020\)](#) extracted RS features (spectral and textural features), GBD features (POI frequency, POI spatial distribution), and other features to train the RF classifier for recognizing urban functions (e.g., residence, business, and industries).

In the feature integration stage, the FI method usually classifies RS and GBD features by using machine learning techniques such as RF classifier, SVM, and DT. For

example, [Gong et al. \(2020\)](#) proposed a research method that is extracting several features from RS and GBD for training the RF classifier, which has proven to be a new way for mapping urban land use over large areas. [Du et al. \(2020\)](#) generated urban functional zones by removing road buffers, and then the functional zones were classified by coupling Latent Dirichlet Allocation (LDA) and SVM. The proposed method is promising for urban land use mapping over large areas. Recently, deep learning techniques such as CNN and AE models are also used for urban land use mapping, which could help improve the classification performance ([Mao et al. 2020](#)).

4.4.2 Decision-level integration

For categorizing urban land use, DI-based classification combines RS-based classification results with GBD-based classification results. To be more specific, RS and GBD features are evaluated independently before being integrated for urban land use mapping utilizing various modes and methods (Figure 4-4).

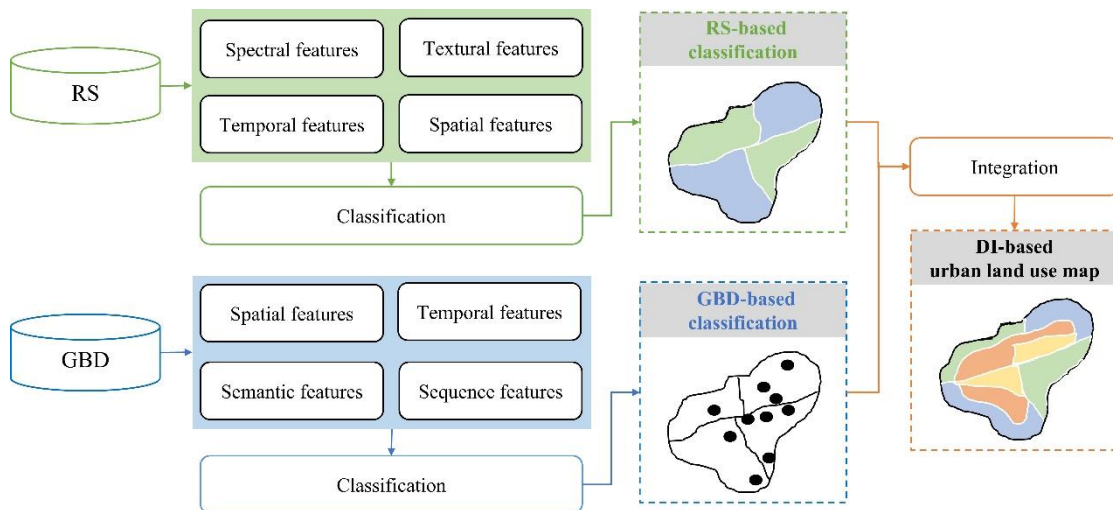


Figure 4-4 DI-based categorization strategies.

Compared to the FI method, the basic unit for the DI method is more diverse ([Chang et al. 2015](#)). The characteristic of DI-based classification is to combine the RS-based classification and GBD-based classification for extracting urban land use

information. It is thus not necessary to unify the spatial units for RS-based classification and GBD-based classification, and the spatial units for the two classification methods will be different. For example, [Jia et al. \(2018\)](#) first classified Gaofen images and mobile phone positioning data by using a support vector machine, and then the two classification results were fused for mapping urban land use based on grid-level units. [Tu et al. \(2018\)](#) utilized the hierarchical cluster analysis to combine RS-based landscape metrics and GBD-based human activity metrics for investigating urban functional patterns in terms of object-based units. In addition, [Zhao et al. \(2019\)](#) delineated the geographical object by using OSM data to train the CNN model. [Chen et al. \(2018\)](#) identified urban green space by utilizing parcels generated by the OSM road networks as the basic units.

There are several methods for the integration of the RS-based classification and GBD-based classification in the DI method based on certain decision fusion strategies such as hierarchical clustering, overlaying, and labeling ([Anugraha et al. 2020](#)). For example, [Xu et al. \(2020\)](#) proposed a framework that extracts spatial geographic characteristics from RS images by using deep neural networks and functional distribution characteristics from POIs, and further normalized the two results to identify urban functional regions. [Song et al. \(2018a\)](#) used an object-based approach to generate urban objects by using RS data. Then, the objects were further classified and aggregated using POIs. Furthermore, [Zhong et al. \(2020\)](#) presented a method by using the rule-based category mapping (RCM) model to integrate RS-oriented results and GBD-oriented results for extracting urban functional zones. In [Zhong et al. \(2020\)](#)'s work, POIs data and RS images were classified through different machine learning methods

based on the parcel-level unit, then the two results with different classification systems were combined by weighting the word frequency within each parcel.

4.5 Discussion

4.5.1 Advantages and disadvantages of FI-based and DI-based classification

The FI method lies in the realization of considerable information compression of RS data and GBD, which is conducive to real-time processing. This method has the advantages of a low level of human intervention and short processing time. In addition, feature optimization and deep interactions between RS and GBD features can be achieved. Several studies have been analyzed to quantify the relative importance of independent features in the FI method ([Zhang et al. 2019](#)). Furthermore, heterogeneity may occur from variations in data quality, time periods, data formats, and data scales, between RS and GBD, resulting in different representations, descriptions, and interpretations of the goal.

The DI-based classification method has been a fundamental contribution of integrating RS and GBD to urban knowledge ([Cai et al. 2017](#); [Zhao et al. 2019](#)). RS and GBD features could be calculated and processed respectively in the DI method, which avoids the feature integrating and conflicting issues. Specifically, RS and GBD with various features could be processed by different methods, and then integrated by certain models. However, the accuracies of DI-based classification are affected by the two kinds of processing procedures because the mapping results are obtained by overlapping RS-based classification and GBD-based classification. It is thus important

to control the procedures for both RS and GBD classification for high accuracy and better performance.

4.5.2 Limitations and future consideration for RS and GBD integration

The main limitation is that the modality gap (data structure, data format, and data quality) between RS data and GBD, which brings difficulty for the integration. GBD has different sources (modern sensors, geo-tagged web, ground surveying, mobile mapping, and social media platforms), spatiotemporal resolution, and structures from that of RS data ([Ali et al. 2017](#)). For example, the data formats of GBD include the image, geo-tagged text, video, and vector. Comparatively, the most commonly used RS data is a raster ([Zhu et al. 2019](#)). Furthermore, GBD is not collected evenly across space because the composition of participation (i.e., sensors, platforms, and social habits) in GBD varies across political, cultural, demographics, and commercial factors, while RS data usually has spatially consistent observation frequencies ([Chen et al. 2020b](#); [Yu et al. 2018](#)). Some areas might have the data sparsity issue of GBD, which might present problems for mapping based on the integration of RS and GBD owing to the lack of GBD. More specifically, social networking platforms such as WeChat, Sina Weibo, and Tencent are widely used in China, while Twitter, Facebook, and Instagram are more popular in Western countries ([Li et al. 2018b](#); [Wang et al. 2016](#); [Zhou and Zhang 2016](#)). The emergence of deep learning technologies provides an opportunity to bridge the gap between different data modalities.

The advances in classification algorithms, computing platforms, and data sources are beneficial for mapping urban land use by integrating RS and GBD. Deep learning, as a novel branch of machine learning, establishes fundamental parameters about the data and trains the computer to learn on its own by detecting patterns using a multi-

layered approach ([Reichstein et al. 2019](#); [Zhu et al. 2017](#)). Several studies have shown that deep learning is particularly effective in integrating RS and GBD for urban region function recognition ([Ma et al. 2019](#); [Qin et al. 2018](#)). Furthermore, the fast growth of cloud computing platforms offers a promising solution for processing large amounts of RS data and GBD. For example, [Yin et al. \(2021b\)](#) processed the Sentinel images and POIs data for urban land use classification by using the Google Earth Engine (GEE). GEE could provide a range of data processing methods as well as RS images at various temporal and spatial scales, which is beneficial for improving data computing difficulties ([Gorelick et al. 2017](#)). In addition to GEE, other platforms such as Earth Observation Data Center and the Amazon Web Services have also been used for analyzing RS data and GBD on urban land use classification. Furthermore, it is necessary to consider auxiliary data (e.g., census data, statistical data, weather data, hydrological data, and digital elevation data) for more space- and time-referenced information on urban land use classification that integrates RS and GBD ([Taubenböck et al. 2009](#)).

4.5.3 Potential applications of urban land use maps derived from RS and GBD integration

Urban land use maps integrating RS and emerging GBD provide more potential in urban management such as urban planning, urban environment assessment, urban disaster monitoring, and urban traffic analysis (Figure 4-5).

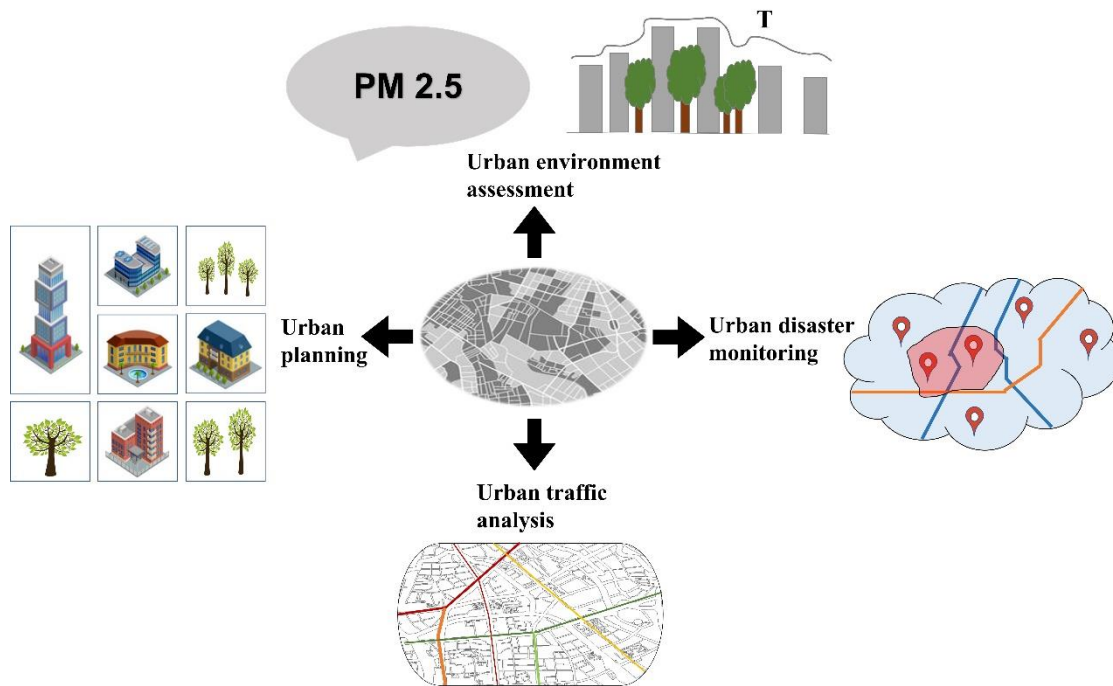


Figure 4-5 Urban land use map integrating RS and GBD for urban management.

A more scientific and efficient urban planning system will benefit from the urban land use map integrating RS and GBD. For example, [Xing and Meng \(2018\)](#) extracted urban land use information by integrating landscape metrics from RS images and semantic features from GBD, which plays as an indicator in urban planning and management. [Chen et al. \(2018\)](#) identified urban green space (e.g., municipal park, community park, etc.) by extracting land cover features from RS data and land use features from POIs for the urban green space planning, which could assist government departments in urban green space planning. In general, integrating RS and GBD would help to improve the city function from the “hard” physical environment and the “soft” services. The addition of GBD for urban land use maps with adequate and timely information could enhance the feedback loops of urban insights for urban governors and planners.

The current advances of urban land use maps that integrate RS data and GBD make it capable of monitoring urban environments such as urban heat island and air pollution

(Masoudi et al., 2021; Halim et al., 2020; Venter et al., 2020; Wang et al., 2021a). For example, [Song et al. \(2019\)](#) analyzed the relationship between urban functional regions and air pollutant emissions and presented a cost-effective way of mapping spatiotemporal patterns of air pollution by utilizing the urban land use map that integrates RS images and POIs. [Luan et al. \(2020\)](#) quantified the impacts of urban natural surfaces and non-surface human activities on urban heat islands by using RS data and GBD, and explored the relationships between urban heat islands and urban land use patterns. This evidence demonstrated that air pollution concentrations are associated with RS-based urban land cover (e.g., industrial layout) as well as GBD-based urban land use (e.g., travel behavior). Other applications have also proved that the spatial patterns of the urban environments have strong relationships with the urban land use patterns ([Pan et al. 2013](#)).

Massive information from urban land use maps that integrate RS and quick-updated GBD is generated continuously and dynamically, providing resources to aid in disaster analysis of historical and future occurrences (earthquakes, fires, or floods) ([Huang et al., 2017](#); [Li et al., 2017c](#)). For instance, [Li et al. \(2019b\)](#) reviewed recent research that utilized RS and GBD for urban disaster information detection including the suffering area, suffering location, and suffering pattern, which provided a useful method for disaster management. Furthermore, [Cervone et al. \(2015\)](#) proposed a novel framework for urban damage assessment under severe weather by using RS images and real-time Twitter data, which were then combined with other GBD in order to get abundant information in disaster areas. Quickly updated urban land use maps that integrate RS and GBD with fine spatial resolution could support urban disaster management by providing unprecedented reference data.

Up to date, urban traffic conditions have become serious problems that threatening human's daily living quality. Several studies have been made on the comprehensive analysis of urban traffic by taking advantage of RS data and GBD on urban land. Applications in traffic quality analysis, for example, classifying Shanghai city into six traffic "source-sink" areas according to the pick-ups and drop-offs of traffic data and LandScan product ([Liu et al. 2012](#)). Improved urban land use maps could also provide technical support for the transportation of the Smart Cities ([Zanella et al. 2014](#)). A thorough perception of urban traffic conditions could be achieved through the integration of RS and GBD.

4.6 Conclusion

This study examined the applications of RS and GBD features in urban land use categorization, as well as methods for RS and GBD integration. The analysis of the existing literature concludes that the emerging GBD provides opportunities for the transformation from urban land cover (physical environment) to urban land use (living environment). Applications on the urban land use maps integrating RS and GBD for urban management mainly include urban planning, urban environment assessment, urban disaster monitoring, and urban traffic analysis. Deeper understandings of the urban surface can be acquired by adding GBD values to the traditional urban RS works. As the integration of RS and GBD has become more generalized, significant progress can be already seen for urban management. Also, integrating RS and GBD on urban land use provides an opportunity for putting people at the center of processes of knowledge and management of the urban planet.



Chapter 5: Decision-level integration and feature-level integration of remote sensing and geospatial big data for urban land use mapping

Stage of publication - Accepted in Remote Sensing (15 April 2021).

5.1 Introduction

China has undergone rapid urbanization since the early 1980s, shown as substantial urban expansion and dynamics in urban land use and structure ([Gong et al. 2019](#)). The timely and accurate urban land use information is important for guiding urban planning and land use management ([Zhu et al. 2019](#)). Remote sensing (RS) techniques were widely used to update urban land use information over the past few decades, by referring to the differences in aspects of texture, spectrum, and context among urban land use categories ([Liu et al. 2018a](#); [Reba and Seto 2020](#)). However, due to the high similarity among urban land use categories in physical attributes, it is hard to identify the complexity and diversity of urban internal structures ([Qi et al. 2019](#)), especially in cities with high-density populations and buildings, such as Hangzhou, Beijing, Shanghai, and Shenzhen ([Cai et al. 2017](#)). Mixed-use land and shadows from high-rise buildings also pose great challenges for the RS data classification ([Zhuo et al. 2019](#)). Big data are defined partly by their large size and their characteristics, in particular their volume, variety, and velocity. Much of this big data are geo-referenced,

or can be geo-referenced, leading to geospatial big data (GBD). The emergence of GBD ([Li et al. 2016](#)), such as mobile phone positioning data ([Ratti et al. 2016](#); [Wu et al. 2020](#)), points-of-interest (POI), social media data ([Ilieva and McPhearson 2018](#); [Yammine et al. 2018](#); [Ye et al. 2016](#)), traffic trajectory data ([Niu et al. 2017](#)), and geotagged photographs ([Cadavid Restrepo et al. 2017](#); [Krylov et al. 2018](#); [Srivastava et al. 2018](#)), provides new opportunities to delineate human dimensions in an urban environment ([Liu et al. 2015](#)). These multi-sourced data contain abundant human activity information, compensating for the lack of socioeconomic attributes of the RS data ([Li et al. 2017b](#)).

The integration of the space-based RS data and the time-based GBD can improve the existing urban land use maps by providing more detailed socioeconomic information and finer spatio-temporal resolution ([Dong et al. 2019](#); [Sarmin and Ismail 2016](#)). Promising progress was made on the applications of integrated RS and GBD on urban land use mapping at different scales and regions ([Goffi et al. 2020](#); [Shi et al. 2019](#); [Zhang et al. 2020c](#); [Zhang et al. 2019](#)). For example, [Liu et al. \(2017\)](#) collected six features (spatial, texture, spatial envelope, rotation-invariant from RS images, Tencent real-time user density, and POI) to identify the urban land use type within each land use parcel. [Jia et al. \(2018\)](#) classified RS data and mobile phone positioning data separately and then fused the two results by using a decision fusion strategy for categorizing urban land use. However, it is still challenging to integrate RS and GBD for urban land use mapping because of the modality gap (i.e., the differences in spatial data quality, technical format, and data structure) and heterogeneity in the data.

In this context, the main goal of this study is to propose the general framework by summarizing the RS and GBD integration approaches used in urban land use mapping

and illustrate the advantages and disadvantages by applying them to map the urban land use situation in Hangzhou city, China. To our knowledge, this is the first study to summarize the existing efforts of integrating RS and GBD for urban land use mapping and analyzing their differences.

5.2 Related Work of RS and GBD Integration Used in Urban Land Use Mapping

Based on the literature review of previous studies, we categorize the various methods of RS and GBD integration used in urban land use mapping into decision-level integration (DI) and feature-level integration (FI) by considering the data integration modes (e.g., integration timing or integration method). A general framework is proposed, as shown in Figure 5-1, to demonstrate and differentiate the two approaches. Briefly, the main difference between DI-based and FI-based urban land use mapping is that DI combines the classification results of the RS and GBD, whereas FI integrates the features extracted from the RS and GBD for classification.

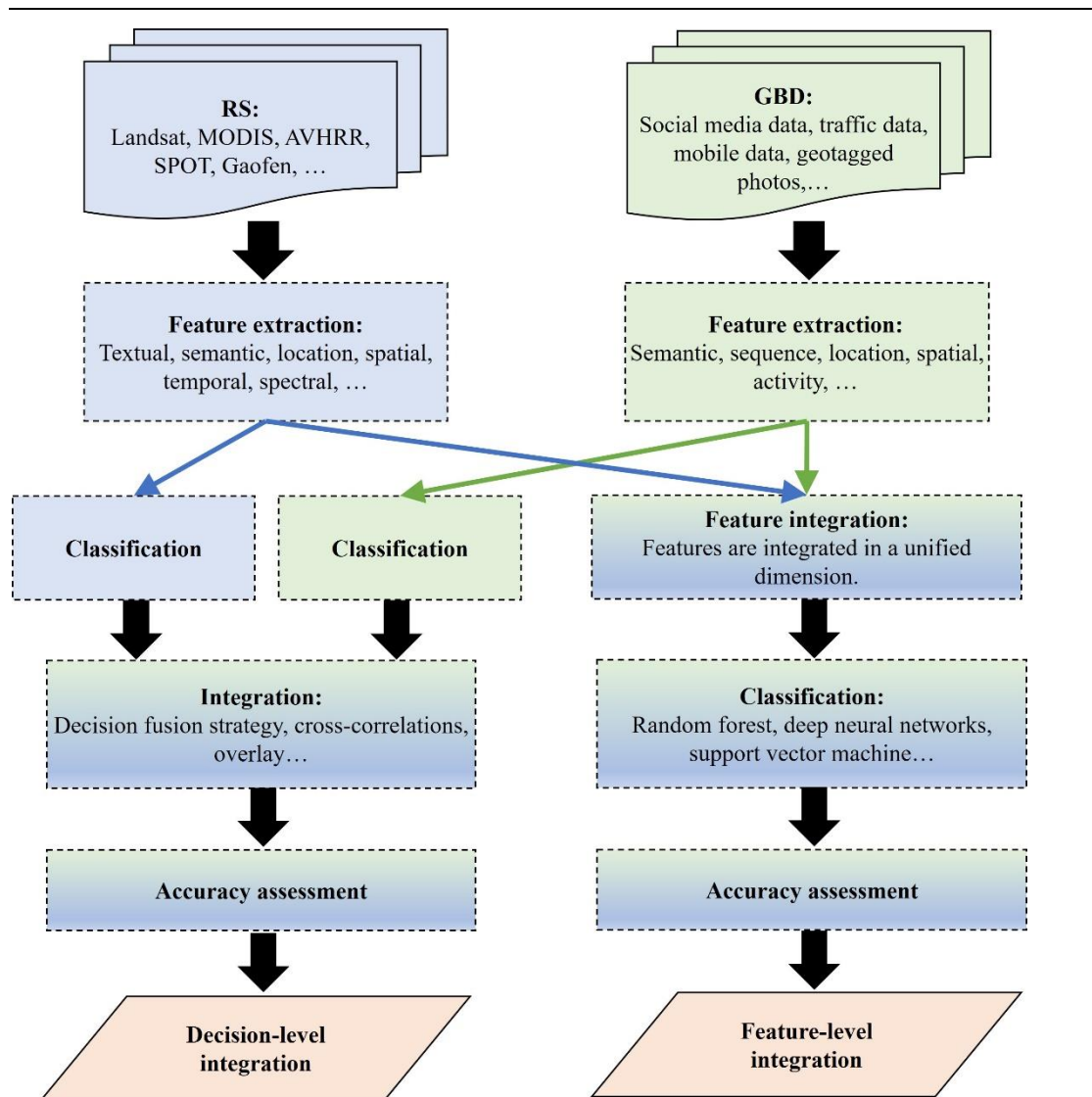


Figure 5-1 A general framework of the two integration strategies: Decision-level integration (DI), and Feature-level integration (FI).

5.2.1 DI-based Urban Land Use Mapping

The DI-based urban land use mapping processes the RS and GBD independently with different models and methods and combines the RS-based and GBD-based results for further generating the urban land use map. The DI-based method was first introduced by [Chang et al. \(2015\)](#) and provides an efficient way for mapping urban land use in Kunming City by integrating POI data and Landsat images. Since then, a series of studies have followed that method. For example, [Song et al. \(2018a\)](#) further

developed the method by fully exploring GBD (Weibo POI, Baidu POI) for more in-depth information and combining them with RS data (Landsat images, Gaofen). [Tu et al. \(2018\)](#) fused landscape metrics (i.e., total class area, patch density, number of patches, and Shannon's diversity index) taken from RS and human activity metrics (i.e., the density of human activities in different functional zones) extracted from mobile phone positioning data for classifying urban functional zones by hierarchical clustering. [Jia et al. \(2018\)](#) integrated RS-based urban land cover maps and GBD-based urban land use maps for mapping urban land use through decision fusion strategy (i.e., certain fusion rules for classifying different urban land use categories). [Zhao et al. \(2019\)](#) generated land cover types by training RS images with the semantic elements derived from OpenStreetMap (OSM) data, then identified each building through semantic classification by using POI. [Xu et al. \(2020\)](#) extracted geographic information from RS and functional distribution from GBD (Gaode POI), then combined them by assigning different weights for urban land use mapping in China. In the DI-based method, the RS and GBD features are calculated and processed separately, avoiding the feature conflicting issues. Furthermore, it has the advantage of easy implementation. In summary, the RS and GBD features reflecting disparate attributes and dimensions are processed through independent methods, then fused by certain models (e.g., cross-correlation, decision fusion strategy).

5.2.2 FI-based Urban Land Use Mapping

In the FI-based method, the features are first extracted from the RS and GBD and are then fused into the integrated feature sets and put into a classifier for urban land use classification. For example, [Zhang et al. \(2017b\)](#) extracted the features from the RS and GBD (i.e., spectral features, texture features, landscape metrics, density, spatial patterns

of POI, and the density of geotagged Weibo posts), based on which urban land use types were determined and fused to train the classifiers for final mapping. [Zhang et al. \(2019\)](#) evaluated the performances of the RS and GBD features and their cross-correlations for classifying urban functions by using a Random Forest (RF) model. [Shi et al. \(2019\)](#) extracted RS features from ZY-3, Landsat8, Sentinel-1A, and GBD (e.g., Wechat user density) for identifying urban land use. [Cao et al. \(2020\)](#) proposed a novel framework that extracted the RS-based and GBD-based features (e.g., temporal features from user-visit data) for recognizing urban land use through deep learning. [Zhang et al. \(2020c\)](#) extracted the RS and GBD features to construct a sparse topic model for identifying representative zones with distinct patterns, then trained these zones to recognize urban functional zones using the deep forest model. Compared to the DI-based approach, the FI-based approach provides a more integrative way of processing the features derived from both RS data and GBD. The RS-based and GBD-based features were fully used for the mining of the urban land use information. However, the modality gap (i.e., differences in data quality, data formats, scales, and timestamps between RS data and GBD) between the RS and GBD features is still a challenging issue.

5.3 Case study

5.3.1 Study Site

Based on the general framework, a detailed methodological framework was designed for testing the two approaches (DI and FI) by using Hangzhou city as a case study. The case study was carried out in Hangzhou city, the capital of Zhejiang Province in China, which is one of the representative cities of urban expansion, population growth, economic development, and land use changes ([Li et al. 2017a](#); [Mao et al. 2020](#)).

We selected the area surrounded by the third ring (i.e., the Hangzhou belt highway) as our study site, which is located in the northeast of Hangzhou city (Figure 5-2c). Besides, the green spaces in the West Lake ecological zone and the croplands in the northeast region were removed to ensure that our study site had the main functional areas related to urban human activities, including institutions, residences, businesses, and open space.

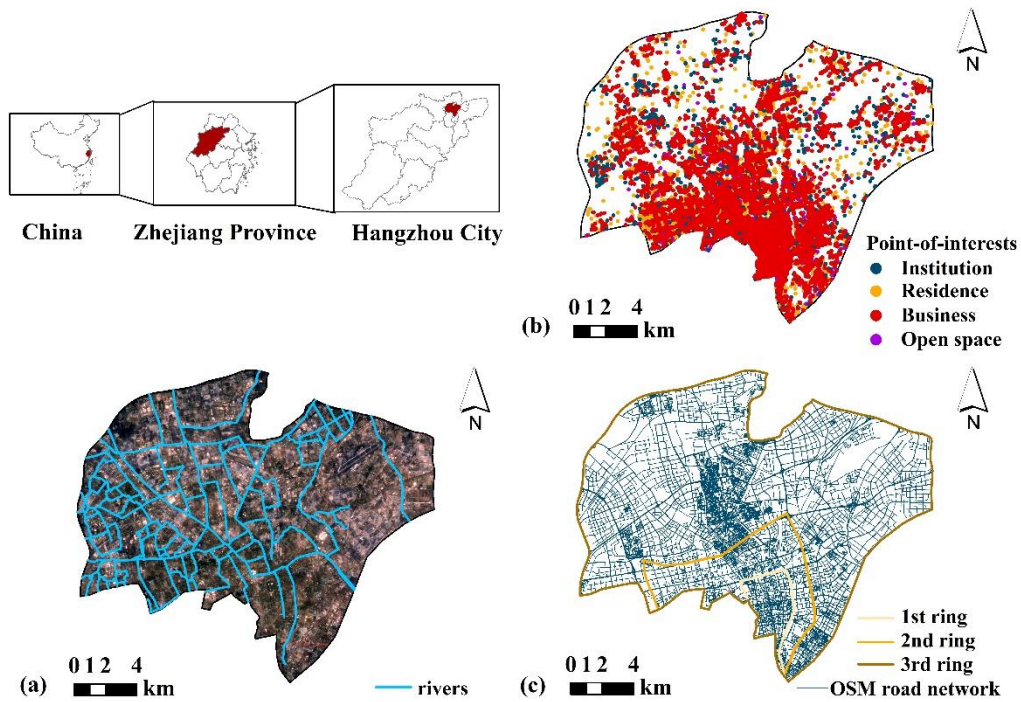


Figure 5-2 The study site (a) Sentinel-2 true-color composition image with OSM river network; (b) Spatial distribution of Point-of-interests (POI); (c) Road network extracted from OpenStreetMap (OSM).

5.3.2 Data Source and Preprocessing

Sentinel-2 is a high-resolution (10 m), multi-spectral, remote-sensing satellite ([Drusch et al. 2012](#); [Zhu et al. 2015](#)). In this paper, the data were integrated from 1 January to 31 December 2019 as mean composites after removing clouds using the Sentinel-2 Quality Assurance (QA) band ([Adiri et al. 2020](#)).

OpenStreetMap (OSM) is a collaborative project providing free editable maps across the world (<https://www.openstreetmap.org>, accessed data:14, December, 2019). The creation and growth of OSM data have mainly been promoted by users and GIS experts ([Liu 2021](#); [Weiss et al. 2018](#)). The high accuracy and free accessibility of OSM in urban areas have been confirmed by many researchers so far ([Helbich et al. 2012](#)). The weekly updated OSM data is very important for capturing the rapid development of urban areas. The OSM road network data are in vector format and contain different classes of roads with different road sizes ([Liu et al. 2020a](#); [Wan et al. 2017](#)). It was proposed as a promising candidate for quick and robust delineations of urban structures and socioeconomic patterns since it is gradually being used to understand Chinese cities ([Liu and Long 2015](#)). Thus, it is reliable to generate urban land use parcels using the OSM road network data.

As a prominent example of the GBD that attracts the most users, POI contain abundant information including land use category, geographic location, and other features (e.g., address, telephone, and postcode) ([Xu et al. 2020](#)). A total number of 28,898 available records of POI within Hangzhou city in 2019 was collected via application programming interfaces (APIs) provided by Gaode Map Services. The urban land use classification system used for aggregating initial POI types were shown in Table 5-1 (<https://lbs.amap.com>, accessed data:14, December, 2019). Moreover, according to previous research, the urban area could be divided into six categories, including institution, residence, commercial land, industrial land, open space, and transportation ([Xu et al. 2020](#); [Zhang et al. 2020c](#)). Since we have excluded the road and river layers for generating urban parcels, there are five categories left. We further combined the commercial and industrial land into business as both have the nature of

the business, and the business level meets the classification requirements of this paper. Therefore, the urban classification system used in this study consists of four categories: institution, residence, business, and open space.

Table 5-1 The framework used for aggregating initial Gaode POI types.

Gaode POI classification	Urban land use classification
Governmental organization	
Medical service	
Finance and Insurance service	Institution
Sports and Recreation	
Culture and Education	
Daily life service	
Commercial house	Residence
Commercial service	
Shopping	
Food and Beverages	Business
Enterprises	
Accommodation service	
Tourist attraction	Open space

5.3.3 Methods

The case study was carried out with the assumption that a parcel divided by urban road networks is homogeneous in terms of urban land use function ([Helbich et al. 2012](#); [Rozenfeld et al. 2011](#)). Quantitatively illustrating the differences between DI-based and FI-based methods requires the following steps: (1) extracting urban land parcels based on the road network and river data (Figure 5-3a) and preparing training and validation parcels (Figure 5-3b); (2) generating urban land use maps by using DI-based (Figure 5-3c) and FI-based methods, respectively, (Figure 5-3d) and evaluating the two based on validation datasets; (3) analyzing the differences between the two methods in terms of

quantitative and qualitative assessment of resulting maps and their data processing modes. The detailed procedures are presented below.

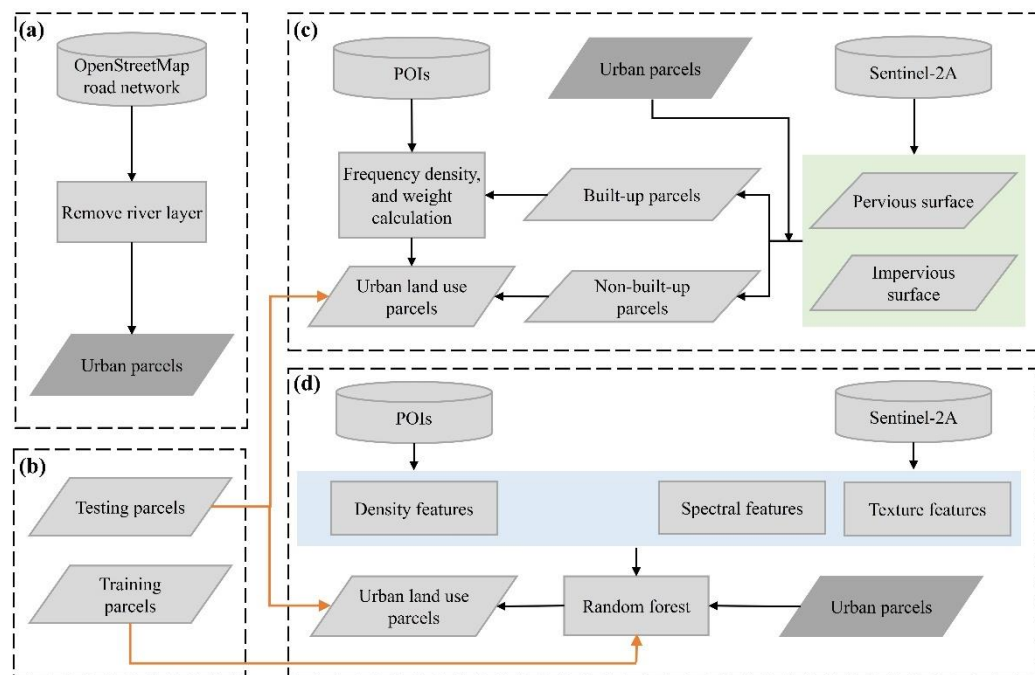


Figure 5-3 This is a figure. Schemes follow the same formatting. Research methodological framework of mapping urban land use with decision-level integration (DI) and feature-level integration (FI). (a) Urban parcel generation based on the road network and river data; (b) Collection of testing and training parcels; (c) DI-based urban land use mapping (DI); (d) FI-based urban land use mapping.

3.3.1. Urban Parcel Generation

The study area was segmented into urban parcels based on the method of the automated identification and characterization of parcels (AICP) ([Liu and Long 2015](#)). Specifically, roads less than 100 m and roads inside communities were considered as redundant information and manually removed. The road network and river layer were then used for generating road buffers and river buffers. The buffer widths of the road and river were determined based on their respective classes according to the Ministry of Housing and Urban-Rural Development (MoHURD) ([Ministry of housing and](#)

[urban-rural development \(MoHURD\) 2012](#)) and the actual situation in Hangzhou city (Table 5-2). The river layer was delineated by visual interpretation based on Google Earth, and the preprocessing of the road network and buffer generation were conducted by using ArcMap 10.2.

Table 5-2 The road and river levels and their corresponding buffer sizes used in this study.

Classes	Road descriptions	Road widths (meters)	River descriptions	River widths (meters)
Level 1	Trunk, primary, railway	40	Main rivers	50
Level 2	Secondary	20	Intermediate rivers	20
Level 3	Tertiary, residential, service, others	10	Small rivers	10

3.3.2. Training and Testing Parcel Collection

We selected parcels for preparing the training and testing parcels, of which the land use types were identified by using the Baidu map, Baidu street view, and a field survey (Figures 5-4 and 5-5). The hybrid satellite map of Baidu map presents a top view of the urban landscape with high-resolution satellite imagery and names of ground features, which were further inspected or confirmed using Baidu street view as it contains high-resolution photos and allows multiple perspectives ([Hoffmann et al. 2019](#); [Kang et al. 2018](#)). An example was presented in Figure 5-4 to illustrate the top view of an institution and residence in the Baidu map and their corresponding Baidu street views. We determined the parcel type while different types of street views were observed for the same parcel, and there is no predominant characteristic referring to the rules in essential urban land use categories in China (EULUC-China) ([Gong et al. 2020](#)).

To be more specific, we recorded a list of items, including the geolocations, classification types, landmark buildings and facilities, and mixed land use situation and their estimated proportions for each urban parcel. Finally, a total of 700 training parcels were identified, including 175 institution parcels, 175 residence parcels, 175 business parcels, and 175 open space parcels. These parcels were selected using stratified random sampling, with an equal sample size (175) for each land use class. Following the approach discussed by [Stehman and Foody \(2019\)](#), we selected 550 parcels for testing. These 550 parcels were selected using simple random sampling from the list of parcels. For each parcel, the land use class was determined by visual interpretation based on the Baidu map. The 550 parcels included 85 institution parcels, 225 residence parcels, 156 business parcels, and 84 open space parcels. The training parcels were used for training the RF classifiers in FI-based classification, while testing parcels were used for accuracy assessment in both DI-based and FI-based classification, respectively.

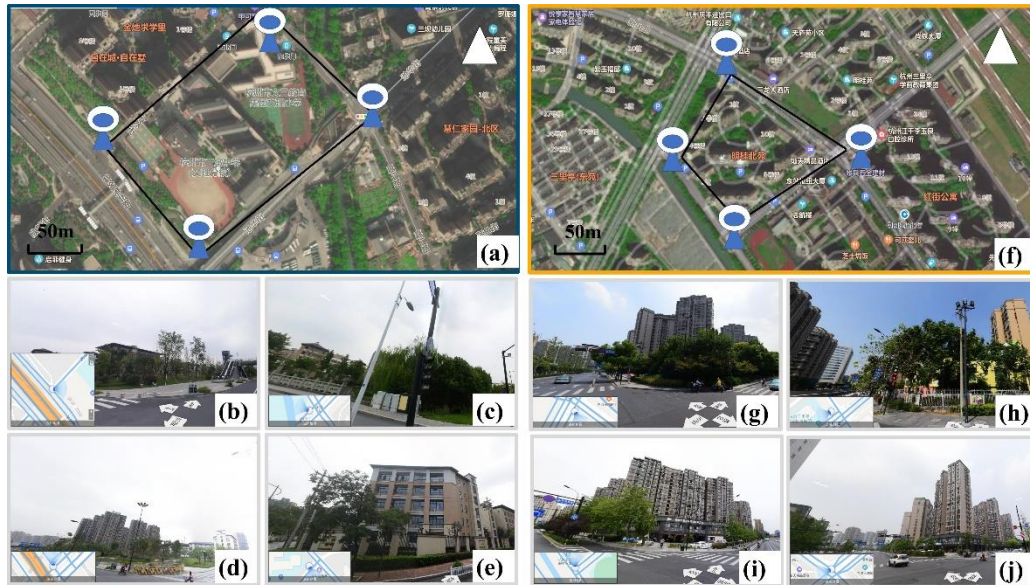


Figure 5-4 Illustration of the institution and residence parcels in Baidu map and Baidu street view: (a) represents the institution parcel in Baidu map; (b), (c), (d) and (e) represent the four perspectives of institution parcel in Baidu street view; (f) represents the residence parcel in Baidu map; and (g), (h), (i) and (j) represent the four perspectives of residence parcel in Baidu street view. The Chinese labels for (a) and (f) represent the name of POIs.

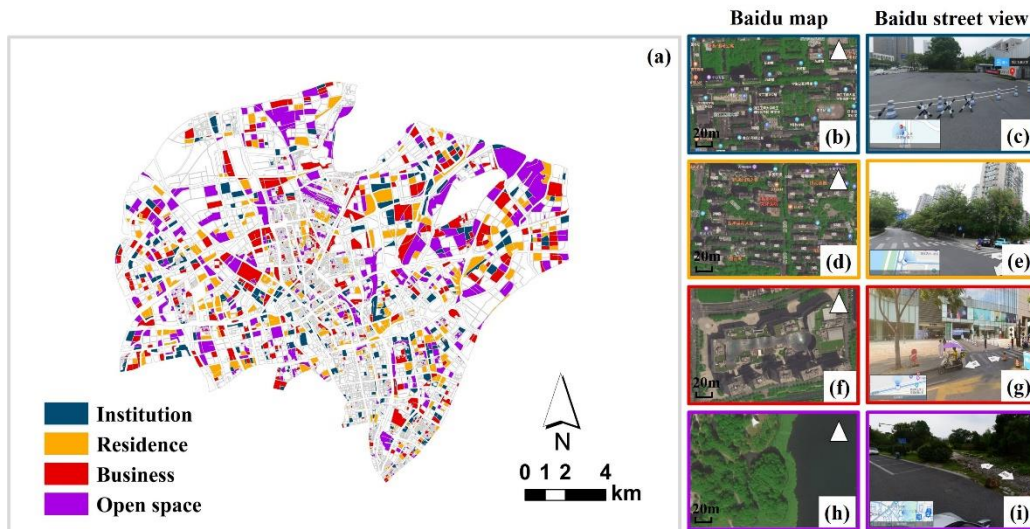


Figure 5-5 Spatial distribution and illustration of training and testing parcels for the four urban land use types. (a) represents the spatial distribution of training and testing parcels; (b), (d), (f), and (h) represents the Baidu map image for institution, residence, business, and open space, respectively; and (c), (e), (g) and (i) represent the Baidu street view for the institution, residence, business, and open space, respectively. The Chinese labels for (b), (d), (f), and (h) represent the name of POIs.

3.3.3. DI-based Urban Land Use Mapping

The DI-based urban land use mapping integrated the results of the RS and GBD classification based on integration rules (Figure 5-3c). First, the Sentinel-2 image was classified for producing the impervious and pervious classes (Figure 5-6a). Specifically, the training samples identified by visual interpretation based on Google Earth and a field survey were put into the RF classifier for the classification of the impervious and pervious surface. The impervious surface is characterized by the artificial structures that are covered by water-resistant materials, while the pervious surface is the opposite (Wang et al. 2015; Weng 2012; Xu 2010). The classification result was then overlaid with the urban parcels to generate the built-up and non-built-up parcels (Figure 5-6b). In this study, urban parcels with impervious surfaces were defined as built-up, otherwise,

the parcels were labeled as non-built-up. This step was carried out under the assumption that built-up parcels require at least partially impervious surfaces (Breiman 2001). The non-built-up parcels were further classified as open space, while built-up parcels were used for further analysis.

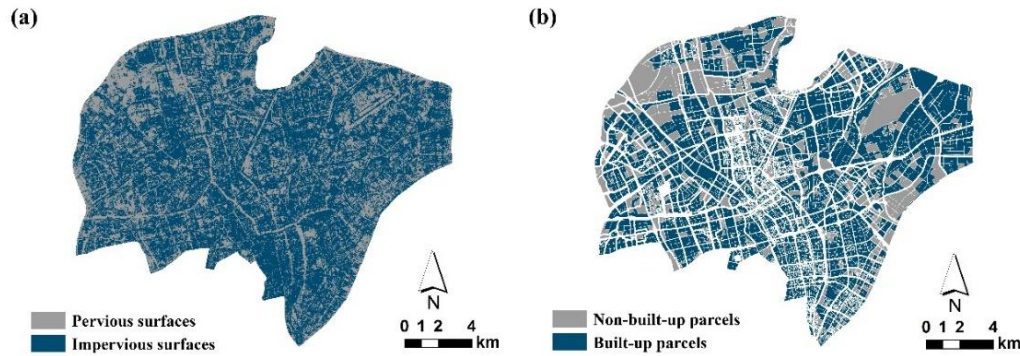


Figure 5-6 (a) Distribution of the pervious and impervious classes; (b) Distribution of the built-up and non-built-up parcels.

Finally, the type of each built-up parcel was determined according to a modified method (Zhao et al. 2011). The method consists of two parts: the frequency density calculation, and the weight calculation. In general, the main type in each parcel was determined after comparing the final frequency density ratios of four types of the POI, and the highest ratio of POI is considered as the final type. The final frequency density ratios were determined by multiplying the initial frequency density ratio and the weight of each types' POI. Specifically, the initial frequency density ratios of the four types' POI were calculated by using Equation 1, Equation 2, and Equation 3.

$$F_i = \frac{n_i}{N_i}, i = 1,2,3,4 \quad (1)$$

where i represents the type of POI, which are institution, residence, business, or open space. n_i demonstrates the quantity of type i in a parcel. N_i represents the total

quantity of type i in all parcels. F_i demonstrates the frequency density which equals the ratio of n_i to N_i .

$$C_i = \frac{F_i}{\sum_1^4 F_i} \times 100\%, i = 1,2,3,4 \quad (2)$$

where C_i represents the initial ratio of type i 's frequency density to the sum of all four types' frequency density.

The initial frequency density ratios of the four types' POI were then respectively multiplied with their weights since various types of POI have different recognition degrees. The weight of institution, residence, business, and open space POI was defined as 2, 1, 1.5, and 2.5, referring to the construction area and the actual situation. The type in each parcel was determined as the one with the highest ratio and the final urban land use map was obtained. The open space parcels produced by the calculation of POI were further combined with the non-built-up parcels as the final open space parcels.

$$D_i = C_i \times W_i \times 100\%, i = 1,2,3,4 \quad (3)$$

where D_i represents the final frequency density ratio which equals the initial ratio of type i 's frequency density multiplying the weight of each types' POI. $W_1 = 2, W_2 = 1, W_3 = 1.5, W_4 = 2.5$.

3.3.4. FI-based Urban Land Use Mapping

Many RS and GBD features were computed for urban land use mapping in the FI-based classification approach. However, a large proportion of these features were inherently highly correlated and redundant because they are all dependent on similar metrics for the classification. This kind of dataset could result in an underperforming and unnecessarily complex classification model. To solve this problem, this research performed a feature selection procedure before the FI-based classification step to construct smaller, more predictive, and parsimonious models. The "Variable Selection

Using Random Forest” (VSURF) algorithm was used for the RS and GBD feature selection and categorized the features into three types: (i) removing useless features, (ii) finding the most predictive set of features which may contain a great amount of redundancy, and (iii) retaining the accuracy while removing redundant features through a stepwise search.

84 RS and GBD features including spectral features, textural features, density features, and spatial features were utilized for the feature selection (Table 5-3). Textural features indicate the adjacency relationships between the gray levels of pixels to characterize the texture of an image ([Puissant et al. 2006](#); [Zhang et al. 2017b](#)). In this study, the textural features (i.e., mean, standard deviation, homogeneity, correlation, angular second moment, contrast, dissimilarity, entropy, and variance) were extracted by computing Gray Level Co-Occurrence Matrix (GLCM) that contain rich information on spatial structure and landscapes ([Lu and Weng 2006](#)). The Red, Blue, Green, and near-infrared (NIR) bands from the Sentinel image were used for computing the spectral and textural features ([Deng et al. 2019](#); [Heiden et al. 2007](#)). Moreover, density features from POIs help to present the differences in spatial patterns among different types of parcels. The Kernel Density Estimation (KDE) tool in ArcMap 10.2 was utilized to generate 4 layers of kernel density of the four land use types of POIs respectively. The spatial features of POIs refer to the POI number of each urban land use type, and the proportion of POI number of each urban land use type to the total number of urban land use types.

Table 5-3. Features selected by feature selection using “VSURF”. These are the remaining features at the “prediction” step.

Feature types	Indices
Spectral features	Mean, Standard deviation, Kurtosis, Skewness, EVI, NDBI, NDVI, NDWI
Textural features	Mean, Standard deviation, Homogeneity, Correlation, Angular second moment, Contrast, Dissimilarity, Entropy, Variance
Density features	Minimum, Maximum, Range, Sum, Mean, Standard deviation
Spatial features	Spatial distribution of POIs

Feature selection was performed on the initial set of features computed and resulted in an impressive reduction of 81.9% (from an initial set of 84 features to 36 remaining features) for the study area (Figure 5-7, Table 5-4). The list of selected features is presented in Table 1. After the feature selection procedure, the RF model was trained using all the above-mentioned selected features 100 times to attenuate the uncertainties and get more stable accuracies. All the operations in the FI-based classification were realized using Google Earth Engine (GEE) platform.

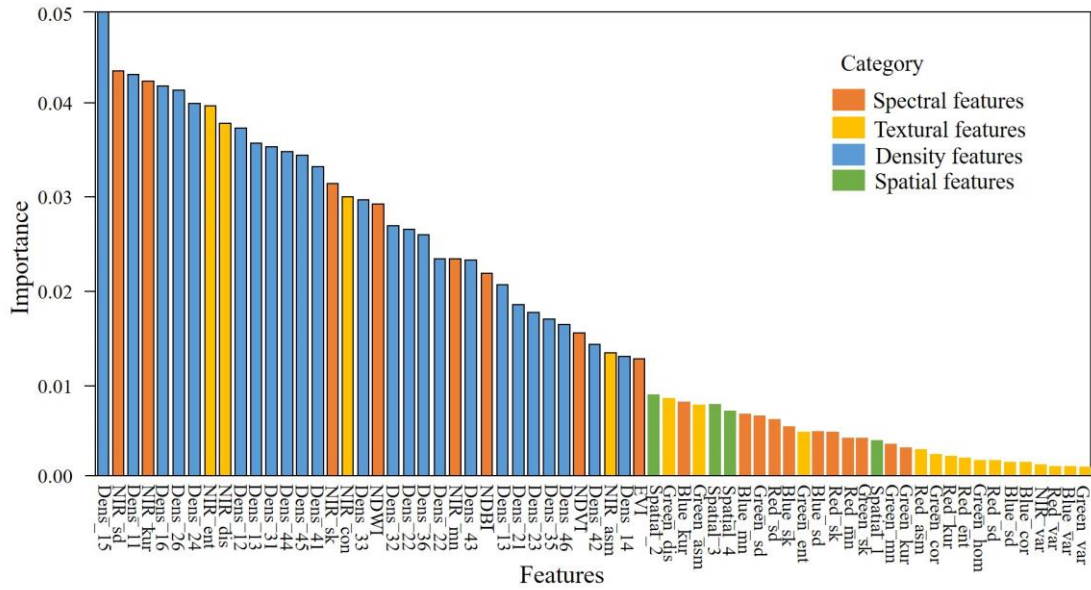


Figure 5-7 Ordered attribute importance for the whole classification.

Table 5-4 Spectral and textural features from RS, and density features from GBD used in FI-based classification.

Feature Types	Indices
Spectral features	Enhanced Vegetation Index (EVI), Normal Difference Built-up Index (NDBI), Normal Difference Vegetation Index (NDVI), Normal Difference Water Index (NDWI), mean_NIR, standard deviation_NIR, kurtosis_NIR, skewness_NIR
Textural features	Angular second moment_NIR, contrast_NIR, dissimilarity_NIR, and entropy_NIR
Density features	Minimum, maximum, range, sum, mean and standard deviation

3.3.5. Analysis of the DI-based and FI-based Classification

To understand the differences between DI-based and FI-based urban land use mapping and describe their advantages and disadvantages, we first evaluated the results derived from the two methods by computing the overall accuracy (OA), user’s accuracy (UA), and producer’s accuracy (PA) based on the confusion matrix (Foody 2002).

Approximately 95% confidence intervals of OA, UA, and PA were also calculated based on the equations proposed by [Card \(1982\)](#) (Supplementary Materials). In the estimation procedure for DI and FI methods, we used a proportion of the number of parcels in each category, relative to all categories in the DI and FI classification results, and denoted this as w_j , following [Olofsson et al. \(2014\)](#), and [Stehman and Foody \(2019\)](#). The column refers to the classified map category while the row refers to the true category based on the reference data (see the Supplementary Materials). For DI, No Data parcels refer to the parcels without POI values that were not involved in the confusion matrix. We then implemented the visual analysis of the DI-based and FI-based urban land use maps. Furthermore, we evaluated the implementation of the two methods by discussing their advantages and disadvantages.

5.4 Results

5.4.1 Quantitative Performance of DI-based and FI-based Urban Land Use Mapping

Table 5-5 showed the confusion matrix of the pervious and impervious map. The estimated OA ($\pm 95\%$ confidence interval) is 0.971 ± 0.007 . The estimated UA and PA of the impervious surface are 0.975 ± 0.008 and 0.986 ± 0.005 . Table 5-6 presents the confusion matrix of DI-based classification. The estimated OA for DI-based urban land use map is 0.635 ± 0.049 . Note that the residence and open space were classified relatively well, with the estimated UA of 0.728 ± 0.069 and 0.714 ± 0.140 , respectively. For institution and business, they have relatively lower UA (0.512 ± 0.140 and 0.522 ± 0.095). Table 5-7 presents the confusion matrix of FI-based classification. The

estimated OA for FI-based classification results is 0.569 ± 0.041 . It can be noted that the residence and open space in FI-based classification can be classified relatively well, with the UA of 0.631 ± 0.059 and 0.640 ± 0.111 , respectively.

Table 5-5 Confusion matrix of the impervious and pervious surface extraction results. UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map.

	Map Category					
	Class	Pervious	Impervious	Total PA	OA	
True Category	Pervious	0.240	0.018	0.258	0.928 ± 0.022	0.971 ± 0.007
	Impervious	0.010	0.732	0.742	0.986 ± 0.005	
	Total(π_j)	0.250	0.750	1.000		
	UA	0.958 ± 0.018	0.975 ± 0.008			

Table 5-6 Confusion matrix of DI-based classification results (I: Institution; R: Residence; B: Business; O: Open Space). UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map.

	Map category					Total PA	OA	
	Class	I	R	B	O			
True Category	I	0.071	0.029	0.042	0.014	0.156	0.453 ± 0.107	0.635 ± 0.049
	R	0.015	0.326	0.072	0.008	0.421	0.774 ± 0.055	
	B	0.034	0.065	0.157	0.010	0.265	0.591 ± 0.082	
	O	0.019	0.029	0.029	0.081	0.158	0.514 ± 0.102	
	Total(π_j)	0.138	0.448	0.300	0.113	1.000		
	UA	0.512 ± 0.140	0.728 ± 0.069	0.522 ± 0.095	0.714 ± 0.140			

Table 5-7 Confusion matrix of FI-based classification results (I: Institution; R: Residence; B: Business; O: Open Space). UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map.

		Map category				Total	PA	OA
		Class	I	R	B			
True Category	I	0.071	0.030	0.025	0.016	0.142	0.496±0.092	0.569±0.041
	R	0.020	0.291	0.108	0.019	0.439	0.663±0.045	
	B	0.036	0.112	0.124	0.012	0.284	0.436±0.062	
	O	0.008	0.028	0.015	0.084	0.135	0.620±0.094	
	Total(π_j)	0.135	0.462	0.272	0.130	1.000		
UA		0.523±0.114	0.631±0.059	0.453±0.080	0.640±0.111			

5.4.2 Qualitative Performance of DI-based and FI-based Urban Land Use

Mapping

Figure 5-8 presents the urban land use maps derived from DI-based and FI-based methods in Hangzhou city. Overall, there is a difference in the number of each urban land use type, and the spatial distribution of the four land use types is not similar. In the DI-based map, the residence parcels were mainly distributed in the city center, while the open space parcels were distributed around Hangzhou city. The distribution of urban land uses in the DI-based map was visually consistent with the actual urban land use referring to the Baidu map. In the FI-based resulting map, the distribution of each land use type was more widely spread out across the study area than that in the DI resulting map. No-value parcels only exist in the resultant map from the DI-based classification, where the built-up parcels determined from the RS images are further labeled based on POI (see Section 5.3.3). No-value parcels arise where there is no POI data within the built-up parcel to support this labeling. On the contrary, the FI-based classification

integrates features from both RS and GBD before land use classification could be classified, as each parcel contains features at least from RS data.

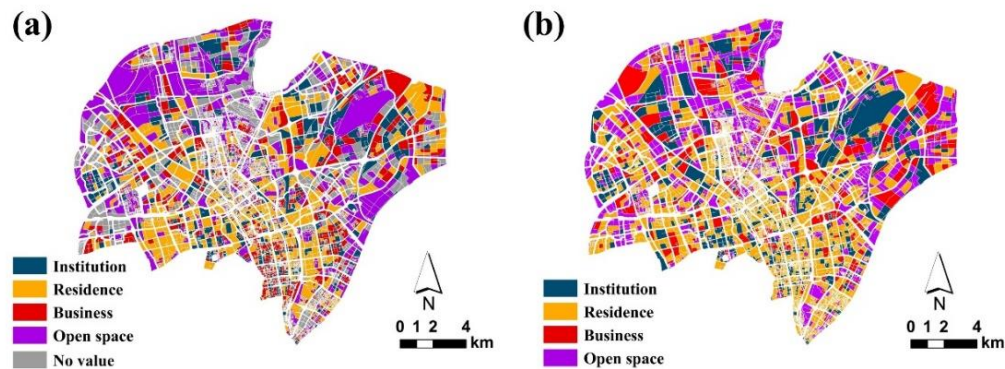


Figure 5-8 Illustration of the (a) DI-based and (b) FI-based urban land use maps in Hangzhou city.

Figure 5-9 presents a visual comparison of the urban land use maps derived from DI-based and FI-based methods and the corresponding Baidu map in four subareas. The four subareas A, B, C, and D are dominated by institution, residence, business and opens space, respectively. Overall, the visual comparison indicates a generally better performance of the DI-based classification than the FI-based classification. For example, some institution parcels were classified as residence for the FI-based classification in subarea A. In subarea B, some residence parcels were misclassified as business by the FI-based classification. The FI-based result map misclassified business parcels as residence parcels in subarea C. For subarea D, the open space parcel was wrongly classified as a residence parcel.

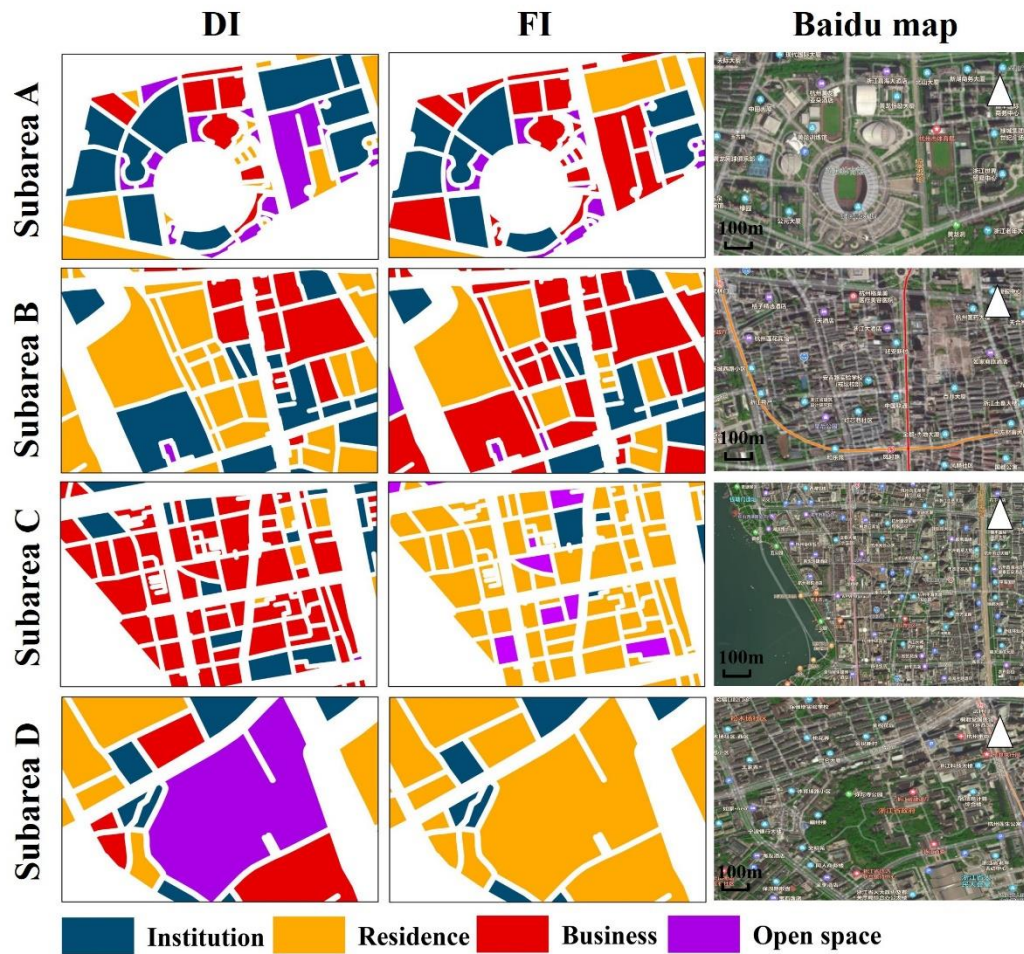


Figure 5-9 Classification results of DI and FI methods and testing samples in the test regions. Subareas A, B, C, and D were selected randomly in Hangzhou city, and each region was shown with DI-based classification, FI-based classification, and Baidu map. The Chinese labels for the subfigure of right column represent the name of POIs.

5.5 Discussion

This study implemented the DI-based and FI-based methods for mapping urban land use types in Hangzhou city, based on the same urban land use classification system (Table 5-1), the same data sources (i.e., Sentinel-2A, OSM road networks, and POI), and the same training and validation urban parcels (Figure 5-5). In our study, the only difference between these two approaches is the manner and timing for integrating the

RS features and GBD features (Figure 5-1 and Figure 5-3). It should be noted that the objective of this study is not to quantitatively compare the performance of the two methods but to illustrate their advantages and disadvantages by applying them to the mapping of the urban land use situation in Hangzhou city. The results in this study, which show that DI-based classification performs better than FI-based classification, may not apply to other circumstances ([Xu et al. 2020](#); [Yu et al. 2019](#)). Given that our study area is located in eastern China, where the diversity of urban land use types has a higher level of complexity than other regions in China, our mapping result is probably not representative of other regions. It should be noted that different data sources, selected features, classifiers, training samples, or land use types may also lead to different results. According to the process of the integration methods, these differences can be summarized into four types: 1) different urban road networks in different regions might lead to different urban parcel unit; 2) The variety of GBD in different regions lead to different mapping results; 3) The number of mixed parcels in different regions varies depending on the tiers of cities; 4) The availability of RS data greatly varies in different cities due to the coverage of clouds and cloud shadows that might be another factor affecting the application of the proposed methods.

5.5.1 Summary of the Advantages and Disadvantages of the Two Methods

In our study, for the DI-based method, the pervious and impervious surfaces map was overlaid with urban parcels for generating built-up and non-built-up parcels, and then they were classified into the four urban land use types by calculating the frequency density and weight of the four types of POI. For the FI-based classification, the RS and GBD features were combined to train the RF model for urban land use mapping based

on the training parcels. Some advantages and disadvantages could be identified through the analysis of classification results and the above-mentioned process:

DI-based urban land use mapping is easy to implement, avoids feature integration, and accompanies conflicting issues. However, it depends largely on the quality and quantity of the GBD in each urban parcel (it depended on POI data in our case study) which might cause the missing value and misclassification for some urban parcels (Figure 5-8 and Figure 5-9). For example, [Zhao et al. \(2019\)](#) indicated that inaccurate POI will produce incorrect labels, as the classification results are directly generated from POI. Furthermore, the classes used in this paper might not match well with the POI classes considering that DI-based classification is based on labeling parcels based on POI. Reclassifying POI classes according to the nomenclature of land use types can result in some uncertainties of urban land use classification. For example, certain POI could be associated with more than one type of urban land use;

The FI-based land use mapping enables the mixture of features from RS and GBD, however, the implementation has challenges due to the modality gap ([Cao et al. 2020](#); [Liu et al. 2015](#)) between the RS and GBD, such as the spatial data quality, technical format, and data structure. Moreover, both feature selection and feature integration in the FI-based classification can contribute to different mapping results. In this study, spectral features and textural features derived from the Sentinel-2 image were integrated with the density features derived from POI for mapping urban land use (Table 5-3). Accordingly, these features have multiple backgrounds and thus can have various understandings of urban land use mapping, leading to different classification results ([Su et al. 2020](#)). The performance of the FI-based classification is probably related to the complexity of urban parcels. To be more specific, a single urban parcel can comprise

different urban land uses, such as office buildings, residential buildings, and shopping centers. Therefore, the differences in features among urban parcels might be hard to distinguish.

5.5.2 The improved Method for Urban Land Use Mapping

Given the framework (Figure 5-1), quantitative and qualitative assessments of the two methods, it is possible to highlight some points for improving the DI-based and FI-based urban land use mapping. First, very-high-resolution (VHR) satellite images provide abundant information about geographical objects in terms of geometrical features and spatial patterns ([Zhao et al. 2019](#)), which can provide more detailed RS information (e.g., textual, contextual, and spectral information) ([Zhang et al. 2017b](#)). Second, the generation of urban parcels relied on the OSM road network, which contains unbalanced classes that are unevenly distributed, leading to heterogeneous urban land use parcels in terms of parcel area and spatial pattern ([Johnson and Iizuka 2016](#)). The accuracy of the urban parcels created based on OSM can be further improved by combining different data sources, and the veracity of urban parcels should be validated. Third, different GBD features have different contributions to the classification of different land use types according to [Zhang et al. \(2017b\)](#). Evaluating the optimal size of the GBD set is a topic for future research. Fourth, the training and testing parcels used in this study were identified through the Baidu map, Baidu street view, and a field survey, which are promising to validate urban land use parcels ([Hu et al. 2016](#)). However, it is inevitable for parcels to contain multiple land use types. A better classification performance would be expected from the adoption of an independent land use map from the local government and urban planners ([Pan et al. 2020](#)). Fifth, the RF model was used for mapping built-up in DI-based classification

and training urban land use parcels in the FI-based classification. However, we need to improve the RF model with transfer learning capability for mapping the urban land use of any region with high accuracy and a different data source. To be more specific, the transfer learning model can be used as a pre-trained model, which might help increase the classification accuracy. Furthermore, how to integrate GBD into the RF model is a worthwhile question to discuss, since there is a significant difference between RS and GBD (Mao et al. 2020). Sixth, the two methods can be used together to improve urban land use maps as they have different performances in classifying different types of land use. For example, the FI-based method can be used for classifying very different urban land use types, such as open space land and industrial land, while the DI-based method can be utilized for classifying urban land use types with very little difference in the RS features but have a significant difference in GBD features, such as commercial land and industrial land.

5.6 Conclusions

With the increase in RS and GBD, integrating the two types of data was widely used in urban land use mapping in recent years as it provides an opportunity to characterize both physical and socioeconomic attributes of urban land. For the first time, this study summarized the diverse methods of RS and GBD integration used in essential urban land use mapping and categorized them into two types, including the DI-based and FI-based methods. A general framework was proposed for explaining the main differences in the process of the two methods. Taking Hangzhou city, China as a case study, we then applied the two methods to produce the urban land use maps consisting of the institution, residence, business, and open space, based on the OSM road network,

Sentinel-2 data, and Gaode POI, and highlighted the differences in classification results and the process of the DI-based and FI-based methods. On this basis, future studies should pay more attention to the integration methods, and further apply the methods to more specific classification-type scenarios, data sources, and other regions.

Chapter 6: Quantifying spatiotemporal patterns of urban land use change: A case study in Hangzhou, China

Stage of publication - First draft prepared. Awaiting feedback from co-authors before submission to the journal Cities.

6.1 Introduction

Since the early 2000s, China has been experiencing the most dramatic transformations of urbanization ([Gong et al. 2019](#)), especially in cities with high-density populations such as Beijing, Shanghai, Guangzhou, Shenzhen, and Hangzhou ([Cai et al. 2017](#)). Rapid urbanization not only alters urban land cover structures (e.g., physical attributes of urban environment) but also change urban land use patterns significantly (e.g., socio-economic attributes) ([Yao et al. 2022](#)). The changing patterns of physical and socio-economic aspects of the urban planet further impact the urban environment as well as the physical and psychological health of urban dwellers ([Gong et al. 2019](#); [Wu et al. 2021a](#)). Previous studies have demonstrated the spatial and temporal patterns of the urban land cover process (e.g., impervious surfaces, green space, water bodies, and bare land) at different scales ([Pan et al. 2020](#); [Zhang et al. 2020a](#)). However, urban land use change and the underlying socioeconomic process are still less understood.

With the increasing population and expanding urban areas, the complexity and diversity of inner-urban functions have changed significantly at spatial and temporal scales ([Pan et al. 2020](#); [Zhang et al. 2020a](#)). Characterizing the changing patterns of urban land use is thus critical, given that urban land use plays an important role in citizens' welfare ([Tu et al. 2021](#); [Yao et al. 2022](#)). Many of the related studies used municipal cadastral data or urban land use data from the urban planning bureau to understand the urban land use dynamics ([Woldesemayat and Genovese 2021](#)), which might introduce large errors due to the low quality and uncertainty of the urban land use data. Recently, the integration of remote sensing (RS) data and geospatial big data (GBD) provide new opportunities for mapping urban land use (e.g., institutional land, commercial land, and residential land) ([Dong et al. 2019](#); [Sarmin and Ismail 2016](#)). Promising progress has been made in the applications of integrated RS and GBD on urban land use mapping at different scales and regions ([Goffi et al. 2020](#); [Shi et al. 2019](#); [Zhang et al. 2020c](#); [Zhang et al. 2019](#)). While accurate and reliable urban land use change product (e.g., the transformation from institutional land to business land) is still limited.

In this context, the main purpose of this paper is to investigate the spatial and temporal patterns of urban land use changes by integrating GBD and RS data. Taking Hangzhou city as the study site, we first classified the urban land use maps in 2017 and 2021 using Sentinel-2 images, OpenStreetMap (OSM) road network data, and Points of Interest (POI) data. The obtained maps were then used to analyze the changing patterns of urban land use from 2017 to 2021. This paper is organized into five sections. In section 6.2, the study site, data source, and methods used in the research were presented. In Section 6.3, this research classified urban land use maps in Hangzhou for

2017 and 2021, and understand the urban land use dynamics from 2017 to 2021. Section 6.4 discussed the issues related to urban land use changes. Section 6.5 concluded the main findings and implications.

6.2 Materials and methods

6.2.1 Study site

Hangzhou is located in the northwestern part of Zhejiang province, which plays an important role in the Yangtze River Delta Urban Agglomerations. It comprises 10 districts, 1 county-level city, and 2 counties. In this paper, we only focused on the most central urban districts, which are Shangcheng District, Gongshu District, and part of Xihu District. The population in this region accounts for 45% of the total population in Hangzhou city. Hangzhou city is one of the representative cities of urbanization, population growth, and economic development in China ([Li et al. 2017a](#); [Mao et al. 2020](#)).

In addition, this research selected the area surrounded by the third ring (i.e., the Hangzhou belt highway) as the study site (Figure 6-1). The urban areas can be divided into three units including 1st ring belt, 2nd ring belt, and 3rd ring belt. To be more specific, the 1st ring represents Huancheng North Road, Huancheng East Road, Jiangcheng Road, Zhongshan South Road, Wansongling Road, Nanshan Road, Hubin Road, and Huancheng West Road. The 2nd ring refers to Desheng Road, Qiutao Road, Fuxing Road, Old Fuxing Road, Hupao Road, Manjuelong Road, Wulaofeng Tunnel, Jiqingshan Tunnel, Jiulisong Tunnel, Lingxi Tunnel, Zijinghua Road, and Wenyi Road. The 3rd ring is the Hangzhou belt highway.

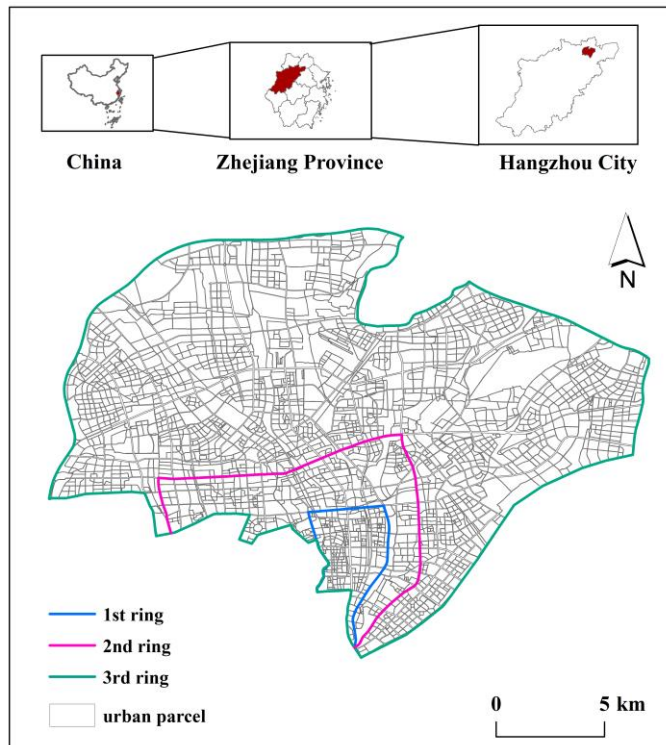


Figure 6-1 The study site.

6.2.2 Data

Sentinel-2 carries an optical instrument payload that includes 13 spectral bands (e.g., 4 bands at 10m, 6 at 20m, and 3 at 60m) ([Drusch et al. 2012](#); [Zhu et al. 2015](#)). The Sentinel-2A Level 1C dataset has been processed with radiometric and geometric corrections, including orthogonal rectification and spatial registration on a global reference system, which is available from the Google Earth Engine (GEE) cloud computing platform. This research selected Sentinel-2A level-1C images, from 1 January 2017 to 31 December 2017 and from 1 January 2021 to 31 August 2021 for urban land use mapping.

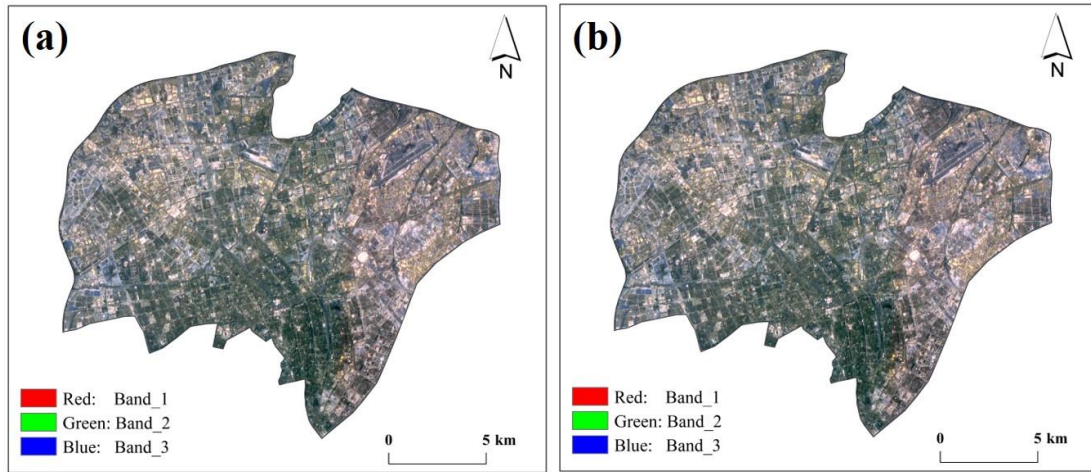


Figure 6-2 Sentinel-2 images. (a) 2017 Sentinel image; (b) 2021 Sentinel image.

OpenStreetMap (OSM) has started as a road network, which was depicted by users voluntarily (<http://www.openstreetmap.org>). It's a dataset composed of vectors, and it mainly includes point, polygon, and polyline formats. Recently, the OSM road network has been utilized as a promising dataset for capturing urban functional patterns ([Liu and Long 2015](#)). This research manually checked the OSM road network data to guarantee the accuracy of the mapping results.

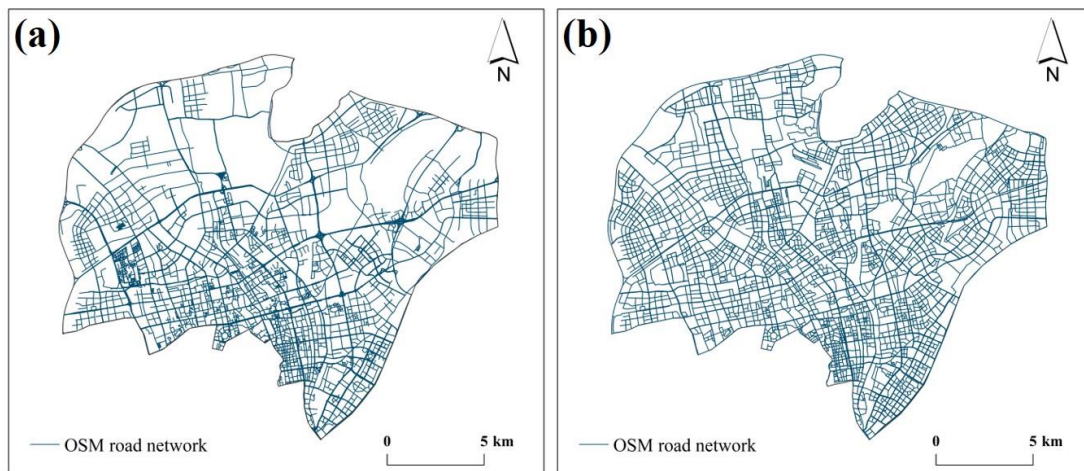


Figure 6-3 OpenStreetMap road network data. (a) 2017 OSM data; 2021 OSM data.

Points of interest (POIs) are the basis for most of the data supporting location-based applications. POIs contain abundant information including land use category,

geographic location, and other features (e.g. address, telephone, and postcode) ([Xu et al. 2020](#)). This research utilized POIs in 2017 and 2021 for mapping urban land use maps in Hangzhou city. The categories of POIs were further combined into four urban land use types (e.g., Institution, Residence, Business, and Open space) according to [Yin et al. 2021b](#)'s work.

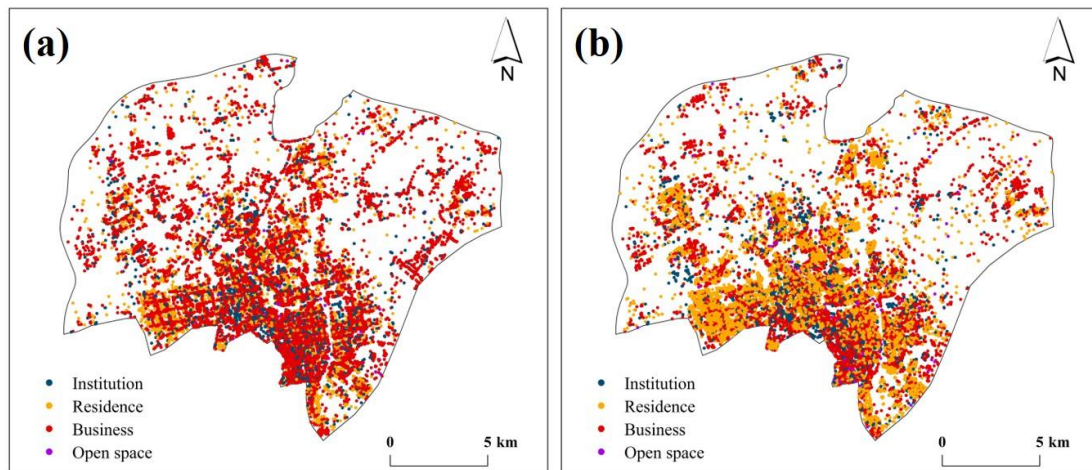


Figure 6-4 Spatial distribution of Points-of-interest (POIs). (a) 2017 POIs; (b) 2021 POIs.

6.2.3 Methods

A methodology framework was proposed referring to [Yin et al. 2021b](#)'s work to map urban land use in 2017 and 2021 (Figure 6-5). However, new data sources have been emerged and used in urban land use classification. In order to improve the classification performance of the existing urban land use maps, an importance analysis of parcel attributes was conducted.

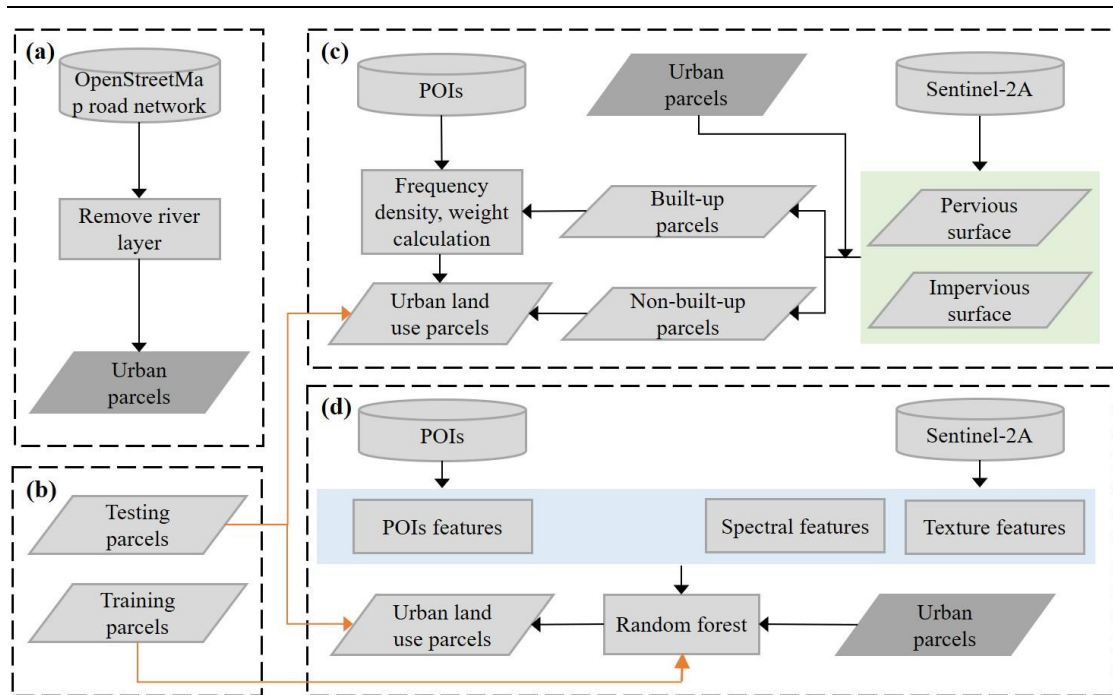


Figure 6-5 Methodology framework. (a) Urban parcel generation; (b) Accuracy assessment; (c) DI-based classification; (d) FI-based classification.

(1) Methodology framework

The urban areas were first segmented into several urban parcels by using OSM road network data based on the method of the automated identification and characterization of parcels (AICP) (Liu and Long 2015). It should be noted that the urban parcels used in 2017 were the same as the urban parcels in 2021 in order to unify the basic units. The reason for this is because the 2021 OSM road network was developed based on the 2017 OSM road network according to the OSM operating system. To be more specific, The road network after 2017 is gradually added on the basis of the 2017 road network. In addition, the road network in many areas was missing due to technical reasons In 2017. Therefore, using a more refined road network in 2021 is more conducive to extracting heterogeneous urban parcels. Based on the urban parcels generated from the OSM road network, the urban land use was then classified by using both DI and FI methods (shown in Figure 6-5).

The FI-based classification method proposed by [Yin et al. \(2021b\)](#) was used for classifying urban land use. However, new RS and GBD features have emerged and applied in urban land use classification. In order to evaluate the performance of different RS and GBD features in urban land use mapping, an importance analysis of parcel attributes was conducted for the improvement of classification performance. Four categories of parcel attributes were evaluated, including spectral features, textural features, density features, and frequency features (Table 6-1). The random forests (RF) model was built to fulfill this purpose, with the dependent variables being respectively the four urban land use types. Specifically, the importance evaluation of parcel attributes was conducted by generating the rankings by using the “Variable Selection Using Random Forests for interpretation” (VSURF) approach proposed by [Genuer et al. \(2015\)](#). RF models were built using different combinations of categories, resulting in two models using three categories. Each model was run 50 times to attenuate the uncertainties associated with RF models so that more stable accuracies can be derived from the error matrix.

Table 6-1 Summary of RS and GBD features used in parcel-level urban land use mapping.

Data source	Features	Variables
Sentinel-2A/B	Spectral features	Mean, standard deviation, kurtosis, skewness of the near-infrared band EVI, NDBI, NDVI, NDWI
	Textural features	Angular second moment, contrast, dissimilarity, and entropy based on GLCM of the near-infrared band
POIs	Density features	Minimum, maximum, range, sum, mean and standard deviation
	Frequency features	Total number of all POIs Total number of each type of POIs The proportion of each type of POIs

(2) Training and testing parcel collection

This paper selected the training and testing parcels by using the Baidu map, Baidu Street View, and a field survey. Following the approach discussed by [Stehman and Foody \(2019\)](#), this research randomly selected 800 parcels in 2021 for training and testing from the list of parcels (Figure 6-6, 6-7). A total of 300 training parcels were identified including 23 institution parcels, 148 residence parcels, 85 business parcels, and 44 open space parcels. The 500 testing parcels included 55 institution parcels, 235 residence parcels, 134 business parcels, and 76 open space parcels.

Furthermore, 300 training parcels in 2017 were identified including 24 institution parcels, 134 residence parcels, 85 business parcels, and 57 open space parcels. The 500

testing parcels in 2017 included 36 institution parcels, 238 residence parcels, 142 business parcels, and 84 open space parcels. Specifically, the testing and training parcels in 2021 were identified by using the Baidu map, Baidu Street View, and a field survey, while the parcels in 2017 were identified through Google Earth and Baidu map. The training parcels were used for training RF classifiers in the FI-based method, while testing parcels were used for accuracy assessment in both DI-based and FI-based approaches, respectively.

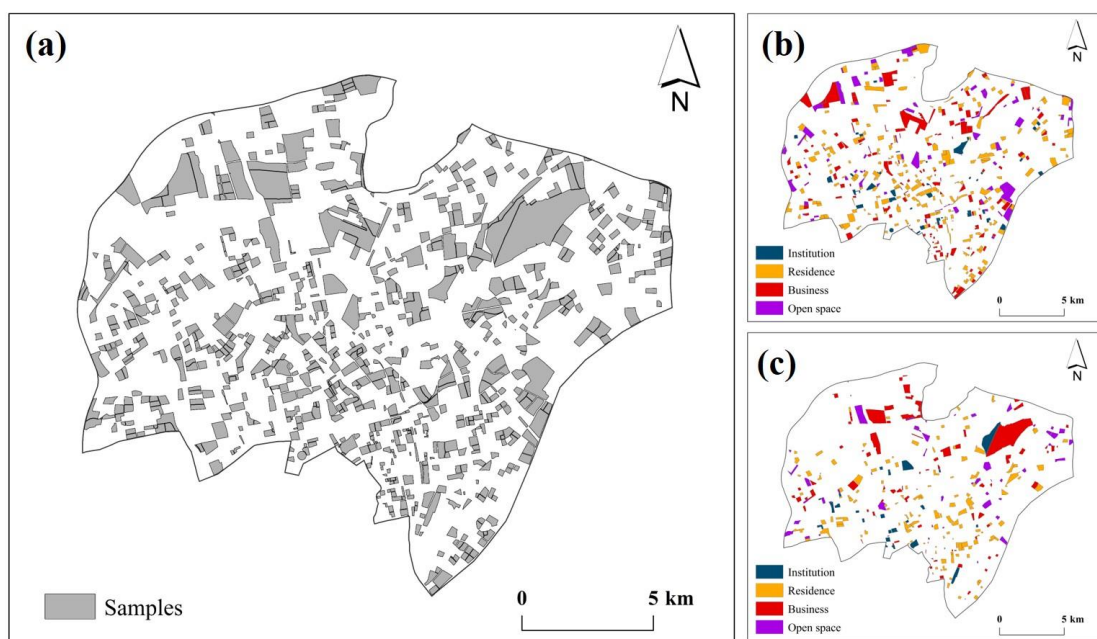


Figure 6-6 Spatial distribution of training and testing parcels in 2017. (a) Spatial distribution of parcels selected randomly; (b) Testing parcels; (c) Training parcels.

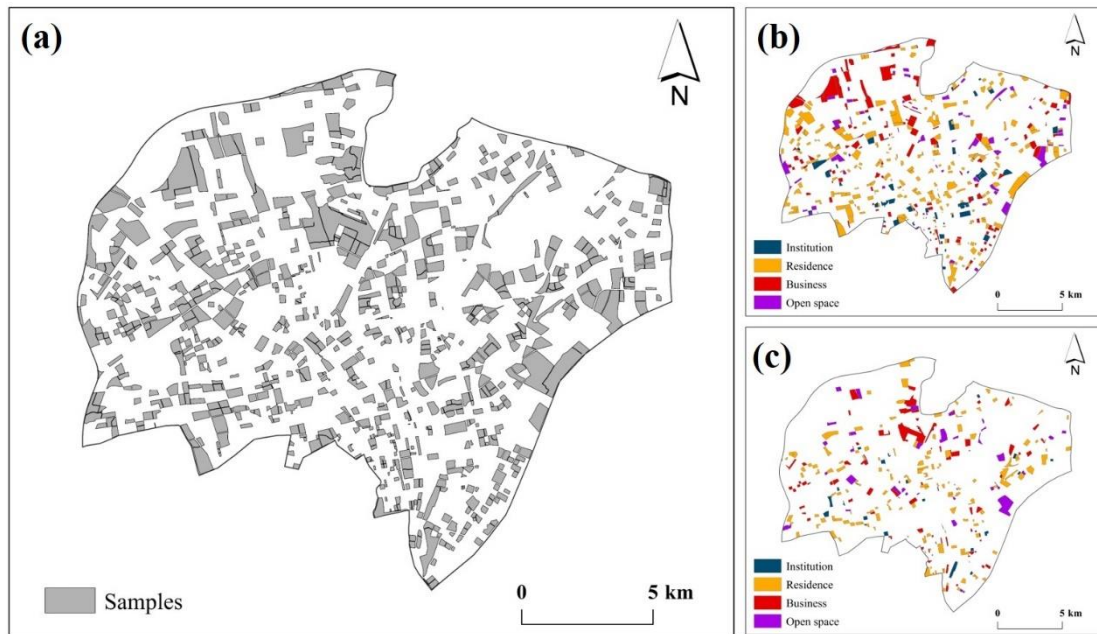


Figure 6-7 Spatial distribution of training and testing parcels in 2021. (a) Spatial distribution of parcels selected randomly; (b) Testing parcels; (c) Training parcels.

(3) Accuracy assessment

In order to evaluate the results derived from the two methods (e.g., DI and FI), this research calculated the confusion matrix of the results with an approximate 95% confidence interval ([Card 1982](#); [Foody 2002](#)). However, it should be noticed that the area proportion in parcel-based urban land use classification cannot represent the proportion of urban land use types. Therefore, we utilized the quantity proportion of urban parcels instead of the area proportion in the confusion matrix. The map category is in the column and the true category is in the row referring to the contingency table (Table 6-2).

Table 6-2 Contingency table for accuracy assessment

		Map category(j)						Total
		1	2	.	.	.	r	
True category(i)	1	n11	n12				n1r	n1.
	2	n21	n22				n2r	n2.
	.							
	.							
	.							
	r	nr1	nr2				nrr	nr.
Total	n.1	n.2				n.r	n	

Marginal distributions of map category, which is the area in category j according to the map. However, the mapping in this research is done on parcel units instead of the pixel unit. This research, therefore, used the ratio between the number of each category N_j and the total number of parcels N according to the resultant land use maps as π_j for the DI and FI classification ([Olofsson et al. 2014](#); [Stehman and Foody 2019](#)). While for impervious surface mapping, the π_j is calculated based on the area in category j according to the resultant map. Some parcels are missing in the DI based classification result because there is no POI data within the built-up parcel to support this labeling.

6.3 Results

6.3.1 Urban land use map in 2021

Table 6-3 showed the confusion matrix of the pervious and impervious maps. The estimated OA ($\pm 95\%$ confidence interval) is 0.903 ± 0.033 . The estimated UA and PA of the impervious surface are 0.944 ± 0.031 and 0.917 ± 0.032 , which are both over 90%. It

can be noted that the impervious and pervious surfaces can be classified relatively well. Table 6-4 presents the confusion matrix of DI-based classification. The estimated OA for the DI-based urban land use map is 0.670 ± 0.039 . Note that the residence and business were classified relatively well, with the estimated UA of 0.892 ± 0.048 and 0.709 ± 0.082 , respectively. For the institution and open space parcels, they have relatively lower UA (0.364 ± 0.093 and 0.522 ± 0.125). Table 6-5 presents the confusion matrix of FI-based classification. The estimated OA for FI-based classification results is 0.802 ± 0.034 . It can be noted that the institution, residence, and business in FI-based classification can be classified relatively well, with the UA of 0.826 ± 0.108 , 0.760 ± 0.048 , and 0.960 ± 0.038 , respectively. However, the open space has a relatively lower UA (0.698 ± 0.139).

Table 6-3 Confusion matrix of the impervious extraction results in 2021. UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map.

		Map Category				
		Pervious	Impervious	Total	PA	OA
True category	Pervious	0.258	0.038	0.296	0.871 ± 0.064	0.903 ± 0.033
	Impervious	0.058	0.645	0.704	0.917 ± 0.032	
Total (π_j)		0.317	0.683	1.000		
UA		0.816 ± 0.078	0.944 ± 0.031			

Table 6-4 Confusion matrix of DI-based classification results in 2021(I: Institution; R: Residence; B: Business; O: Open Space). UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map.

		Map category						
	Class	I	R	B	O	Total	PA	OA
True Category	I	0.084	0.007	0.002	0.020	0.113	0.741±0.11	0.670±0.039
	R	0.058	0.326	0.068	0.036	0.488	0.667±0.046	
	B	0.051	0.019	0.188	0.010	0.268	0.702±0.068	
	O	0.037	0.014	0.007	0.072	0.130	0.555±0.108	
	Total (π_j)	0.231	0.365	0.265	0.139	1		
	UA	0.364±0.093	0.892±0.048	0.709±0.082	0.522±0.125			

Table 6-5 Confusion matrix of FI-based classification results in 2021 (I: Institution; R: Residence; B: Business; O: Open Space). UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map.

		Map category						
	Class	I	R	B	O	Total	PA	OA
True Category	I	0.078	0.029	0.000	0.005	0.111	0.697±0 .102	0.802±0 .034
	R	0.002	0.471	0.004	0.006	0.484	0.974±0 .019	
	B	0.008	0.050	0.194	0.014	0.266	0.729±0 .062	
	O	0.006	0.070	0.004	0.058	0.139	0.421±0 .086	
	Total (π_j)	0.094	0.620	0.202	0.084	1.000		
	UA	0.826±0 .108	0.760±0 .048	0.960±0 .038	0.698±0 .139			

Figure 6-8 presented the urban land use maps derived from DI-based (Figure 6-8a) and FI-based (Figure 6-8b) methods in Hangzhou city. Overall, there is a difference in the number of each urban land use type, and the spatial distribution of the four land use types is not similar. In the DI-based map, the residence parcels were randomly distributed in the city, while the open space parcels were distributed around Hangzhou city. As for the institution parcels, they were concentrated in several areas in the city. Large business patches were found mainly around the city, while small patches were mainly distributed in the city center. The distribution of urban land uses in the FI-based map was visually consistent with the actual urban land use referring to the Baidu map. In the FI-based resulting map, the connectivity of each urban land use type is higher

than the DI-based classification result. Especially for the residence parcels, they were mainly distributed in the city center. The business and open space parcels were distributed around Hangzhou city.

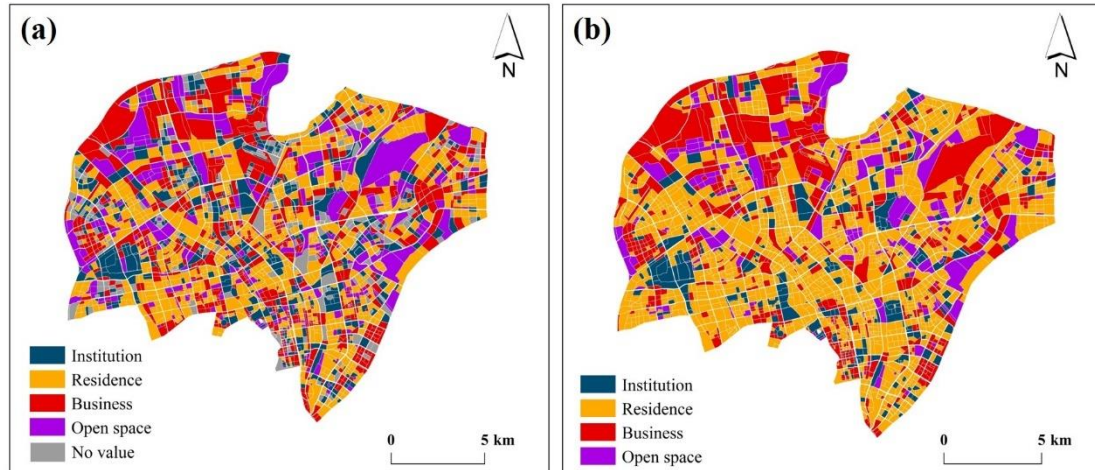


Figure 6-8 Illustration of the (a) DI-based and (b) FI-based urban land use maps in Hangzhou city in 2021.

6.3.2 Urban land use map in 2017

The FI-based classification result in 2021 performed better than the DI-based classification result according to the quantitatively and qualitatively analysis in Hangzhou city. This research thus utilized the FI method to classify urban land use types in 2017. The only difference between the 2017 FI-based classification and the 2021 FI-based classification is the reference map. This research selected parcels for preparing the training and testing parcels of which the land use types were identified by using the Baidu map, Baidu Street View, and a field survey in the 2021 urban land use map. However, the Baidu map and Baidu Street View are not available for 2017. This research thus utilized the OSM map and Google map as the reference maps to identify the training and testing parcels.

Table 6-6 presents the confusion matrix of FI-based classification. The estimated OA for FI-based classification results is 0.817 ± 0.032 . It can be noted that the residence,

business, and open space in FI-based classification can be classified relatively well, with the UA of 0.872 ± 0.043 , 0.758 ± 0.067 , and 0.981 ± 0.035 , respectively. However, the institution has a relatively lower UA (0.587 ± 0.133).

Table 6-6 Confusion matrix of FI-based classification results in 2017 (I: Institution; R: Residence; B: Business; O: Open Space). UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map.

		Map category						
	Class	I	R	B	O	Total	PA	OA
True category	I	0.062	0.002	0.016	0.000	0.080	0.776 ± 0.118	0.817 ± 0.032
	R	0.016	0.408	0.036	0.002	0.462	0.883 ± 0.037	
	B	0.025	0.023	0.238	0.000	0.286	0.831 ± 0.053	
	O	0.002	0.035	0.024	0.111	0.172	0.645 ± 0.079	
	Total (π_j)	0.106	0.468	0.313	0.113	1.000		
	UA	0.587 ± 0.133	0.872 ± 0.043	0.758 ± 0.067	0.981 ± 0.035			

In the 2017 FI-based resulting map (Figure 6-9), the residence parcels were randomly distributed in the city. It can be noted that the number of the residence parcels was larger than the other urban land use types (e.g., institution, business, and open space). Furthermore, the residence parcels were more concentrated in the city center, while more spread out around the city. Most business and open space parcels were distributed around Hangzhou city. Also, the size of the business and open space parcels

around the city was larger than the parcels distributed in the city center. The number of institution parcels was smaller than the other urban land use types (e.g., institution, business, and open space). The institution parcels were mainly concentrated in several areas in the city center.

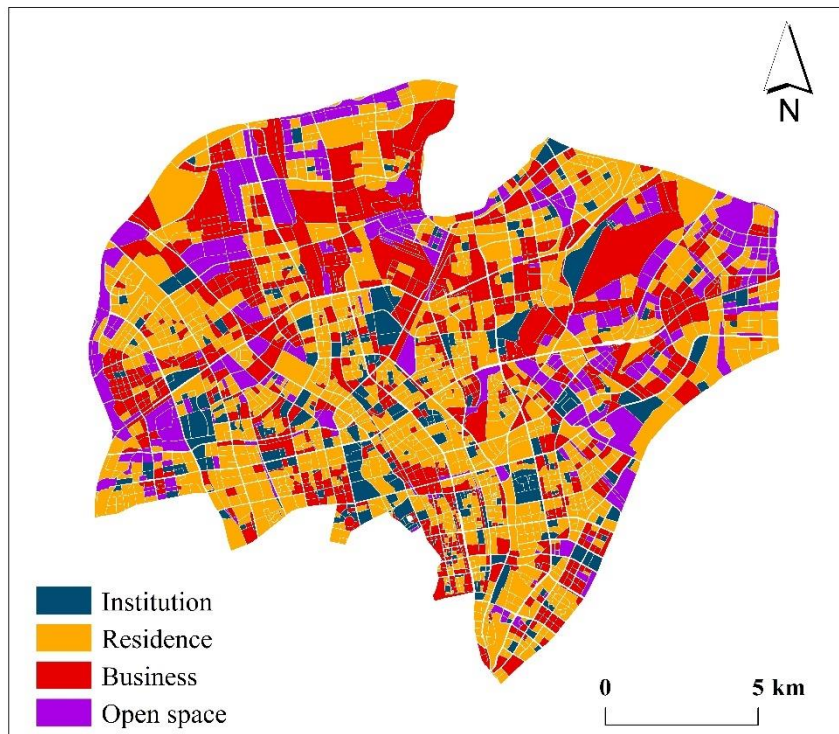


Figure 6-9 Illustration of the FI-based urban land use map in Hangzhou city in 2017.

6.3.3 Urban land use changes from 2017 to 2021

Although the accuracies of the 2017 and 2021 urban land use maps were relatively well, the maps should be post-processed for obtaining high accuracy products and analyzing the impact of urban land use on green space. The post-processed urban land use map in 2017 was shown in Figure 6-10(a). The large open space parcels were distributed around Hangzhou city, while small parcels were distributed in the city center. The institution parcels were concentrated in several areas. As for the residence parcels, they were randomly distributed in the city. The large business parcels were mainly

distributed in the northern city, while the small parcels were distributed in the southern city. The distribution of the 2021 urban land use map (Figure 6-10(b)) was visually consistent with the 2017 urban land use map. However, parcels of each urban land use type were more concentrated in the 2021 urban land use map. To be more specific, the institution, business, and open space parcels were distributed more concentrated in the city.

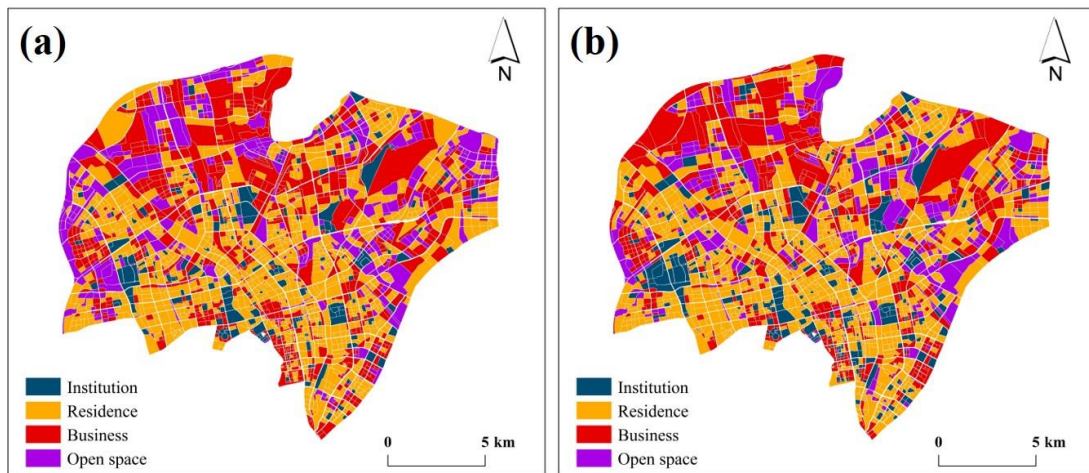


Figure 6-10 Post-processing of the urban land use map in 2017 (a) and 2021 (b).

Table 6-7 showed the numbers of urban land use types within different rings in 2017 and 2021. In general, the number of institution and residence parcels increased, while the number of business parcels and open space parcels decreased from 2017 to 2021. Furthermore, the number of the residence parcels was larger than the others (e.g., institution, business, and open space), and the institution has the lowest number of parcels. In the 1st ring area, the number of the institution and business parcels changed significantly. For example, the number of institution parcels increased from 18 in 2017 to 39 in 2021. It can be noted that the residence and business parcels changed obviously from 2017 to 2021 in the 2nd ring area. The number of parcels within the 3rd ring was the largest, while the number of parcels within the 1st ring belt was the lowest.

Table 6-7 Quantitative in urban land use types in 2017 and 2021.

Class	1 st ring		2 nd ring		3 rd ring		Total	
	2017	2021	2017	2021	2017	2021	2017	2021
Institution	18	39	52	55	132	193	202	287
Residence	120	119	225	235	826	895	1171	1249
Business	114	94	98	79	583	544	795	717
Open space	6	6	32	38	432	341	470	385
Total	258	258	407	407	1973	1973	2638	2638

Figure 6-11 presents the spatial patterns of urban land use changes from 2017 to 2021. It can be noted that the main changes in the 1st ring area were residence and business parcels (Figure 6-11). A large number of residence and business parcels was transformed into institution parcels, which happened in the southern city. To be more specific, these changes include the transformation from residence to the institution, the business to the institution, and business to the residence. In the 2nd ring area, the main changes were also from the residence and business parcels. However, fewer changes happened in the 2nd ring area compared with the 1st ring area. The biggest change occurred within the 3rd ring as shown in Figure 6-11. A large number of open space parcels were transformed into residence and business parcels. This is probably because there are a lot of farmlands in the periphery city and these farmlands were occupied with impervious surfaces due to the urbanization process.

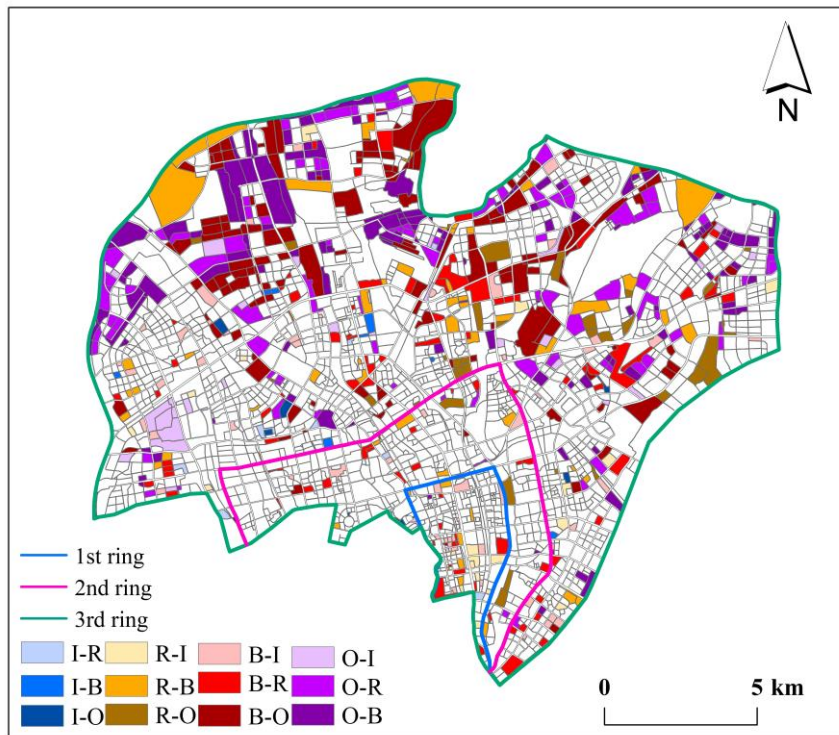


Figure 6-11 Urban land use changes from 2017 to 2021. I:Institution; R: Residence; B:Business; O:Open space.

Table 6-8 shows the quantitative analysis of urban land use changes from 2017 to 2021 in Hangzhou city. In general, the business and open space parcels changed significantly. It can be noted that a large number of business parcels were transformed into residence and institution parcels in the 1st ring area. To be more specific, 15 business parcels changed to institution parcels, and 12 business parcels changed to residence parcels. Also, 9 of the residence parcels were transformed into institution parcels. In the 2nd ring area, 17 business parcels were transformed into residence parcels. The other urban land use types have no obvious change. 140 business parcels in the 3rd ring area were transformed into open space, while 148 open space parcels were transformed into business parcels.

Table 6-8 Quantitative analysis of different urban land use changes from 2017 to 2021

Type	1 st ring		2 nd ring		3 rd ring		Total	
	Quantity	Ratio	Quantity	Ratio	Quantity	Ratio	Quantity	Ratio
1-1	15	0.38	47	0.82	113	0.59	175	0.61
2-1	9	0.23	2	0.04	11	0.06	22	0.08
3-1	15	0.38	8	0.14	39	0.20	62	0.21
4-1	0	0	0	0	30	0.16	30	0.10
1-2	3	0.03	2	0.01	6	0.01	11	0.01
2-2	104	0.87	213	0.90	735	0.82	1052	0.84
3-2	12	0.10	17	0.07	65	0.07	94	0.08
4-2	0	0.00	4	0.02	89	0.10	93	0.07
1-3	0	0.00	0	0.00	8	0.01	8	0.01
2-3	7	0.07	7	0.09	49	0.09	63	0.09
3-3	86	0.91	72	0.90	339	0.62	497	0.69
4-3	1	0.01	1	0.01	148	0.27	150	0.21
1-4	0	0.00	5	0.13	5	0.01	10	0.03
2-4	0	0.00	6	0.15	31	0.09	37	0.10
3-4	1	0.17	2	0.05	140	0.41	143	0.37
4-4	5	0.83	27	0.68	165	0.48	197	0.51

6.4 Discussion

The results from the urban land use changes highlight three key points.

The uncertainty of urban land use changes should be considered during the urban land use mapping process. The emergence of GBD provides new opportunities for urban functional patterns due to GBD containing abundant socioeconomic characteristics that could compensate for the lack of anthropogenic information of RS data (Yao et al. 2022). However, the uncertainty of GBD makes the integration process challenging and also brings difficulties to the quality of urban land use products. In this

regard, the different time periods of GBD used in this research might cause uncertainties in the investigation of spatial and temporal patterns of urban land use change. To be more specific, this research utilized POIs in 2017 and 2021 to classify urban land use types, while the data quality (e.g., data quantity, data attributes, and data accuracy) of POIs in 2017 and 2021 might be different. To solve this problem, we manually checked the quality of POIs in terms of the data quantity and data accuracy on our study site. In addition, our study site is located in the central area of Hangzhou city, which avoids the data missing issues. Therefore, the different time periods of POIs might not affect the analysis result in this research.

According to the analysis of urban land use changes in Hangzhou city, it can be noted that the change in land use around the city is more obvious than that in the city center. Furthermore, the change in land use inside the city is mainly concentrated on commercial and institutional land, while the urban periphery is focused on residential land and open space. The development of urban land use in Hangzhou city is consistent with urban planning and management. The “Hangzhou Urban Master Plan (2001–2020)” has published the layout of urban land use to realize the reorganization, differentiation, and reaggregation of the functions of Hangzhou city ([Lou et al. 2019](#)). The central and south parts of the city should be the commercial regions, and the northern city could play as the industrial areas. Also, the eastern part is zoned for municipal and transportation land, while the western region is zoned for educational and residential land. Since the West Lake tourist area is located in the southern part of Hangzhou city, the lakeside area is planned as tourist and business areas. Therefore, the change in urban land use was significantly affected by the urban planning and management in Hangzhou city.

The experience of urban land use change in Hangzhou city is valuable for small cities with a low development pace of urbanization. This study addressed the problem that how urban functional patterns change in developed cities in China. Although different cities have different patterns of urban utilization change due to different scales, different economic sources, and different geographical locations. The experience of urban land use change in Hangzhou city still has reference value for other cities. The reason is that urban land use change in China is mainly regulated by relevant policies and laws and the urban land use changes thus correspond to different regulations and different development goals. Hangzhou city is now in an advantageous position in terms of policy support, and economic investment. The practical research in Hangzhou city is valuable for urban planning and management.

6.5 Conclusion

This research classified the urban land use in 2017 and 2021 for Hangzhou city and investigated the spatial-temporal dynamics of urban land use from 2017 to 2021. A methodology framework was proposed to classify urban land use types (e.g., institution, residence, business, and open space), based on OpenStreetMap (OSM) road network data, 10 m Sentinel-2A images, and Gaode Points of interest (POIs). Specifically, certain RS and GBD features were improved referring to the importance analysis of parcel attributes. The results demonstrated that the FI-based result (overall accuracy: 80.2%) in 2021 performed better than the DI-based classification result (overall accuracy: 67.0%) according to the quantitative and qualitative analysis. The FI-based approach was thus used to map 2017 urban land use in Hangzhou city, with an overall accuracy of 81.7%. The classification results were post-processed by visual

interpretation for obtaining highly accurate urban land use products. The results also showed that the number of institution and residence parcels increased, while the number of business and open space parcels decreased from 2017 to 2021. A large number of open space parcels were transformed into the residence and business parcels around the urban periphery. A few business parcels were changed into residence and institution parcels in the urban core area. The understanding of the spatial-temporal patterns of urban land use changes is valuable for guiding cities to improve their urban green space patterns and system.

Chapter 7: Investigating the urban green space distribution impacted by urban land use changes: A case study in Hangzhou, China

Stage of publication - First draft prepared. Awaiting feedback from co-authors before submission to the journal Landscape and Urban Planning.

7.1 Introduction

Urban green spaces are considered the most important surface landscapes in urban settings ([Haas and Ban 2014](#); [Hepcan 2012](#); [Wolch et al. 2014](#)). It contains a variety of benefits including physical health benefits, psychological health benefits, socio-economic benefits, and environmental benefits ([Haas and Ban 2014](#); [Kwon et al. 2021](#)([Kondo et al. 2018](#); [Kwon et al. 2021](#))). The rapid urbanization in China significantly affects urban green spaces ([Sperandelli et al. 2013](#); [Yu et al. 2017](#)), causing an increase in exposure to the ecosystem and environmental hazards (e.g., urban heat island, urban flooding, urban heat waves) ([Aram et al. 2019](#); [Heidt and Neef 2008](#)). Recently, many developed cities, especially Beijing, Shanghai, Shenzhen, Guangzhou, and Hangzhou, are dedicated themselves to protecting urban green space to realize the sustainable development of cities ([Kuang and Dou 2020](#); [Song et al. 2021](#)). Urban land use and greening policies such as “the old town transformation” and “urban green space system plan” are formulated to develop the distributions of urban green space ([Liu et](#)

[al. 2021](#)). As a result, it's critical to look at how urban green space distribution changes in response to urbanization.

Previous efforts provide an improved understanding of the urban green space variations in response to urban land cover changes (e.g., physical aspects of urbanization) ([Liu et al. 2021](#) [Tang et al. 2018](#)). However, few studies have examined the effects of urban land use changes (e.g., socio-economic aspects of urbanization) on urban green space. In addition, the ecological and socio-economic functions of urban green spaces have proven to be highly dependent on their spatial patterns ([Hernández-Moreno and Reyes-Paecke 2018](#); [Wang et al. 2020](#); [Woldesemayat and Genovese 2021](#); [Xu et al. 2018](#)). For example, the increasing area of urban green space is highly correlated with the reduction of urban heat island ([Meng et al. 2018](#)). Also, there is a strong correlation between urban green space quantity and gross domestic product ([Chen et al. 2017](#)). It is therefore important to understand the spatial characteristics of urban green space impacted by urban land use changes, which can guide future urban planning and management.

Several approaches have been utilized to analyze the impact of urbanization on urban green space distributions ([Chan and Vu 2017](#); [Gavriliadis et al. 2019](#); [Yang et al. 2018](#)). These efforts can be categorized into statistical analysis, landscape analysis, and geospatial analysis. Among them, the statistical analysis could reveal the spatial patterns of urban green space at the landscape level ([Feng and Astell-Burt 2018](#); [Wüstemann et al. 2017](#)), which has the potential to understand the distribution of urban green space. Few studies, on the other hand, have sought to combine these methodologies to investigate the spatial-temporal patterns of urban green space and its response to urban land use changes.

In this context, this research analyzed the variations of urban green space distribution in response to urban land use changes (e.g., from open space to residence) by integrating statistical analysis, landscape analysis, and geospatial analysis in Hangzhou city from 2017 to 2021. The understanding of urban green space patterns influenced by urban land use changes is valuable for small cities to develop their urban green space system. This research is organized into five sections. In section 7.2, the study site, data source, and methods used in the research were presented. Section 7.3 analyzed the urban green space distribution within different urban land use changes. Section 7.4 discussed the related issues and challenges. Section 7.5 concluded the main findings and implications.

7.2 Materials and methods

7.2.1 Study site

Hangzhou is located in the northwestern part of Zhejiang province, which plays an important role in the Yangtze River Delta Urban Agglomerations. Hangzhou comprises 10 districts, 1 county-level city, and 2 counties. In this paper, we only focused on the most central urban districts, which are Shangcheng District, Gongshu District, and part of Xihu District. The population in this region accounts for 45% of the total population in Hangzhou city. Hangzhou city is one of the representative cities of urbanization, population growth, and economic development in China ([Li et al. 2017a](#); [Mao et al. 2020](#)).

In addition, this research selected the area surrounded by the third ring (i.e., the Hangzhou belt highway) as the study site (Figure 7-1). The urban areas can be divided into three units including 1st ring belt, 2nd ring belt, and 3rd ring belt. To be more specific,

the 1st ring represents Huancheng North Road, Huancheng East Road, Jiangcheng Road, Zhongshan South Road, Wansongling Road, Nanshan Road, Hubin Road, and Huancheng West Road. The 2nd ring refers to Desheng Road, Qiutao Road, Fuxing Road, Old Fuxing Road, Hupao Road, Manjuelong Road, Wulaofeng Tunnel, Jiqingshan Tunnel, Jiulisong Tunnel, Lingxi Tunnel, Zijinghua Road, and Wenyi Road. The 3rd ring is the Hangzhou belt highway.

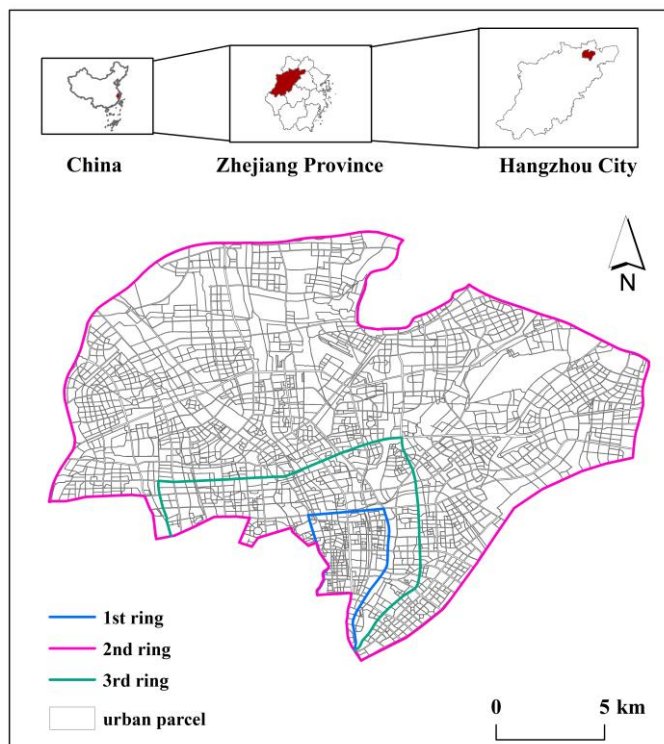


Figure 7-1 Study area

7.2.2 Data

(1) Gaofen-2 high-resolution images

GF-2 satellite was designed and developed by CAST (China Academy of Space Technology) (Table 7-1). It employs the CAST-CS-L3000A bus and two PAN/MS cameras, capable of collecting images with a GSD (Ground Sampling Distance) of 0.81 m in panchromatic and 3.24 m in the multispectral bands on a swath of 45 km ([Han et al. 2020](#)). The GF-2 spacecraft was launched on August 19, 2014 (03:15 UTC) with a

CZ-4B (Long March-4B) vehicle from TSLC (Taiyuan Satellite Launch Center) in the Shanxi Province of northern China ([Tong et al. 2018](#)).

Table 7-1 Gaofen-2 Satellite Sensor Specifications

Parameters	Panchromatic	Multispectral
Spectral range (μm)	0.45-0.89	B1/blue:0.45-0.52 B2/green:0.52-0.59 B3/red:0.62-0.69 B4/NIR:0.77-0.89
Swath width (km)	45	45
Viewing angle	0°-25°	0°-25°
Repetition Cycle (days)	5	5
Spatial resolution (m)	0.8	3.2
Global-mode Coverage Ability (days)	60	60

This research used two scenes of GF-2 high-resolution images in Hangzhou (Figure 7-2), which were acquired on 10 October 2017, and 27 September 2021. Both scenes were preprocessed with a quick atmospheric correction method and geometrical rectification in ENVI 5.3. GF-2 imagery is an ideal data source for urban green space classification.

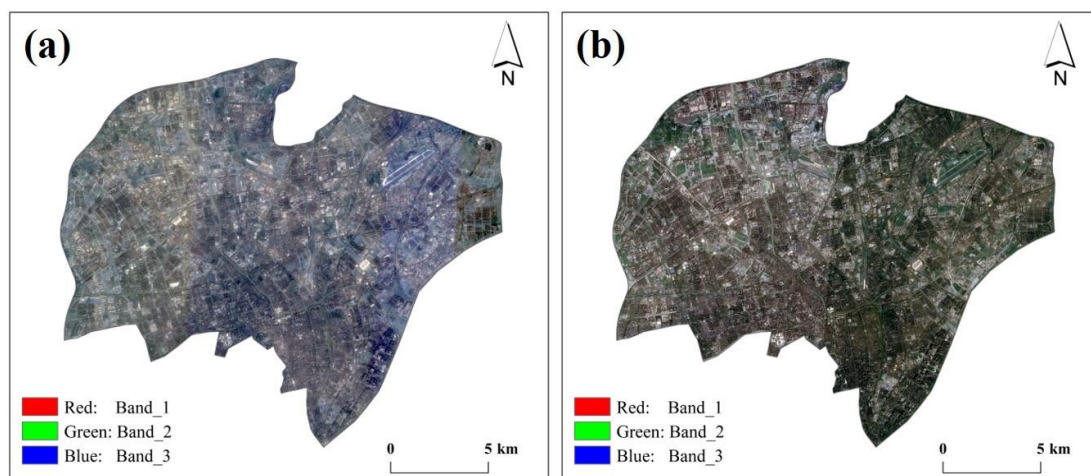


Figure 7-2 Gaofen-2 images

(2) Urban land use data

Notably, the difference between urban land use and urban land cover and their implications in ecology is often overlooked, because most traditional ecological research concern the biological features of land surface other than the socio-economic features. Urban land use maps were produced by using the feature integration (FI) approach (Yin et al. 2021b). Four urban land use types were selected, namely, institution, residence, business, and open space, for analysis because they represented the most important socio-economic activities of Hangzhou city, and the green spaces in these urban land use types played important social, environmental, and ecological roles for urban residents.

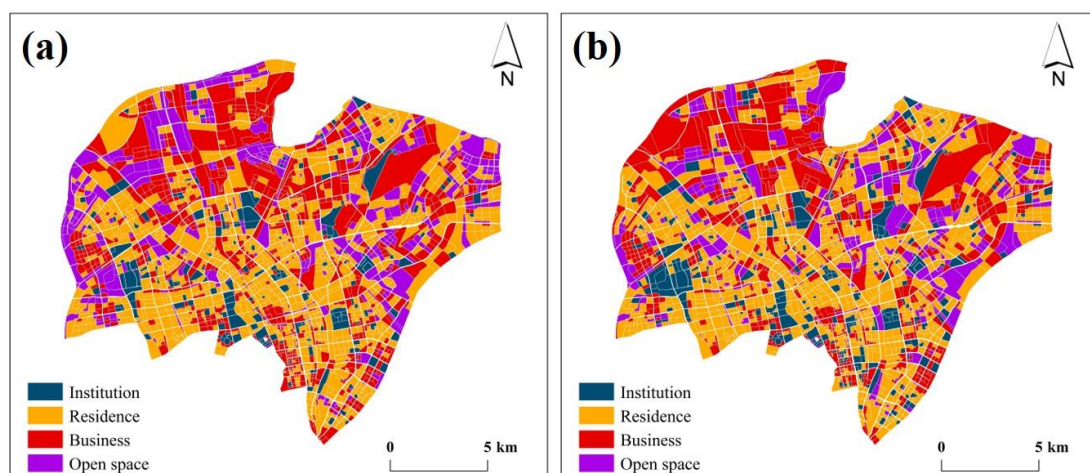


Figure 7-3 Urban land use maps in 2017 and 2021

7.2.3 Methods

This research integrated statistical analysis (e.g., green space ratio and proportion), landscape analysis (e.g., landscape metrics), and geospatial analysis (hotspot analysis) to investigate the variations of urban green space (Figure 7-4).

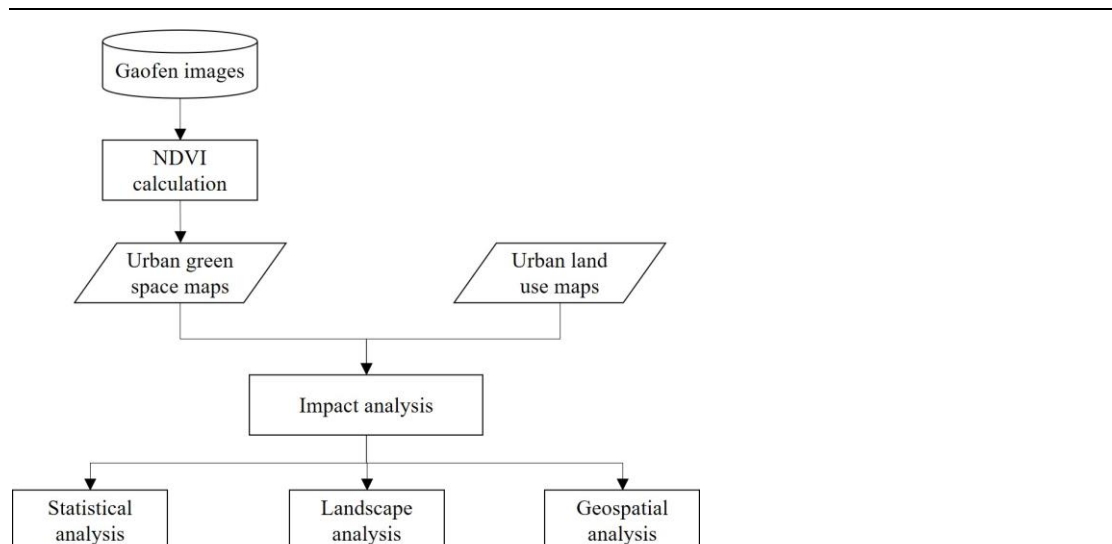


Figure 7-4 Methodology framework

(1) Urban green space mapping

This research mapped urban green space by calculating the normalized difference vegetation index (NDVI). NDVI is one of the most commonly used indices for extracting vegetation greenness. According to previous studies, vegetation surfaces absorb most of the red light that hits them while reflecting much of the near-infrared light. On the contrary, non-vegetation surfaces reflect more red light and less near-infrared light. The NDVI index was calculated using equations as follows.

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)}$$

where NIR represents the near-infrared channel and Red is the red channel. The result for the calculation of NDVI generally ranges from minus one (-1) to plus one (+1). Results close to +1(0.8-0.9) indicate the highest possible of urban green spaces.

(2) Green space area and ratio analysis

This research calculated the urban green space area and ratio for each urban land use type (Institution, Residence, Business, and Open space) to capture the spatial-temporal patterns of urban green space in Hangzhou city. To be more specific, the urban green space area and ratio for the total area, 1st ring belt, 2nd ring belt, and 3rd ring belt

were calculated respectively. The urban green space area and ratio for each urban land use type were also analyzed.

(3) Linear regression

Simple linear regression is a linear regression model with a single explanatory variable. That is, it concerns two-dimensional sample points with one independent variable and one dependent variable and finds a linear function that, as accurately as possible, predicts the dependent variable values as a function of the independent variables. The relationship between urban green space areas and patch areas of different urban land use types were calculated and analyzed through linear regression models with t-test at the 5% significance level. Simple linear regressions were carried out in the platform of SPSS.

(4) Hotspot analysis

This research utilized the Getis-Ord hotspot of ArcGIS tools to identify the hotspots and coldspots of urban green space patches within different in Hangzhou city. Hotspots analysis utilized a series of weighted features and identify statistically significant hot spots and cold spots using Getis-Ord G_i^* statistics, which calculates the $G_iZScore$ and $G_iPValue$ for the selected parameters ([Wang et al. 2021a](#)). In general, hot spots and cold spots have very high or very low z-scores, with very small p-values, respectively. Here, this research used the percentage of confidence level above 90% to identify hot/cold spots.

(5) Landscape analysis

Landscape metrics are often used to detect and quantify the landscape patterns of urban green space patches ([Herold et al. 2002](#); [Kumar et al. 2018](#); [Uuemaa et al. 2009](#)). This research selected six landscape metrics including six class-level metrics and three

landscape-level metrics by referencing earlier work reported in many research (Table 7-2). These metrics are used to reflect fragmentation, connectivity, and diversity of landscape patterns. The six class-level metrics include patch density (PD), largest patch index (LPI), Mean shape index (SHAPE_MN), and patch cohesion index (COHESION). The three landscape-level metrics include Contagion index (CONTAG) and Shannon's Evenness Index (SHEI). Among these, PD, LPI, and SHAPE_MN can represent the fragmentation and complexity of urban green space patches ([Bosch et al. 2020](#); [Zhang et al. 2020b](#)); COHESION could reflect the connectivity ([Inkoom et al. 2018](#)); CONTAG presents the connectivity of the total landscape ([Miller et al. 2020](#)); SHEI presents the distribution of landscapes ([Pindral et al. 2020](#)).

Table 7-2 Landscape metrics

Metrics		Description	Equation
Class-level	Patch Density (PD)	PD measures the density patches for each class.	①
	Largest Patch Index (LPI)	LPI quantifies the percentage of the total landscape area comprised by the largest patch.	②
	Mean shape index (SHAPE_MN)	Average evaluates the complexity of urban green space patches.	③
	Patch Cohesion Index (COHESION)	COHESION measures the connectedness of urban green space patches.	④
Landscape-level	Contagion Index (CONTAG)	CONTAG measures the fragmentation in the entire landscape.	⑤
	Shannon's Evenness Index (SHEI)	SHEI measures the dominance of urban green space patches in the entire landscape.	⑥

①

$$PD = \frac{n_i}{A} (10,000)(100)$$

n_i = number of patches in landscape of patch type (class) i .
 A = total landscape area (m²).

②

$$LPI = \frac{\max_{j=1} (a_{ij})}{A} (100)$$

a_{ij} = area(m²) of patch ij .
 A = total landscape area (m²).

③

$$MN = \frac{\sum_{j=1}^n X_{ij}}{n_i}$$

MN equals the sum, across all patches of the corresponding patch type, of the corresponding patch metric values, divided by the number of patches of the same type.

④

$$COHESION = \left[1 - \frac{\sum_{j=1}^n P_{ij}^*}{\sum_{j=1}^n P_{ij}^* \sqrt{a_{ij}^*}} \right] \cdot \left[1 - \frac{1}{\sqrt{Z}} \right]^{-1} \cdot (100)$$

P_{ij}^* = perimeter of patch ij in terms of number of cell surfaces.
 a_{ij}^* = area of patch ij in terms of number of cells; Z = total number of cells.

⑤

$$CONTAG = \left[1 + \frac{\sum_{i=1}^m \sum_{k=1}^m \left[P_i \cdot \frac{g_{ik}}{\sum_{k=1}^m g_{ik}} \right] \cdot \left[\ln \left(P_i \cdot \frac{g_{ik}}{\sum_{k=1}^m g_{ik}} \right) \right]}{2 \ln(m)} \right] (100)$$

P_i = proportion of the landscape of (class) i ; m = number of patch types.
 g_{ik} = number of adjacencies(joins) between pixels of patch types (classes) i and k .

⑥

$$SHEI = \frac{-\sum_{i=1}^m (P_i \cdot \ln P_i)}{\ln m}$$

P_i = proportion of the landscape occupied by patch type(class) i .
 m = number of patch types(classes)present in the landscape, excluding the landscape border of present.

7.3 Results

7.3.1 Urban green space distribution

Figure 7-5 represents the urban green space maps in 2017 (7-5a) and 2021 (7-5b). In general, urban green spaces in the 2017 map were distributed mainly around the urban periphery, especially in the northwest regions and northeast regions of Hangzhou city. Compared with the urban green spaces in the urban periphery, the urban green spaces in the urban core areas were significantly less. As for the urban green space distribution in the 2021 map, the area of urban green spaces has increased significantly in the whole city. It should be noted that the urban green spaces in the urban core area have increased. While some urban green spaces in the urban periphery have decreased especially in the northeast regions.

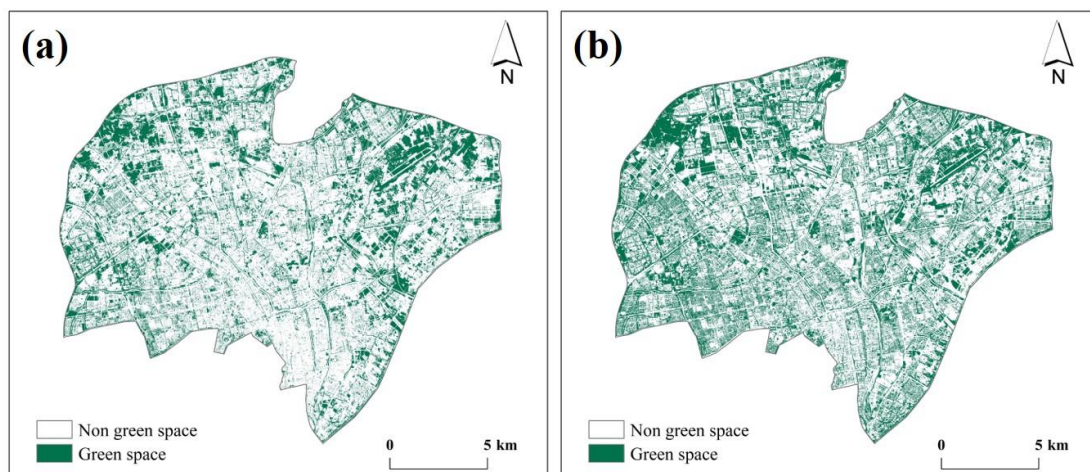


Figure 7-5 Urban green space distribution in 2017 (a) and 2021 (b).

7.3.2 Spatial distribution of urban green space

(1) Hotspot analysis

Hotspot analysis was used to identify the clusters of the increasing and decreasing urban green space patches in 2017 and 2021. Figure 7-6(a) presents the hotspot analysis of the urban green space ratio in Hangzhou city in 2017. In general, the high-value regions (e.g., hot spots) of urban green space coverage had a range of 5%-10%, while the low-value regions (e.g., cold spots) had about 20%. The distribution of spatial clustering differed by region in Hangzhou city in 2017. Specifically, the clusters of hot spots tended to occur in the urban periphery except for the southern part of the city. The hot spots were mostly distributed in the 3rd ring belt. The cold spots of low-value regions were mostly found in the city center and the northern city. Most of the cold spots were distributed in the 1st ring belt and 2nd ring belt. In Figure 7-6(b), the high-value regions of urban green space coverage had approximately 15%, while around 5% to 10% of urban land use parcels were low-value regions. Compared with the 2017 hotspots map, the number of high-value and low-value urban land use parcels in 2021 was more likely to undergo a greater change in urban green space. As for the distribution patterns, the hot spots of urban green space coverage were mostly concentrated in the urban periphery. The clusters of cold spots tended to occur in the city core and the northwest and northeast parts of the city. The hot/cold spots of urban green space coverage were distributed more spread out in Hangzhou city. In addition, the clusters of hot spots have increased in the 2nd ring belt, and cold spots have decreased in the 1st ring belt and 2nd ring belt.

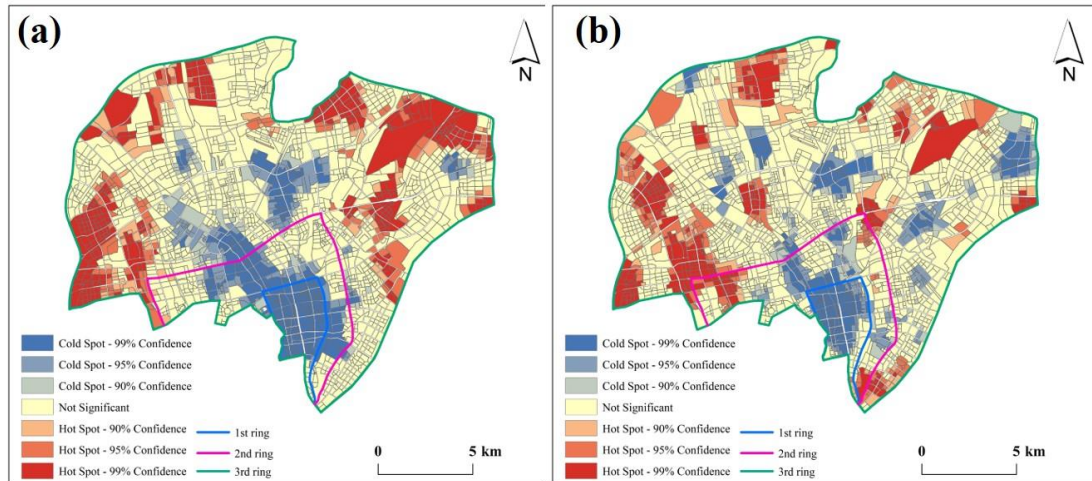


Figure 7-6 Hotspot analysis of urban green space distribution at ring-level in 2017 (a) and 2021 (b)

Figure 7-7 presents the increase and decrease of urban green space coverage in Hangzhou city. It can be shown that approximately 5% of urban land use parcels decreased, while the increase in urban green space coverage had a range of 10%-15%. Furthermore, the changes in urban green space tended to be clustered in space. To be more specific, the clusters of increased urban green space coverage differed greatly by region from that of the decreased urban green space coverage. The cold spots mostly occurred in the urban periphery, while the hot spots were concentrated in the urban core and the west of the city. The clusters of the increased and decreased urban green space coverage mostly occurred in the 3rd ring belt, while some of the increased coverage was distributed in the 1st and 2nd ring belts.

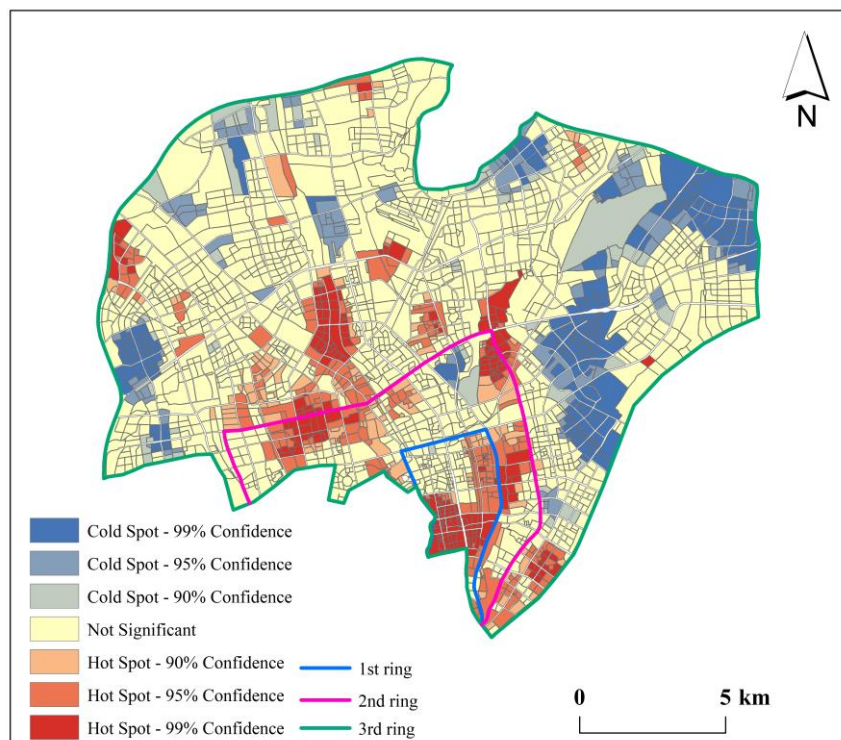


Figure 7-7 Hotspot analysis of urban green space distribution at ring-level from 2017 to 2021

(2) Urban green space area and proportion analysis

The areas and proportions of urban green space patches from the 1st ring belt to the 3rd ring belt are shown in Table 7-3. It can be shown that the overall area of urban green space increased significantly from 59.52 km² (22%) in 2017 to 81.29 km² (36%) in 2021. In the 1st ring belt, the area of urban green space has increased from 0.79 km² in 2017 to 1.99 km² in 2021, and the proportion has increased from 10% to 23%. The area of urban green space in the 2nd ring belt has increased from 2017 to 2021, from 5.16 km² to 9.02 km². The area of urban green space patches in the 3rd ring belt has increased from 53.56 km² to 70.27 km². Among the different ring belts, the growing proportion of urban green space patches in the 1st ring belt was the highest (13.7%), while the proportion of the increased urban green space in the 3rd ring belt was the lowest (8.2%).

Table 7-3 Analysis of urban green space area and proportion at ring-level in 2017 and 2021

Region	2017		2021	
	Area (km ²)	Proportion	Area (km ²)	Proportion
Total	59.52	0.22	81.29	0.36
1 st ring	0.79	0.10	1.99	0.23
2 nd ring	5.16	0.19	9.02	0.32
3 rd ring	53.56	0.28	70.27	0.37

According to Figure 7-8, Figure 7-9, Table 7-4, and Table 7-5, the areas of urban green spaces and areas of urban land use parcels are highly correlated. It can be noted that the urban green space areas within the urban parcels have increased from 2017 to 2021, except for the open space parcels. Among them, the urban green spaces within the institution parcels has the largest growth, while the growth of urban green spaces within the institution parcels is less obvious. The reason for the loss of urban green spaces within the open space parcels is probably due to the transformation from farmland to impervious surfaces around the urban periphery.

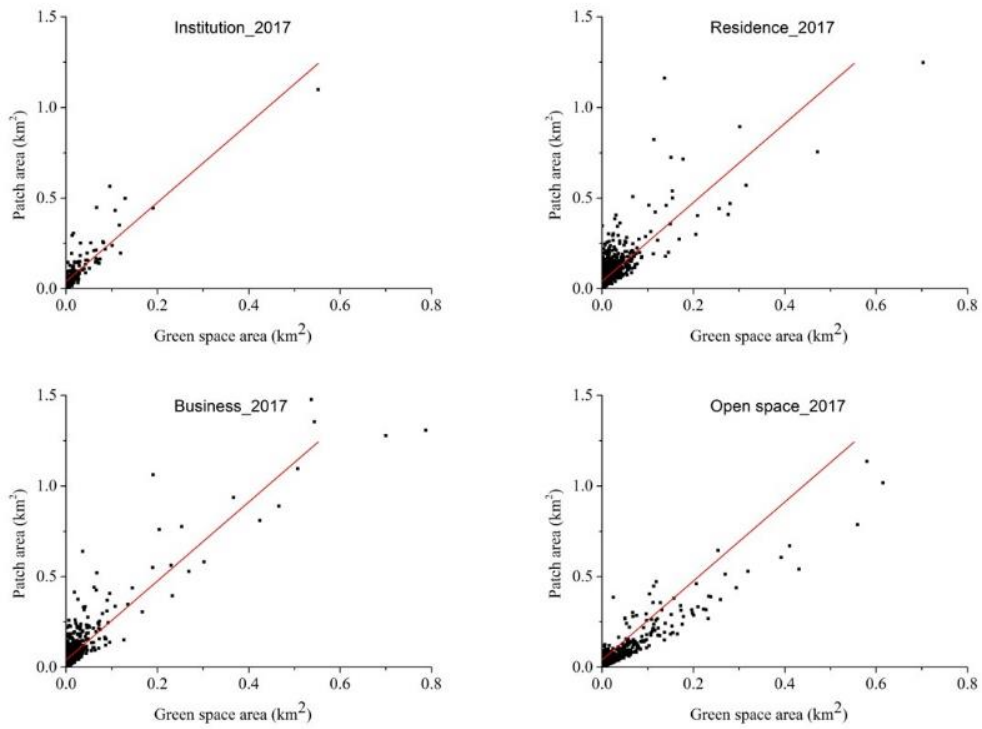


Figure 7-8 Regression analysis of the relationship between urban green space areas and patch areas of different urban land use types in 2017.

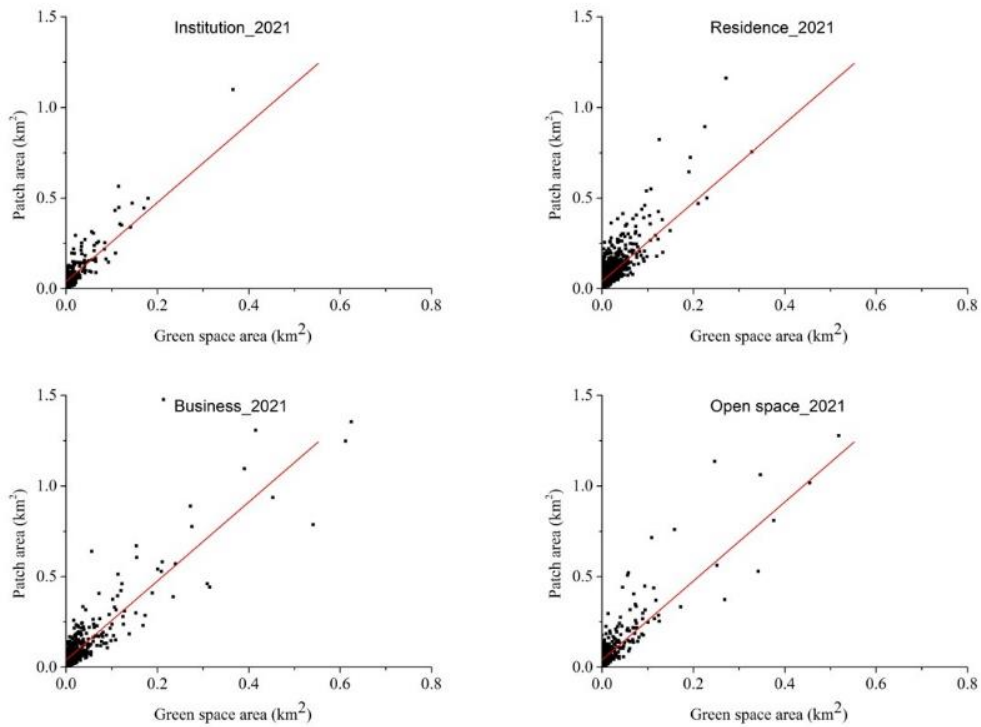


Figure 7-9 Regression analysis of the relationship between urban green space areas and patch areas of different urban land use types in 2021.

Table 7-4 Regression analysis of the relationship between urban green space areas and patch areas of different urban land use types in 2021.

Types	Equation	P Value
Institution_2017	$y=2.2x+0.04$	0.002
Residence_2017	$y=1.5x+0.06$	0.001
Business_2017	$y=1.8x+0.05$	0.004
Open space_2017	$y=1.6x+0.02$	0.001

Table 7-5 Regression analysis of the relationship between urban green space areas and patch areas of different urban land use types in 2021.

Types	Equation	P Value
Institution_2021	$y=1.7x+0.03$	0.004
Residence_2021	$y=1.2x+0.04$	0.001
Business_2021	$y=1.6x+0.04$	0.002
Open space_2021	$y=2.1x+0.04$	0.002

(3) Landscape analysis

The Fragstats 4.2 was utilized to calculate the class-level and landscape-level metrics of urban green space patches in Hangzhou city for 2017 and 2021 (Figure 7-10). Specifically, PD in 2017 was higher than that in 2021, which indicates that the number of urban green space patches has decreased. However, the PD of urban green space in the 1st ring belt has increased. The LPI of urban green space decreased from 2017 to 2021, which means the size of the largest urban green space patch is decreasing. However, the LPI in the 2nd ring belt and 3rd ring belt has increased. The SHAPE_MN increased during the period, indicating the shapes of urban green space patches became more complex. The COHESION of urban green space patches increased slightly from 2017 to 2021, which represents that there are no obvious changes in the connectivity among urban green space patches. The CONTAG displayed a downward trend from 2017 to 2021, indicating more interspersed and increasingly fragmented landscape patches. The SHEI increased, which presents the distribution of landscape patches become more even and equitable.

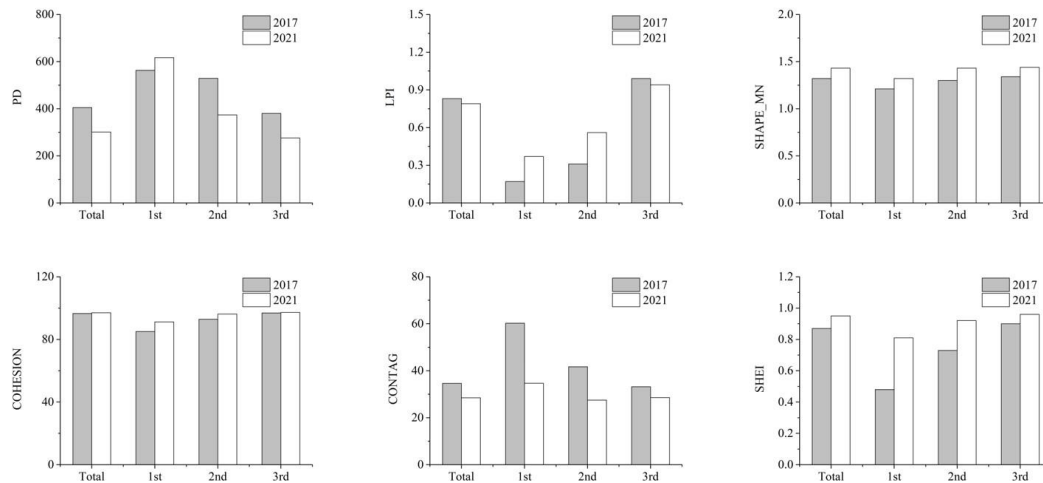


Figure 7-10 Landscape analysis of urban green space distribution at ring-level in 2017 and 2021

7.3.3 Spatial distribution of urban green space in different urban land use changes

(1) Urban green space area and proportion analysis

The areas and proportions of urban green spaces were calculated in different urban land use changes in 2017 and 2021. It can be noted from Figure 7-11 that the increase of urban green space is largely dependent on unchanged institution parcels, indicating well-planned urban green spaces within institution parcels. The unchanged institution parcels in the 1st ring belt have the largest proportion growth of urban green space, while it has relatively small urban green space areas. Also, urban green space areas within the institution parcels transformed from the residence and business parcels have increased. In the 2nd ring belt, there were fewer changes in urban green space areas within the institution parcels transformed from other land use parcels.

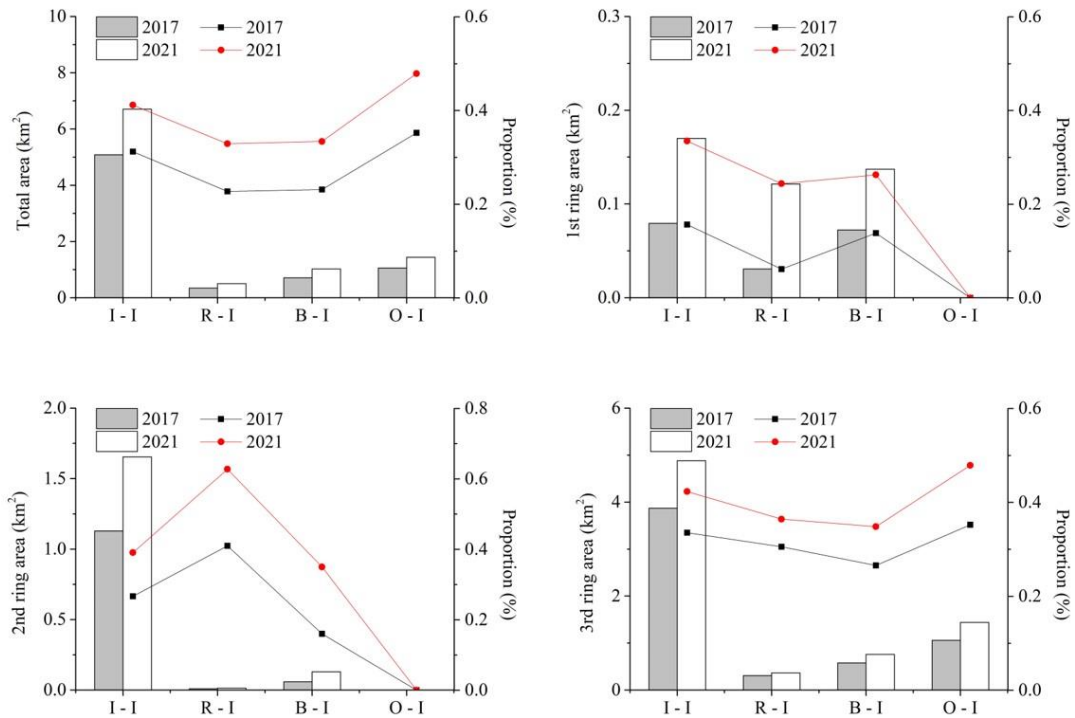


Figure 7-11 Areas and proportion analysis of urban green space in the institution parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space)

Figure 7-12 showed the areas and proportions of urban green spaces in the residence parcels changed from the four urban land use types in 2017 and 2021. In general, the increase of urban green space areas is mainly dependent on unchanged residence parcels, and the increased area is significant. It can be noticed that the urban green space areas have increased from 0.26 km² (0.08) in 2017 to 0.95 km² (0.29) in 2021, indicating the urban green space of the unchanged residence parcels has been greatly improved in the 1st ring belt. In addition, the urban green space areas within the residence parcels transformed from the open space parcels have decreased in the 2nd ring belt and 3rd ring belt. This is probably because these residence parcels were still under construction.

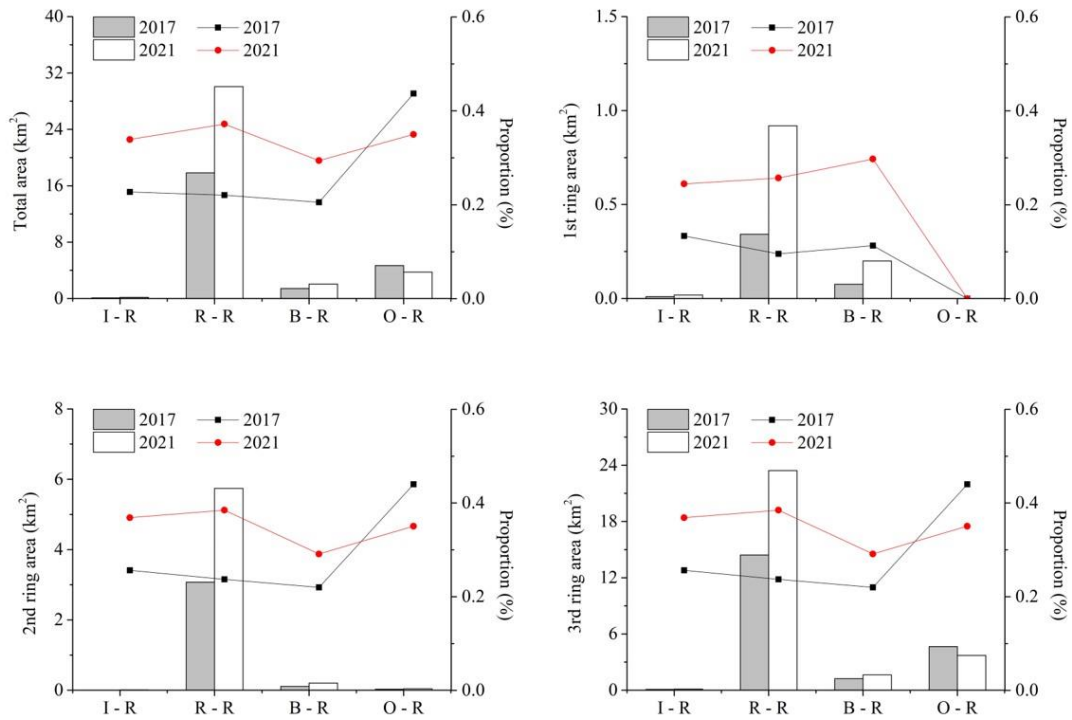


Figure 7-12 Areas and proportion analysis of urban green space in the residence parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space)

Figure 7-13 presented the areas and proportions of urban green spaces in the business parcels changed from the four urban land use types in 2017 and 2021. Generally, the increase of urban green space areas is mainly dependent on unchanged business parcels, from 10.04 km² (0.23) in 2017 to 17.25 km² (0.38) in 2021. Though the proportions of business parcels transformed from the residence and open space parcels were relatively high, the changes in urban green space areas were less obvious. The proportion growth of urban green space within unchanged business parcels was significant in the 1st ring belt and the 2nd ring belt.

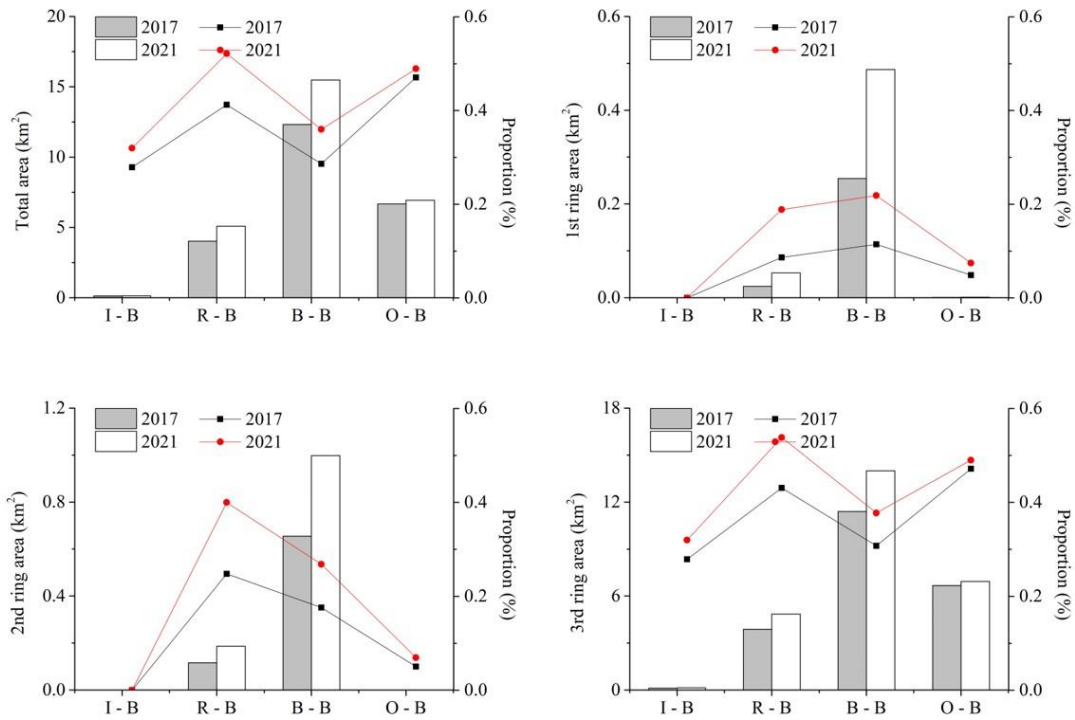


Figure 7-13 Areas and proportion analysis of urban green space in the business parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space)

Figure 7-14 presented the areas and proportions of urban green spaces in the open space parcels changed from the four urban land use types in 2017 and 2021. In this research, the open space parcels mainly include bare land, parks, squares, and unused land. In general, the urban green space area has decreased from 0.15 km² in 2017 to 0.11 km² in 2021 within the unchanged residence parcels, which is related to the decrease of farmland in the 3rd ring belt. The urban green space area also has decreased in 2021 within the residence parcels transformed from the institution parcels. It should be noted that the proportion growth of urban green space within unchanged open space parcels was significant.

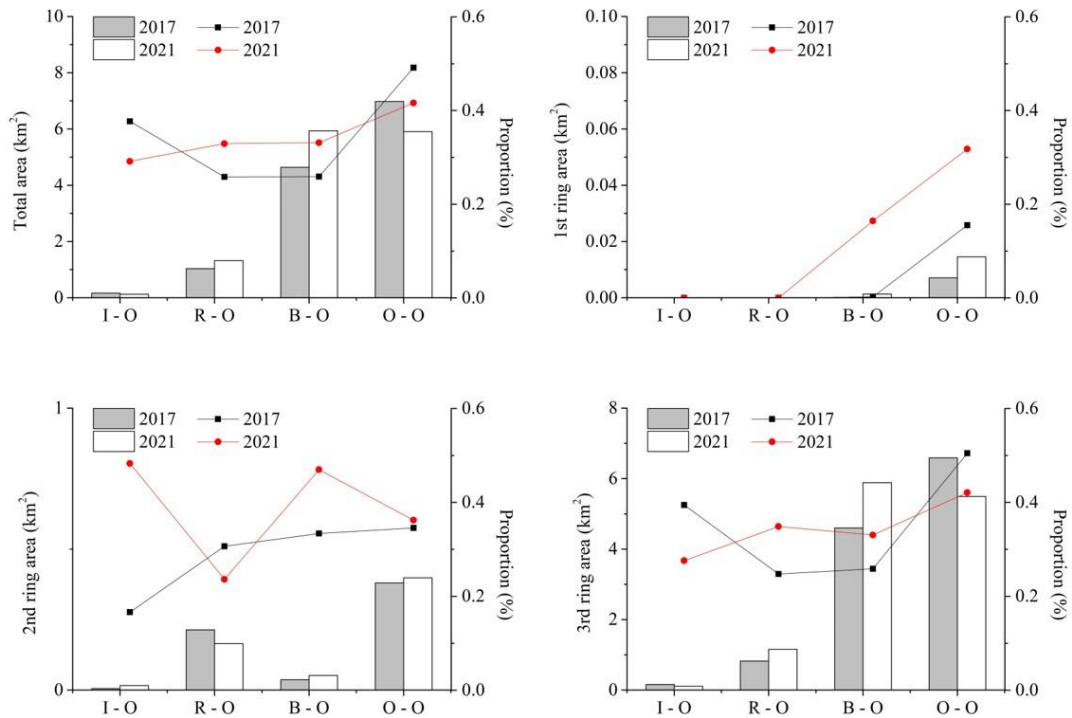


Figure 7-14 Areas and proportion analysis of urban green space in the open space parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space)

(2) Landscape analysis

Table 7-6 represented the landscape analysis of urban green space in the institution parcels changed from the four urban land use types in 2017 and 2021. Among them, PD can be used to evaluate the fragmentation of urban green space patches. It can be noted that the PD values for the institution parcels changed from the four urban land use types have all decreased, suggesting the fragmentation reduced. In addition, the SHAPE_MN, COHESION, and SHEI values for all patches have increased, indicating the more complex, more aggregated, and less dominant urban green space patches in the institution parcels in 2021. The CONTAG values for all green space patches have decreased, which means the patches become uneven across the landscape. The LPI values for all patches have increased except for the institution parcels changed from the

residence parcels. In summary, urban green space patches for the institution parcels in 2021 are less fragmented, more complex, and more aggregated, which indicates the distribution of urban green space patches is more reasonable.

Table 7-6 Landscape analysis of urban green space in the institution parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space)

Change types		Class-level				Landscape-level	
		PD	LPI	SHAPE-MN	COHESION	CONTAG	SHEI
I-I	2017	388.00	2.14	1.37	96.23	31.28	0.89
	2021	292.18	2.61	1.47	97.21	25.53	0.97
R-I	2017	441.25	6.67	1.28	95.19	41.67	0.77
	2021	372.97	3.94	1.42	95.40	29.45	0.91
B-I	2017	477.52	1.16	1.33	92.83	39.66	0.78
	2021	366.05	2.07	1.43	95.04	30.04	0.92
O-I	2017	324.82	5.43	1.34	96.97	32.61	0.94
	2021	283.55	6.06	1.41	97.14	27.70	1.00

Table 7-7 represented the landscape analysis of urban green space in the residence parcels changed from the four urban land use types in 2017 and 2021. The landscape results for the residence parcels were similar to the institution parcels. It can be noted that the LPI, SHAPE_MN, COHESION, and SHEI values for all patches have increased except for the residence parcels changed from the open space parcels, indicating the more complex, more aggregated, and less dominant urban green space patches in the institution parcels in 2021. The PD and CONTAG values for all green space patches have decreased except for the residence parcels changed from the open space parcels, which means the patches become uneven across the landscape. The urban green space patches within the residence parcels changed from the open space parcels becoming

more fragmented, more even, and more dominant. This is probably because the residence parcels transformed from the residence parcels were still in the demolition stage.

Table 7-7 Landscape analysis of urban green space in the residence parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space)

Change types	Class-level				Landscape-level		
	PD	LPI	SHAPE-MN	COHESION	CONTAG	SHEI	
I-R	2017	476.21	4.00	1.37	92.47	39.30	0.77
	2021	421.84	10.17	1.45	95.59	26.73	0.92
R-R	2017	503.95	0.22	1.31	93.50	40.45	0.75
	2021	355.74	0.29	1.46	96.28	25.33	0.95
B-R	2017	428.76	1.47	1.31	93.85	44.16	0.73
	2021	383.14	1.51	1.39	94.97	33.70	0.87
O-R	2017	280.90	2.13	1.35	97.31	28.86	0.99
	2021	306.09	1.69	1.43	96.30	31.10	0.93

Table 7-8 represented the landscape analysis of urban green space in the business parcels changed from the four urban land use types in 2017 and 2021. According to the landscape analysis, the urban green space distribution within the business parcels was more complex, less fragmented, more evenly, less aggregated, and less dominant. In addition, the urban green space distribution in the unchanged business parcels showed less evenly. The business parcels changed from the institution and residence parcels showed less aggregated distribution of urban green space. As for the business parcels changed from the open space parcels, the PD and SHEI values of urban green space patches were decreased, which means the patches become less fragmented and less dominant. The green space patches within the business parcels in 2021 showed

relatively high complexity and low fragmentation. Unlike the institution and residence parcels in 2021, the urban green space patches within the business parcels in 2021 showed less dominance.

Table 7-8 Landscape analysis of urban green space in the business parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space)

Change types		Class-level				Landscape-level	
		PD	LPI	SHAPE-MN	COHESION	CONTAG	SHEI
I-B	2017	488.29	9.01	1.37	95.16	31.49	0.85
	2021	377.00	5.00	1.38	93.76	31.56	0.90
R-B	2017	348.63	14.64	1.29	99.00	28.95	0.98
	2021	202.79	18.52	1.40	98.97	29.77	1.00
B-B	2017	370.47	4.42	1.32	97.58	36.54	0.86
	2021	276.38	3.84	1.41	97.82	31.91	0.94
O-B	2017	273.12	2.33	1.39	97.39	27.70	1.00
	2021	194.85	3.37	1.43	97.66	30.94	1.00

Table 7-9 represented the landscape analysis of urban green space in the open space parcels changed from the four urban land use types in 2017 and 2021. The green space distribution within open space parcels changed from the residence and business parcels showed less fragmented, more complex, less evenly, and less dominant. While the green space distribution within unchanged open space parcels and open space parcels changed from the institution parcels showed more fragmented, less complex, more evenly, and more dominant. It should be noted that the urban green space distribution in the open space parcels changed from the institution parcels showed more fragmented.

Table 7-9 Landscape analysis of urban green space in the open space parcels changed from the four urban land use types. (I: Institution; R: Residence; B: Business; O: Open space)

Change types	Class-level				Landscape-level		
	PD	LPI	SHAPE-MN	COHESION	CONTAG	SHEI	
I-O	2017	411.56	15.45	1.34	97.00	28.98	0.95
	2021	446.24	6.93	1.37	94.47	33.70	0.86
R-O	2017	510.16	2.19	1.27	94.46	37.71	0.82
	2021	253.34	5.26	1.37	96.77	36.20	0.91
B-O	2017	351.58	1.42	1.34	95.86	38.72	0.82
	2021	251.63	3.16	1.43	97.02	34.20	0.92
O-O	2017	250.58	3.65	1.38	97.81	29.31	1.00
	2021	243.89	2.10	1.42	97.04	31.30	0.98

7.4 Discussion

7.4.1 Differences in the distribution of urban green space

According to the analysis of urban green space distribution given in Section 7.3, diverse urban land use changes (e.g., human activities) had significantly different effects on urban green space distribution in Hangzhou from 2017 to 2021.

Generally, the percentage of urban green space coverage was 29% in 2017 and 38% in 2021, respectively, suggesting that the urban land use changes have important impacts on the urban green space system. Among them, the green space contribution from different urban land use changes was varied. It should be highlighted that residence and business parcels have the potential to have a greater impact on the urban green space system than other urban land use categories. To be more specific, the contribution of urban green space from residence parcels accounted for the lowest

proportion (24%) of total green space coverage in 2017, lower than that of institution parcels (31%), business parcels (26%), and open space parcels (46%), which is due to the lack of green infrastructure in the residence parcels. However, the urban residents have a strong willingness to access urban green space, which indicates the increase of the urban green space coverage in the residence parcels in 2021. The proportion of urban green space coverage in business parcels has increased from 26% in 2017 to 41% in 2021. The design of green space layout in business parcels focused on ornamental and entertainment functions rather than on ecological values. While the growth of urban green space coverage in institution parcels from 2017 to 2021 accounted for 9%, lower than that of residence parcels (12%) and business parcels (14%). Institution parcels usually have more space for urban green space construction compared with other urban land use types, indicating that they have more potential for urban green space development. As for the open space parcels, the proportion of urban green space coverage decreased from 46% to 36%. The contribution of open space parcels to total urban green space area was limited because the large proportion of open space in the urban periphery transformed into other land use types.

The distribution of urban green space among different urban land use changes varies, which has significant consequences for urban green planning and management. From the 1st ring belt to the 3rd ring belt, the overall proportion of urban green space coverage gradually increased in 2017, from 10%, 20%, to 31%. The percentage of green space cover from the 1st ring belt to the 3rd ring belt has also increased in 2021, from 25%, 34%, and 39%. It can be noted that the urban green space development in the urban environment especially the urban core is relatively well. This is not only due to the horizontal greenery management, but also relate to the vertical greenery including

green roofs, skyscraper farms, and sky gardens ([Larkin and Hystad 2019](#); [Ricci et al. 2022](#)). The implementation of vertical greenery solved the problem of the lack of space for greenery in some developing countries. In addition, the distribution of urban green space among different urban land use types was spatially changed at the ring level for 2017 and 2021. For instance, the total percentage of urban green space in institution parcels increased significantly from 1st ring belt to 3rd ring belt both in 2017 and 2021. This suggests that the closer to the urban periphery, the greater contribution of residential green space to the environment. A large number of the residence and business parcels are located in the urban periphery and the size of these business parcels is larger than that in the urban core. This is also related to the reason that the contribution of green space in the 3rd ring belt is greater than that in the 1st ring belt. However, the percentage of residential green space in the 1st ring belt increased from 2017 to 2021, indicating that greenery management is effective. By contrast, the proportion of green space in the open space parcels in the 3rd ring belt has decreased from 2017 to 2021, revealing the green space around the city periphery contributes to the construction of the other land use types. Accordingly, the urban green space in the business parcels should consider more environmental and ecological values, not aesthetic values only. In addition, the urban green space coverage from the four urban land use types showed an increasing trend from the 1st ring belt to the 3rd ring belt. This result revealed that the new regions in the urban periphery should have more green space coverage for the four urban land use types, while the old regions in the urban core should pay more attention to ecological urban green space design.

7.4.2 The development of urban green space distribution in China

Based on literature review and empirical evidence, this Ph.D. research explored the characteristics of the development of urban green space distribution in different stages ([Li et al. 2017a](#); [M’Ikiugu et al. 2012](#)). Urbanization and socio-economic change affect the development of urban green space distribution in different ways by urban planners or spontaneously formed based on people's real lifestyles ([Qian et al. 2015](#); [Zhou et al. 2018a](#)). As a result, the development stage of urban green space in different cities might also be different according to the urban hierarchy, the urban scale, or the economic status ([Wang et al. 2020](#)). In summary, the urban green space development in China can be divided into three stages in terms of urban green space distribution and patterns. An improved understanding of the development of urban green space distribution can guide urban planning and prioritize designs that fit current and future demands.

The level of urban green space development in Stage 1 is relatively high. According to [Wang et al. \(2020\)](#)'s work, the development of urban green space distribution in most first-tier cities can be categorized into Stage 1. In the beginning, the urban green spaces decrease with the enlargement of the city-scale under the background of accelerated urbanization and industrialization. The urban economy of cities in Stage 1 mainly depends on secondary industry (e.g., extensive manufacturing). During this period, a large number of urban green spaces have been converted into business land in the city periphery. In the meantime, the demand for residential land increases because of the massive increase in the urban population. After a period of development, the demand for urban green space increases gradually due to the transformation of the city's economic structure (e.g., from extensive manufacturing to

the service industry and high-tech industry) and the development of the optimization of the urban land use structure. Furthermore, the demand of urban residents for urban green spaces (public gardens, squares, and parks) has increased. As a result, urban green space began to increase in the main city, while the urban green space around the city still decreased with urbanization. Due to technical progress, although urban green space areas decrease continually, the decreasing rate is slowing.

In Stage 2, the level of urban green space development is medium. The development of urban green space distribution in most second-tier cities can be categorized into Stage 2 ([Haas and Ban 2014](#)). In the beginning, the urbanization in Stage 2 is proceeding slowly, which leads to a decrease in urban green space, especially in the urban core area. During this period, a large number of rural populations swarmed into cities, thus stimulating the urban demand for residential land and commercial land. However, the rate of urban expansion is less obvious and the urban green space around the city remains a large proportion. Subsequently, with the acceleration of urbanization, local governments introduce many industries for their political achievements. Therefore, the loss of urban green space is accelerated both in the urban core and urban periphery. The urban economy of cities in this period still depends on secondary industry (e.g., extensive manufacturing such as mining, manufacturing, electricity, and gas industry). The next step of urban green space planning for Stage 2 can refer to Stage 1.

The level of urban green space development in Stage 3 is relatively low. The development of urban green space distribution in most third-tier cities can be categorized into Stage 2 ([Zhou et al. 2018a](#)). In the beginning, the urbanization in Stage 3 is relatively low and urban green spaces remain stable. The urban economy of cities in this period mainly depends on the husbandry industry, light industry, and service

industry. The development of urban land use expands gradually due to the low rate of urbanization. Correspondingly, benefiting from the low-carbon economic growth and green lifestyle, the change in urban green space is not obvious. With the further decline of the economy in urban areas, the loss of employment opportunities is setting off partial out-migration. Further urban shrinkage may lead to further increases in urban green space.

7.5 Conclusion

This research investigated the impact of various urban land use changes (e.g., from open space to residence) on urban green space distribution by integrating statistical analysis, landscape analysis, and geospatial analysis. The results demonstrated that the urban land use changes significantly affect the distribution of urban green space in Hangzhou city. In general, urban green spaces increased from 2017 to 2021, and the growth areas were mainly concentrated in the urban core, indicating the city has made remarkable achievements in the planning of green spaces in the city center. Specifically, the increasing urban green space patches were distributed in the parcels that transformed from the business parcels, which showed more evenly distributed, and less aggregation. In addition, the proportion of urban green space patches in the unchanged residence parcels presented a significant increase, especially in the 1st ring belt. A large number of open space parcels have changed to business parcels around the city, while the urban green space changes within them were less obvious. The understanding of urban green landscapes could help provide regulations and guidelines for current and future planning of urban development.

Chapter 8: Synthesis

8.1 Research chapter summary

The objective of this Ph.D. research is to improve the existing urban land use maps for investigating the dynamics of urban land use and assess the impact of urban land use changes on urban green space distribution. The main findings of the research chapters (**Chapter 4**, **Chapter 5**, **Chapter 6**, and **Chapter 7**) are summarized as follows.

Chapter 4 has furthered the understanding of the integration of RS and GBD on urban land use mapping. The summarization of the existing literature concludes that the emerging GBD provides new opportunities for the transformation from urban land cover (e.g., physical environment) to urban land use (e.g., living environment). A deeper understanding of the urban surface can be acquired by adding GBD values to the traditional urban RS works. Specifically, the commonly used RS features (e.g., spectral, textural, temporal, and spatial features) and GBD features (e.g., spatial, temporal, semantic, and sequence features) were identified and analyzed in urban land use classification. In addition, this research categorized the various methods of RS and GBD integration used in urban land use mapping into decision-level integration (DI) and feature-level integration (FI). As the integration of RS and GBD has become more generalized, significant progress can be already seen in urban management (e.g., urban planning, urban environment assessment, urban disaster monitoring, and urban traffic

analysis). Integrating RS and GBD provides an opportunity for improving the existing urban land use maps.

Chapter 5 proposed a general framework to demonstrate and differentiate the DI and FI approaches and then proposed a methodology framework based on the general framework. The methodology framework was applied in Hangzhou city in 2019 for mapping urban land use types (e.g., institution, residence, business, and open space), based on OpenStreetMap (OSM) road network data, 10 m Sentinel-2A images, and Gaode Points of interest (POIs). The corresponding classification results were validated quantitatively and qualitatively using the same testing dataset. Overall, the results quantified the performance of the DI and FI methods for urban land use mapping and illustrated their advantages and disadvantages. It also should be noted that different data sources, selected features, classifiers, training samples, or land use types may also lead to different results. The diversity of urban land use types in Hangzhou city has a higher level of complexity than other regions in China, the mapping result is probably not representative of other regions. This result provides an improved understanding of urban land use mapping in terms of the RS and GBD integration strategy.

Chapter 6 classified the urban land use in 2017 and 2021 for Hangzhou city and investigated the spatial-temporal dynamics of urban land use changes from 2017 to 2021. This research improved the methodology framework proposed in **Chapter 5** by conducting an importance analysis of parcel attributes. The results demonstrated that the FI-based result (overall accuracy: 80.2%) performed better than the DI-based classification result (67.0%) in 2021 according to the quantitative analysis. Then the FI method was utilized to classify urban land use types in 2017, with an overall accuracy of 81.7%. The 2017 and 2021 maps were post-processed for obtaining high-accuracy

urban land use change products from 2017 to 2021. The results also showed that the number of institution and residence parcels increased, while the number of business and open space parcels decreased from 2017 to 2021. A large number of open space parcels were transformed into the residence and business parcels around the urban periphery. A few business parcels were changed into residence and institution parcels in the urban core area. This chapter analyzed the spatial-temporal patterns of urban land use changes in Hangzhou city, which could help investigate the urban green space distribution in response to urban land use changes.

Chapter 7 investigated the impact of various urban land use changes (e.g., from open space to residence) on urban green space distribution by integrating statistical analysis, landscape analysis, and geospatial analysis. The results demonstrated that the urban land use changes significantly affect the distribution of urban green space in Hangzhou city. In general, urban green spaces increased from 2017 to 2021, and the growth areas were mainly concentrated in the urban core, indicating the city has made remarkable achievements in the planning of green spaces in the city center. Specifically, the increasing urban green space patches were distributed in the parcels that transformed from the business parcels, which showed more evenly distributed, and less aggregation. In addition, the proportion of urban green space patches in the unchanged residence parcels presented a significant increase, especially in the 1st ring belt. A large number of open space parcels have changed to business parcels around the city, while the urban green space changes within them were less obvious. The understanding of urban green landscapes could help provide regulations and guidelines for current and future planning of urban development.

8.2 Future work

8.2.1 The improved method for urban land use mapping

Given the methodology framework, quantitative and qualitative assessments of the DI-based and FI-based classification methods (shown in **Chapter 5**), it is possible to highlight some points for improving the integration methods of urban land use mapping in future work.

(1) With the development of urban areas around the world increasing at an exponential rate, new images are needed for helping with efficient urban land use planning. VHR imagery from satellites is the most effective way to gain an accurate picture of landscapes of the urban environment's entirety and complexity ([Xu et al. 2021b](#)). VHR satellite images provide abundant information (e.g., textual, contextual, and spectral information) about geographical objects in terms of geometrical features and spatial patterns ([Du et al., 2021](#); [Zhang et al., 2017b](#)[Zhao et al. 2019](#)). These images are so detailed that users have been using them to find abandoned warehouses and empty spaces inside cities ([Morell-Monzó et al. 2021](#)). In addition, more spatial details can be obtained from VHR images compared with medium and low-resolution satellite images such as AVHRR, Landsat, and SPOT images ([Xu et al. 2021b](#)). VHR imagery represents a significant advance in categorizing urban land use and will be of great use for urban planning and management.

(2) OSM data contain multiple and finer information on urban land use, especially social functional information that is rarely captured from RS images ([Wang et al. 2021b](#)). Each functional unit delineated by OSM data could serve as the most basic unit for urban land use analysis because of its greatest degree of functional homogeneity

and the greatest degree of functional heterogeneity with surrounding units ([Vu et al. 2021](#)). The basic units make it possible to efficiently delineate urban land properties and discover the distribution characteristics of different land use types and further assist with urban spatial planning and urban management. However, the OSM road network might contain unbalanced classes that are unevenly distributed, leading to heterogeneous urban land use parcels in terms of parcel area and spatial pattern ([Johnson and Iizuka 2016](#)). Several studies have tried to improve the identification of urban land use parcels by adding other data sources such as transportation data, social media data, and so on ([Vargas Muñoz et al. 2021](#)). Furthermore, the accuracy of the urban parcels created based on OSM can be further improved by combining different data sources, and the veracity of urban parcels should be validated.

(3) The training and testing samples are important for urban land use classification because different classifiers and machine learning algorithms highly depend on the input reference with varying qualities and quantities. Generally, these samples used for urban land use classification were visually identified through the Baidu map, Baidu Street View, and a field survey, which are promising to validate urban land use parcels ([Hu et al. 2016](#)). However, sample collection is the most time-consuming step and it is inevitable for urban land use parcels to contain multiple land use types. The automatic sampling method can label samples within urban scenes that consist of heterogeneous urban land uses that can hardly be resolved by classical sampling methods ([Zhang et al. 2020c](#)). In this regard, sample labeling for urban land use classification should not only depend on visual interpretation but also consider automatic approaches to guarantee high accuracy. In addition, better classification performance would be expected from

the adoption of an independent land use map by local government and urban planners ([Pan et al. 2020](#)).

(4) The RF model was used for mapping built-up in DI-based classification and training urban land use parcels in FI-based classification ([Wu et al. 2021a](#); [Wu et al. 2021b](#)). The RF model can be improved by using the transfer learning model as a pre-trained model, which might help increase the classification accuracy ([Kang et al. 2021](#)). Furthermore, how to integrate GBD into the RF model is a worthwhile question to discuss since there is a significant difference between RS and GBD ([Mao et al. 2020](#)). Recently, great progress in deep learning techniques has been made in computer vision applications, and breakthroughs in semantic segmentation tasks especially facilitate the urban land use classification problem ([Chen and Tsou 2021](#)). Deep learning-based approaches have rapidly developed in the field of urban land use classification because of their capability of transforming the raw input imagery and low-level features into classes at a higher and abstract level ([Xu et al. 2021a](#)). The unique deep hierarchical structure of deep learning algorithms has proven its powerful capability of learning and generalization using deep convolutional neural networks (CNNs), which is suitable for urban land use classification.

(5) The DI-based and FI-based classification methods can be used together for improving urban land use maps as they have different performances in classifying different types of land use (e.g., institution, residence, business, and open space) ([Yin et al. 2021a](#); [Yin et al. 2021b](#)). To be more specific, the FI-based method can be used for classifying very different urban land use types. It can be noted from the 2021 DI-based classification result that the institution and business in FI-based classification can be classified relatively well, with the UA of 0.826 ± 0.108 and 0.960 ± 0.038 , respectively.

While the DI-based method can be utilized for classifying urban land use types with very little difference in RS features but have a significant difference in GBD features such as institution parcels and business parcels. According to the urban land use classification result in Hangzhou city for 2021, the residence and business were classified relatively well in the DI-based urban land use map, with the estimated UA of 0.892 ± 0.048 and 0.709 ± 0.082 , respectively.

8.2.2 The mixed urban land use parcels

Urban land use mix refers to a situation in which different urban land use types of land uses co-exist within the same space (e.g., residential commercial mix building) (Tu et al. 2021). Under the current shortage of land resources and the tight supply of urban land in areas of rapid urbanization (Li 2021), the proposed urban mixed space utilization is of great significance for improving the efficiency of urban land space use and reducing resource consumption. However, the mixing of urban land use parcels has also caused significant problems in cities all over the world, especially in China (Lang et al. 2018).

From a historical perspective of development, the idea of urban land use mix began to spring up in China in the 1990s to solve the problems of traffic congestions, the decline of city centers, and social isolation (Shi et al. 2022). The mixed-use of urban land has its inevitable requirements. First, the rapid pace of urbanization necessitates a diversity of urban building functions, and single land uses find it difficult to adapt to changing urban land requirements (Sun et al. 2022). Second, with the arrival of the Internet era, the original industrial structure has gained many new functions, resulting in the formation of several new industries as well as the deep integration of cross-industry activities (Fang et al. 2021). Through the mixing of different urban land use

types, it gets rid of the traditional urban planning method and expands the urban space to various dimensions.

Mixed land use has been a big challenge in mapping urban land use that integrates multiple functional uses ([Bodhankar et al. 2022](#); [He et al. 2021](#)). The differentiation from the mixed land uses will be quite difficult if the spatial resolution of the input data sources and characteristics is not equal to or significantly lower than the spatial resolution of the targeted objects ([Jiao et al. 2021](#)). Previous research solely looked at "dominant land use" rather than "proportional land use," eliminating a wealth of information on the microstructure of urban land use ([Yao et al. 2022](#)).

Future research can overcome the inadequacies of mixed urban land use recognition in two important directions. First, the basic units for urban land use classification should be generated with the greatest degree of functional homogeneity and the greatest degree of functional heterogeneity. In order to obtain fine-resolution urban land use parcels, more detailed road network data or image segmentation can be used to delineate the urban areas. Second, traffic data, cell phone data, social media data, and municipal data are among the new GBD that may be used to estimate mixed urban land use patterns. In addition, vertical mixed land uses continue to be a concern because of the restricted options of partitioning different land use categories inside building groups.

8.2.3 The configuration and composition of urban green space

The configuration and composition of urban green space were not considered accordingly in this research. Studies have shown that the ability of urban green spaces to provide their expected environmental and socio-economic benefits largely depends

on their distributions, compositions, and configurations ([Woldesemayat and Genovese 2021](#)).

The configuration generally refers to the shape, size, and structure of urban green spaces ([Coban et al. 2021](#)). With the development of urbanization, the configurations of urban green space are likely to have a variety of changes in different urban functional zones ([Threlfall et al. 2016](#)). These changes in urban green spaces significantly influence the ecology of cities, yet they are generally poorly understood in urban landscapes. To be more specific, the configurations of urban green space are changed by various human activities within different urban functional regions, such as exotic plant invasion and the changed traits of native plants, which play important roles in urban ecosystem functions ([Miralles-Guasch et al. 2019](#)). The ecological sustainability and function of urban landscapes are strongly influenced by the configuration and structure of the urban green spaces. Therefore, further studies that investigate the urban green space configurations of different urban land use and their variations are needed.

Urban green space composition refers to the identity of the species and types comprising the urban green spaces (e.g., grass, trees, shrubs, flowers). Numerous studies have analyzed the relationships between urbanization and urban green space composition ([Ghosh and Das 2018](#); [Grafius et al. 2018](#)). However, it is still unclear what impact inner cities have on the urban green space compositions because of the increased complexity of the urban landscape. Further understanding of variations in the composition of urban green space within different urban land use types will greatly improve our ability to create urban landscapes that enhance biodiversity in the urban environment.

In addition, urban functional green spaces such as community gardens, parks, and greenways are highly correlated with socio-economic activities. Urban functional green space maps produced by using high-resolution images and deep learning techniques can be utilized to provide suggestions for future urban planning, design, and management. Furthermore, The accessibility of different urban functional green spaces can be considered in the future study.

8.3 Concluding remarks

This Ph.D. research improved the existing urban land use maps for investigating the dynamics of urban land use and assessed the impact of urban land use changes on urban green space distribution. Among them, the research confirms that the integration of RS and GBD provides new opportunities for urban land use mapping. RS data, coupled with GBD, can combine the socio-economic features with the spatial and temporal coverage afforded by RS. In addition, the urban green space distribution was significantly affected by the urban land use change in Hangzhou city. It is hoped that the practical research and experience of urban green space distribution in this thesis will provide valuable suggestions for small cities in developing their urban green space system for the city, and make the city achieve the SDGs.

References:

Adiri, Z., Lhissou, R., El Harti, A., Jellouli, A., & Chakouri, M. (2020). Recent advances in the use of public domain satellite imagery for mineral exploration: A review of Landsat-8 and Sentinel-2 applications. *Ore Geology Reviews*, *117*, 103332-103365.

Ali, A.L., Falomir, Z., Schmid, F., & Freksa, C. (2017). Rule-guided human classification of Volunteered Geographic Information. *ISPRS Journal of Photogrammetry and Remote Sensing*, *127*, 3-15.

Amoatey, P., Sulaiman, H., Kwarteng, A., & Al-Reasi, H.A. (2018). Above-ground carbon dynamics in different arid urban green spaces. *Environmental Earth Sciences*, *77*, 1-10.

Andrade, R., Alves, A., & Bento, C. (2020). POI mining for land use classification: A case study. *ISPRS International Journal of Geo-Information*, *9*, 493-514.

Angel, S., Parent, J., Civco, D.L., Blei, A., & Potere, D. (2011). The dimensions of global urban expansion: estimates and projections of all countries, 2000–2050. *Progress in planning*, *75*, 53-107.

Anugraha, A.S., Chu, H.J., & Ali, M.Z. (2020). Social sensing for urban land use identification. *ISPRS International Journal of Geo-Information*, *9*, 550-569.

Aram, F., García, E.H., Solgi, E., & Mansournia, S. (2019). Urban green space cooling effect in cities. *Heliyon*, *5*, e01339.

Bao, H.Q., Ming, D.P., Guo, Y., Zhang, K., Zhou, K.Q., & Du, S.G. (2020). DFCNN-based semantic recognition of urban functional zones by integrating remote sensing data and POI data. *Remote Sensing*, *12*, 1088-1114.

Bao, T., Li, X.M., Zhang, J., Zhang, Y.J., & Tian, S.Z. (2016). Assessing the distribution of urban green spaces and its anisotropic cooling distance on urban heat island pattern in Baotou, China. *ISPRS International Journal of Geo-Information*, *5*, 12-25.

Batty, M. (2013). Big data, smart cities and city planning. *Dialogues in Human Geography*, *3*, 274-279.

Bodhankar, S., Gupta, K., Kumar, P., & Srivastav, S.K. (2022). GIS-based multi-objective urban land allocation approach for optimal allocation of urban land uses. *Journal of the Indian Society of Remote Sensing*, *50*, 763–774.

Bosch, M., Jaligot, R., & Chenal, J. (2020). Spatiotemporal patterns of urbanization in three Swiss urban agglomerations: insights from landscape metrics, growth modes and fractal analysis. *Landscape Ecology*, *35*, 879-891.

Braubach, M., Egorov, A., Mudu, P., Wolf, T., Thompson, C.W., & Martuzzi, M. (2017). Effects of urban green space on environmental health, equity and resilience. *Nature-based solutions to climate change adaptation in urban areas* (pp. 187-205): Springer, Cham.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5-32.

Byomkesh, T., Nakagoshi, N., & Dewan, A.M. (2011). Urbanization and green space dynamics in Greater Dhaka, Bangladesh. *Landscape and Ecological Engineering*, 8, 45-58.

Cadavid Restrepo, A.M., Yang, Y.R., Hamm, N.A.S., Gray, D.J., Barnes, T.S., Williams, G.M., Soares Magalhaes, R.J., McManus, D.P., Guo, D., & Clements, A.C.A. (2017). Land cover change during a period of extensive landscape restoration in Ningxia Hui Autonomous Region, China. *Science of the Total Environment*, 598, 669-679.

Cai, J.X., Huang, B., & Song, Y.M. (2017). Using multi-source geospatial big data to identify the structure of polycentric cities. *Remote Sensing of Environment*, 202, 210-221.

Callaghan, A., McCombe, G., Harrold, A., McMeel, C., Mills, G., Moore-Cherry, N., & Cullen, W. (2021). The impact of green spaces on mental health in urban settings: a scoping review. *Journal of Mental Health*, 30, 179-193.

Campbell, A. (2009). The rise of people-centric sensing. In V. Garg, R. Wattenhofer, & K. Kothapalli (Eds.), *Distributed Computing and Networking* (pp. 9-9). Berlin, Heidelberg: Springer Berlin Heidelberg.

Cao, R., Tu, W., Yang, C.X., Li, Q., & Qiu, G.P. (2020). Deep learning-based remote and social sensing data fusion for urban region function recognition. *ISPRS Journal of Photogrammetry and Remote Sensing*, 163, 82-97.

Cao, R., Zhu, J.S., Tu, W., Li, Q.Q., Cao, J.Z., Liu, B.Z., Zhang, Q., & Qiu, G.P. (2018). Integrating aerial and street view images for urban land use classification. *Remote Sensing*, *10*, 1553-1576.

Card, D.H. (1982). Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogrammetric Engineering and Remote Sensing*, *48*, 431-439.

Caspersen, O.H., Konijnendijk, C.C., & Olafsson, A.S. (2006). Green space planning and land use: An assessment of urban regional and green structure planning in Greater Copenhagen. *Geografisk Tidsskrift-Danish Journal of Geography*, *106*, 7-20.

Cervone, G., Sava, E., Huang, Q.Y., Schnebele, E., Harrison, J., & Waters, N. (2015). Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study. *International Journal of Remote Sensing*, *37*, 100-124.

Chan, K.M., & Vu, T.T. (2017). A landscape ecological perspective of the impacts of urbanization on urban green spaces in the Klang Valley. *Applied Geography*, *85*, 89-100.

Chang, C., Ye, Z.W., Huang, Q.Y., & Wang, C.X. (2015). An integrative method for mapping urban land use change using "geo-sensor" data. In, *1st International ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics* (pp. 47-54)

Chang, S.Z., Wang, Z.M., Mao, D.H., Guan, K.H., Jia, M.M., & Chen, C.Q. (2020). Mapping the essential urban land use in Changchun by applying Random Forest and Multi-Source Geospatial Data. *Remote Sensing*, *12*, 2488-2507.

Chen, B., Song, Y.M., Huang, B., & Xu, B. (2020a). A novel method to extract urban human settlements by integrating remote sensing and mobile phone locations. *Science of Remote Sensing*, *1*, 100003.

Chen, B., Tu, Y., Song, Y.M., Theobald, D.M., Zhang, T., Ren, Z.H., Li, X.C., Yang, J., Wang, J., Wang, X., Gong, P., Bai, Y.Q., & Xu, B. (2021). Mapping essential urban land use categories with open big data: Results for five metropolitan areas in the United States of America. *ISPRS Journal of Photogrammetry and Remote Sensing*, *178*, 203-218.

Chen, F.H., & Tsou, J.Y. (2021). Mapping urban form and land use with deep learning techniques: A case study of Dongguan City, China. *International Journal of Oil, Gas and Coal Technology*, *29*, 306–328.

Chen, T., Huang, Q.H., Liu, M., Li, M.C., Qu, L.A., Deng, S.L., & Chen, D. (2017). Decreasing net primary productivity in response to urbanization in Liaoning Province, China. *Sustainability*, *9*, 162-179.

Chen, W., Huang, H.P., Dong, J.W., Zhang, Y., Tian, Y.C., & Yang, Z.Q. (2018). Social functional mapping of urban green space using remote sensing and social sensing data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *146*, 436-452.

Chen, W., Zhou, Y.Y., Wu, Q.S., Chen, G., Huang, X., & Yu, B.L. (2020b). Urban building type mapping using geospatial data: A case study of Beijing, China. *Remote Sensing*, *12*, 2805-2823.

Cheng, L., Yuan, Y., Xia, N., Chen, S., Chen, Y.M., Yang, K., Ma, L., & Li, M.C. (2018). Crowd-sourced pictures geo-localization method based on street view images and 3D reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, *141*, 72-85.

Çoban, S., Yener, Ş.D., & Bayraktar, S. (2021). Woody plant composition and diversity of urban green spaces in Istanbul, Turkey. *Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology*, *155*, 83-91.

Daniels, B., Zaunbrecher, B.S., Paas, B., Ottermanns, R., Ziefle, M., & Roß-Nickoll, M. (2018). Assessment of urban green space structures and their quality from a multidimensional perspective. *Science of the Total Environment*, *615*, 1364-1378.

Daz, B.S., Ghaffari Gilandeh, A., & Azizi, A. (2020). Analyzing of urban green space changes in Gorgan City using landscape metrics. *Journal of Environmental Science and Technology*, *22*, 167-181.

de Jalón, S.G., Chiabai, A., Quiroga, S., Suárez, C., Ščasný, M., Máca, V., Zvěřinová, I., Marques, S., Craveiro, D., & Taylor, T. (2021). The influence of urban greenspaces on people's physical activity: A population-based study in Spain. *Landscape and Urban Planning*, *215*, 104229.

Defries, R.S., & Townshend, J.R.G. (2007). NDVI-derived land cover classifications at a global scale. *International Journal of Remote Sensing*, *15*, 3567-3586.

Deng, J.S., Huang, Y.B., Chen, B.J., Tong, C., Liu, P.B., Wang, H.Q., & Hong, Y. (2019). A methodology to monitor urban expansion and green space change using a time series of Multi-Sensor SPOT and Sentinel-2A images. *Remote Sensing*, *11*, 1230-1249.

Deng, J.S., Wang, K., Deng, Y.H., & Qi, G.J. (2008). PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data. *International Journal of Remote Sensing*, *29*, 4823-4838.

Dong, J.W., Metternicht, G., Hostert, P., Fensholt, R., & Chowdhury, R.R. (2019). Remote sensing and geospatial technologies in support of a normative land system science: status and prospects. *Current Opinion in Environmental Sustainability*, *38*, 44-52.

Dong, X.Y., Xu, Y.D., Huang, L.P., Liu, Z.G., Xu, Y., Zhang, K.Y., Hu, Z.W., & Wu, G.F. (2020). Exploring impact of spatial unit on urban land use mapping with multisource data. *Remote Sensing*, *12*, 3597-3616.

Donner, A., Shoukri, M.M., Klar, N., & Bartfay, E. (2000). Testing the equality of two dependent kappa statistics. *Statistics in medicine*, *19*, 373-387.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O.,

Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, *120*, 25-36.

Du, S.J., Du, S.H., Liu, B., Zhang, X.Y., & Zheng, Z.J. (2020). Large-scale urban functional zone mapping by integrating remote sensing images and open social data. *Giscience & Remote Sensing*, *57*, 411-430.

Eckert, S., Kohler, S. (2014). Urbanization and health in developing countries: a systematic review. *World Health and Population*, *15*, 7-20.

Eisenman, T.S., Churkina, G., Jariwala, S.P., Kumar, P., Lovasi, G.S., Pataki, D.E., Weinberger, K.R., & Whitlow, T.H. (2019). Urban trees, air quality, and asthma: An interdisciplinary review. *Landscape and Urban Planning*, *187*, 47-59.

Ellis, E.C., & Ramankutty, N. (2008). Putting people in the map: anthropogenic biomes of the world. *Frontiers in Ecology and the Environment*, *6*, 439-447.

Elmqvist, T., Andersson, E., Frantzeskaki, N., McPhearson, T., Olsson, P., Gaffney, O., Takeuchi, K., & Folke, C. (2019). Sustainability and resilience for transformation in the urban century. *Nature Sustainability*, *2*, 267-273.

Esch, T., Taubenboeck, H., Roth, A., Heldens, W., Felbier, A., Thiel, M., Schmidt, M., Mueller, A., & Dech, S. (2012). TanDEM-X mission-new perspectives for the inventory and monitoring of global settlement patterns. *Journal of Applied Remote Sensing*, *6*, 061702.

Fang, F., Yu, Y.F., Li, S.W., Zuo, Z.J., Liu, Y.Y., Wan, B., & Luo, Z.W. (2021). Synthesizing location semantics from street view images to improve urban land-use

classification. *International Journal of Geographical Information Science*, 35, 1802-1825.

Feng, X.Q., & Astell-Burt, T. (2018). Residential green space quantity and quality and symptoms of psychological distress: a 15-year longitudinal study of 3897 women in postpartum. *BMC psychiatry*, 18, 1-11.

Fischer, F. (2012). VGI as Big Data: A new but delicate geographic data-source. *GeoInformatics*, 15, 46-47.

Flouri, E., Papachristou, E., & Midouhas, E. (2019). The role of neighbourhood greenspace in children's spatial working memory. *British journal of educational psychology*, 89, 359-373.

Foody, G.M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80, 185-201.

Foody, G.M. (2004). Thematic map comparison. *Photogrammetric Engineering and Remote Sensing*, 70, 627-633.

Frias-Martinez, V., & Frias-Martinez, E. (2014). Spectral clustering for sensing urban land use using Twitter activity. *Engineering Applications of Artificial Intelligence*, 35, 237-245.

Gascon, M., Sánchez-Benavides, G., Dadvand, P., Martínez, D., Gramunt, N., Gotsens, X., Cirach, M., Vert, C., Molinuevo, J.L., & Crous-Bou, M. (2018). Long-term exposure to residential green and blue spaces and anxiety and depression in adults: A cross-sectional study. *Environmental research*, 162, 231-239.

Gavrilidis, A.A., Niță, M.R., Onose, D.A., Badiu, D.L., & Năstase, I.I. (2019). Methodological framework for urban sprawl control through sustainable planning of urban green infrastructure. *Ecological Indicators*, *96*, 67-78.

Genuer, R., Poggi, J.M., & Tuleau-Malot, C. (2015). VSURF: an R package for variable selection using random forests. *The R Journal*, *7*, 19-33.

Ghosh, S., & Das, A. (2018). Modelling urban cooling island impact of green space and water bodies on surface urban heat island in a continuously developing urban area. *Modeling Earth Systems and Environment*, *4*, 501-515.

Goffi, A., Bordogna, G., Stroppiana, D., Boschetti, M., & Brivio, P.A. (2020). Knowledge and data-driven mapping of environmental status indicators from remote sensing and VGI. *Remote Sensing*, *12*, 495-514.

Goldewijk, K.K. (2001). Estimating global land use change over the past 300 years: The HYDE Database. *Global Biogeochemical Cycles*, *15*, 417-433.

Gong, L., Liu, X., Wu, L., & Liu, Y. (2015). Inferring trip purposes and uncovering travel patterns from taxi trajectory data. *Cartography and Geographic Information Science*, *43*, 103-114.

Gong, P., Chen, B., Li, X.C., Liu, H., Wang, J., Bai, Y.Q., Chen, J.M., Chen, X., Fang, L., Feng, S.L., Feng, Y.J., Gong, Y.L., Gu, H., Huang, H.B., Huang, X.C., Jiao, H.Z., Kang, Y.D., Lei, G.B., Li, A.N., Li, X.T., Li, X., Li, Y.C., Li, Z.L., Li, Z.D., Liu, C., Liu, C.X., Liu, M.C., Liu, S.G., Mao, W.L., Miao, C.H., Ni, H., Pan, Q.S., Qi, S.H., Ren, Z.H., Shan, Z.R., Shen, S.Q., Shi, M.J., Song, Y.M., Su, M., Suen, H.P., Sun, B.,

Sun, F.D., Sun, J., Sun, L., Sun, W.Y., Tian, T., Tong, X.H., Tseng, Y., Tu, Y., Wang, H., Wang, L., Wang, X., Wang, Z.M., Wu, T.H., Xie, Y.W., Yang, J., Yang, J., Yuan, M., Yue, W.Z., Zeng, H.D., Zhang, K., Zhang, N., Zhang, T., Zhang, Y., Zhao, F., Zheng, Y.C., Zhou, Q.M., Clinton, N., Zhu, Z.L., & Xu, B. (2020). Mapping essential urban land use categories in China (EULUC-China): preliminary results for 2018. *Science Bulletin*, *65*, 182-187.

Gong, P., & Howarth, P.J. (1989). Performance analyses of probabilistic relaxation methods for land-cover classification. *Remote Sensing of Environment*, *30*, 33-42.

Gong, P., Li, X.C., & Zhang, W. (2019). 40-Year (1978–2017) human settlement changes in China reflected by impervious surfaces from satellite remote sensing. *Science Bulletin*, *64*, 756-763.

Goodchild, M.F., & Glennon, J.A. (2010). Crowdsourcing geographic information for disaster response: a research frontier. *International Journal of Digital Earth*, *3*, 231-241.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18-27.

Grafius, D.R., Corstanje, R., & Harris, J.A. (2018). Linking ecosystem services, urban form and green space configuration using multivariate landscape metric analysis. *Landscape Ecology*, *33*, 557-573.

Grippa, T., Georganos, S., Zarougui, S., Bognounou, P., Diboulo, E., Forget, Y., Lennert, M., Vanhuysse, S., Mboga, N., & Wolff, E. (2018). Mapping urban land use at street block level using OpenStreetMap, remote sensing data, and spatial metrics. *ISPRS International Journal of Geo-Information*, 7, 246-267.

Haas, J., & Ban, Y.F. (2014). Urban growth and environmental impacts in Jing-Jin-Ji, the Yangtze, River Delta and the Pearl River Delta. *International Journal of Applied Earth Observation and Geoinformation*, 30, 42-55.

Hagenauer, J.L., & Helbich, M. (2012). Mining urban land-use patterns from volunteered geographic information by means of genetic algorithms and artificial neural networks. *International Journal of Geographical Information Science*, 26, 963-982.

Han, Z.M., Dian, Y.Y., Xia, H., Zhou, J.J., Jian, Y.F., Yao, C.H., Wang, X., & Li, Y. (2020). Comparing fully deep convolutional neural networks for land cover classification with high-spatial-resolution Gaofen-2 images. *ISPRS International Journal of Geo-Information*, 9, 478.

Hao, J.W., Zhu, J., & Zhong, R. (2015). The rise of big data on urban studies and planning practices in China: Review and open research issues. *Journal of Urban Management*, 4, 92-124.

Haralick, R.M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *Studies in Media and Communication*, 3, 610-621.

Harvey, F. (2013). To volunteer or to contribute locational information? Towards truth in labeling for Crowdsourced Geographic Information. In D. Sui, S. Elwood, & M. Goodchild (Eds.), *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice* (pp. 31-42). Dordrecht: Springer Netherlands.

He, D., Shi, Q., Liu, X., Zhong, Y., & Zhang, X. (2021). Deep subpixel mapping based on semantic information modulated network for urban land use mapping. *IEEE Transactions on Geoscience and Remote Sensing*, *59*, 10628-10646.

Heiden, U., Segl, K., Roessner, S., & Kaufmann, H. (2007). Determination of robust spectral features for identification of urban surface materials in hyperspectral remote sensing data. *Remote Sensing of Environment*, *111*, 537-552.

Heidt, V., & Neef, M. (2008). Benefits of urban green space for improving urban climate. *Ecology, Planning, and Management of Urban Forests* (pp. 84-96): Springer.

Helbich, M., Amelunxen, C., Neis, P., & Zipf, A. (2012). Comparative spatial analysis of positional accuracy of OpenStreetMap and proprietary Geodata. In, *GI_Forum 2012: Geovisualization, Society and Learning, Salzburg, Germany* (pp. 4–6)

Hepcan, Ş. (2012). Analyzing the pattern and connectivity of urban green spaces: A case study of Izmir, Turkey. *Urban Ecosystems*, *16*, 279-293.

Hernández-Moreno, Á., & Reyes-Paecke, S. (2018). The effects of urban expansion on green infrastructure along an extended latitudinal gradient (23°S–45°S) in Chile over the last thirty years. *Land Use Policy*, *79*, 725-733.

Herold, M., Scepan, J., & Clarke, K.C. (2002). The use of remote sensing and landscape metrics to describe structures and changes in urban land uses. *Environment and planning A*, 34, 1443-1458.

Hersperger, A.M., Oliveira, E., Pagliarin, S., Palka, G., Verburg, P., Bolliger, J., & Grădinaru, S. (2018). Urban land-use change: The role of strategic spatial planning. *Global Environmental Change*, 51, 32-42.

Hoffmann, E.J., Wang, Y.Y., Werner, M., Kang, J., & Zhu, X.X. (2019). Model fusion for building type classification from Aerial and Street View Images. *Remote Sensing*, 11, 1259-1279.

Howarth, P.J., & Boasson, E. (1983). Landsat digital enhancements for change detection in urban environments. *Remote Sensing of Environment*, 13, 149-160.

Hu, T.Y., Yang, J., Li, X.C., & Gong, P. (2016). Mapping urban land use by using Landsat images and open social data. *Remote Sensing*, 8, 151-169.

Hu, Y.J., Gao, S., Janowicz, K., Yu, B.L., Li, W.W., & Prasad, S. (2015). Extracting and understanding urban areas of interest using geotagged photos. *Computers Environment and Urban Systems*, 54, 240-254.

Huang, B., Zhao, B., & Song, Y.M. (2018a). Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. *Remote Sensing of Environment*, 214, 73-86.

Huang, C.B., Huang, P., Wang, X.S., & Zhou, Z.X. (2018b). Assessment and optimization of green space for urban transformation in resources-based city—A case study of Lengshuijiang city, China. *Urban Forestry & Urban Greening*, 30, 295-306.

Huang, Q.Y., & Wong, D.W.S. (2016). Activity patterns, socioeconomic status and urban spatial structure: what can social media data tell us? *International Journal of Geographical Information Science*, 30, 1873-1898.

Huang, R., Taubenbock, H., Mou, L.C., & Zhu, X.X. (2018c). Classification of settlement types from tweets using LDA and LSTM. In, *IEEE International Geoscience and Remote Sensing Symposium* (pp. 6408-6411)

Huang, X., Huang, J.Y., Wen, D.W., & Li, J.Y. (2021). An updated MODIS global urban extent product (MGUP) from 2001 to 2018 based on an automated mapping approach. *International Journal of Applied Earth Observation and Geoinformation*, 95, 102255.

Huang, X., & Wang, Y. (2019). Investigating the effects of 3D urban morphology on the surface urban heat island effect in urban functional zones by using high-resolution remote sensing data: A case study of Wuhan, Central China. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, 119-131.

Huang, Z., Qi, H.J., Kang, C.G., Su, Y.L., & Liu, Y. (2020). An ensemble learning approach for urban land use mapping based on remote sensing imagery and social sensing data. *Remote Sensing*, 12

Hussainzad, E.A., Yusof, M.J.M., & Maruthaveeran, S. (2021). Identifying women's preferred activities and elements of private green spaces in informal settlements of Kabul city. *Urban Forestry & Urban Greening*, 59, 127011.

Ilieva, R.T., & McPhearson, T. (2018). Social-media data for urban sustainability. *Nature Sustainability*, 1, 553-565.

Inkoom, J.N., Frank, S., Greve, K., Walz, U., & Fürst, C. (2018). Suitability of different landscape metrics for the assessments of patchy landscapes in West Africa. *Ecological Indicators*, 85, 117-127.

Jennings, V., & Bamkole, O. (2019). The relationship between social cohesion and urban green space: An avenue for health promotion. *International Journal of Environmental Research and Public Health*, 16, 452.

Jepsen, M.R., & Levin, G. (2013). Semantically based reclassification of Danish land-use and land-cover information. *International Journal of Geographical Information Science*, 27, 2375-2390.

Jia, Y.X., Ge, Y., Ling, F., Guo, X., Wang, J.H., Wang, L., Chen, Y.H., & Li, X.D. (2018). Urban land use mapping by combining remote sensing imagery and mobile phone positioning data. *Remote Sensing*, 10, 446-467.

Jiang, Q.J., Kresin, F., Bregt, A.K., Kooistra, L., Pareschi, E., van Putten, E., Volten, H., & Wesseling, J. (2016). Citizen sensing for improved urban environmental monitoring. *Journal of Sensors*, 2016, 1-9.

Jiao, J.C., Rollo, J., & Fu, B.B. (2021). The hidden characteristics of land-use mix indices: An overview and validity analysis based on the land use in Melbourne, Australia. *Sustainability*, *13*, 1898.

Jiao, L.M., Xu, G., Xiao, F.T., Liu, Y.L., & Zhang, B.E. (2017). Analyzing the impacts of urban expansion on green fragmentation using constraint gradient analysis. *The Professional Geographer*, *69*, 553-566.

Jin, H.R., Mountrakis, G., & Stehman, S.V. (2014). Assessing integration of intensity, polarimetric scattering, interferometric coherence and spatial texture metrics in PALSAR-derived land cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *98*, 70-84.

Johnson, B.A., & Iizuka, K. (2016). Integrating OpenStreetMap crowdsourced data and Landsat time-series imagery for rapid land use/land cover (LULC) mapping: Case study of the Laguna de Bay area of the Philippines. *Applied Geography*, *67*, 140-149.

Johnson, B.A., Iizuka, K., Bragais, M.A., Endo, I., & Magcale-Macandog, D.B. (2017). Employing crowdsourced geographic data and multi-temporal/multi-sensor satellite imagery to monitor land cover change: A case study in an urbanizing region of the Philippines. *Computers, Environment and Urban Systems*, *64*, 184-193.

Kang, J., Körner, M., Wang, Y.Y., Taubenböck, H., & Zhu, X.X. (2018). Building instance classification using street view images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *145*, 44-59.

Kang, Y., Cho, N., Yoon, J., Park, S., & Kim, J. (2021). Transfer learning of a deep learning model for exploring tourists' urban image using geotagged photos. *ISPRS International Journal of Geo-Information*, *10*, 137.

Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*, *3*, 1-10.

Koch, F., & Krellenberg, K. (2018). How to contextualize SDG 11? Looking at indicators for sustainable urban development in Germany. *ISPRS International Journal of Geo-Information*, *7*, 464.

Kondo, M.C., Fluehr, J.M., McKeon, T., & Branas, C.C. (2018). Urban green space and its impact on human health. *International Journal of Environmental Research and Public Health*, *15*, 445.

Kong, F.H., & Nakagoshi, N. (2006). Spatial-temporal gradient analysis of urban green spaces in Jinan, China. *Landscape and Urban Planning*, *78*, 147-164.

Kothencz, G., Kolcsár, R., Cabrera-Barona, P., & Szilassi, P. (2017). Urban green space perception and its contribution to well-being. *International Journal of Environmental Research and Public Health*, *14*, 766.

Krylov, V., Kenny, E., & Dahyot, R. (2018). Automatic discovery and geotagging of objects from street view imagery. *Remote Sensing*, *10*, 661-681.

Kuang, W.H., & Dou, Y.Y. (2020). Investigating the patterns and dynamics of urban green space in China's 70 major cities using satellite remote sensing. *Remote Sensing*, *12*, 1929-1943.

Kumar, M., Denis, D.M., Singh, S.K., Szabó, S., & Suryavanshi, S. (2018). Landscape metrics for assessment of land cover change and fragmentation of a heterogeneous watershed. *Remote Sensing Applications: Society and Environment*, *10*, 224-233.

Kwon, O.H., Hong, I., Yang, J., Wohn, D.Y., Jung, W.S., & Cha, M. (2021). Urban green space and happiness in developed countries. *EPJ Data Science*, *10*, 28-41.

Lang, W., Long, Y., & Chen, T.T. (2018). Rediscovering Chinese cities through the lens of land-use patterns. *Land Use Policy*, *79*, 362-374.

Larkin, A., & Hystad, P. (2019). Evaluating street view exposure measures of visible green space for health research. *Journal of Exposure Science & Environmental Epidemiology*, *29*, 447-456.

Lee, A.C.K., & Maheswaran, R. (2011). The health benefits of urban green spaces: a review of the evidence. *Journal of public health*, *33*, 212-222.

Li, D.Q., Lu, D.S., Wu, M., Shao, X.X., & Wei, J.H. (2017a). Examining land cover and greenness dynamics in Hangzhou bay in 1985–2016 using Landsat time-series data. *Remote Sensing*, *10*, 32-54.

Li, D.R., Ma, J., Cheng, T., van Genderen, J.L., & Shao, Z.F. (2018a). Challenges and opportunities for the development of MEGACITIES. *International Journal of Digital Earth*, *12*, 1382-1395.

Li, F., Sun, W.W., Yang, G., & Weng, Q.H. (2019a). Investigating spatiotemporal patterns of surface urban heat islands in the Hangzhou metropolitan area, China, 2000–2015. *Remote Sensing*, *11*, 1553-1572.

Li, J., Benediktsson, J.A., Zhang, B., Yang, T., & Plaza, A. (2017b). Spatial technology and social media in remote sensing: challenges and opportunities. *Proceedings of the IEEE*, *105*, 1583-1585.

Li, J., Ying, L., & Dang, A.R. (2018b). Live-Work-Play Centers of Chinese cities: Identification and temporal evolution with emerging data. *Computers Environment and Urban Systems*, *71*, 58-66.

Li, K.N., & Chen, Y.H. (2018). A genetic algorithm-based urban cluster automatic threshold method by combining VIIRS DNB, NDVI, and NDBI to monitor urbanization. *Remote Sensing*, *10*, 277.

Li, M.M., Stein, A., & de Beurs, K.M. (2020a). A Bayesian characterization of urban land use configurations from VHR remote sensing images. *International Journal of Applied Earth Observation and Geoinformation*, *92*, 102175.

Li, S.N., Dragicevic, S., Castro, F.A., Sester, M., Winter, S., Coltekin, A., Pettit, C., Jiang, B., Haworth, J., Stein, A., & Cheng, T. (2016). Geospatial big data handling theory and methods: A review and research challenges. *ISPRS Journal of Photogrammetry and Remote Sensing*, *115*, 119-133.

Li, W.F., Bai, Y.Q., Zhou, W.Q., Han, C.M., & Han, L.J. (2015). Land use significantly affects the distribution of urban green space: Case study of Shanghai, China. *Journal of Urban Planning and Development*, *141*, A4014001.

Li, W.J., Dong, R.M., Fu, H.H., Wang, J., Yu, L., & Gong, P. (2020b). Integrating Google Earth imagery with Landsat data to improve 30-m resolution land cover mapping. *Remote Sensing of Environment*, *237*, 111563-111579.

Li, W.L. (2021). Mapping urban land use by combining multi-source social sensing data and remote sensing images. *Earth Science Informatics*, *14*, 1537-1545.

Li, X.C., Zhou, Y.Y., Zhu, Z.Y., & Cao, W.T. (2020c). A national dataset of 30 m annual urban extent dynamics (1985-2015) in the conterminous United States. *Earth System Science Data*, *12*, 357-371.

Li, X.T., Hu, T.Y., Gong, P., Du, S.H., Chen, B., Li, X.C., & Dai, Q. (2021). Mapping essential urban land use categories in Beijing with a fast Area of Interest (AOI)-based method. *Remote Sensing*, *13*, 477-503.

Li, Z.L., Huang, Q.Y., & Emrich, C.T. (2019b). Introduction to social sensing and big data computing for disaster management. *International Journal of Digital Earth*, *12*, 1198-1204.

Liang, H.L., Chen, D., & Zhang, Q.P. (2017). Assessing urban green space distribution in a compact megacity by landscape metrics. *Journal of Environmental Engineering and Landscape Management*, *25*, 64-74.

Liebelt, V., Bartke, S., & Schwarz, N. (2018). Hedonic pricing analysis of the influence of urban green spaces onto residential prices: the case of Leipzig, Germany. *European Planning Studies*, 26, 133-157.

Liu, D.D., Chen, N.C., Zhang, X., Wang, C., & Du, W.Y. (2020a). Annual large-scale urban land mapping based on Landsat time series in Google Earth Engine and OpenStreetMap data: A case study in the middle Yangtze River basin. *ISPRS Journal of Photogrammetry and Remote Sensing*, 159, 337-351.

Liu, H.M., Xu, Y.Y., Tang, J.B., Deng, M., Huang, J.C., Yang, W.T., & Wu, F. (2020b). Recognizing urban functional zones by a hierarchical fusion method considering landscape features and human activities. *Transactions in GIS*, 24, 1359-1381.

Liu, H.X., Remme, R.P., Hamel, P., Nong, H.F., & Ren, H. (2020c). Supply and demand assessment of urban recreation service and its implication for greenspace planning-A case study on Guangzhou. *Landscape and Urban Planning*, 203, 103898.

Liu, J.Z., Li, J., Li, W.F., & Wu, J.S. (2016). Rethinking big data: A review on the data quality and usage issues. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115, 134-142.

Liu, S.R., Zhang, X.S., Feng, Y.J., Xie, H., Jiang, L., & Lei, Z.K. (2021). Spatiotemporal dynamics of urban green space influenced by rapid urbanization and land use policies in Shanghai. *Forests*, 12, 476-495.

Liu, X., & Long, Y. (2015). Automated identification and characterization of parcels with OpenStreetMap and points of interest. *Environment and Planning B: Planning and Design*, 43, 498-510.

Liu, X.P., He, J.L., Yao, Y., Zhang, J.B., Liang, H.L., Wang, H., & Hong, Y. (2017). Classifying urban land use by integrating remote sensing and social media data. *International Journal of Geographical Information Science*, 31, 1675-1696.

Liu, X.P., Hu, G.H., Chen, Y.M., Li, X., Xu, X.C., Li, S.Y., Pei, F.S., & Wang, S.J. (2018a). High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform. *Remote Sensing of Environment*, 209, 227-239.

Liu, X.P., Niu, N., Liu, X.J., Jin, H., Ou, J.P., Jiao, L.M., & Liu, Y.L. (2018b). Characterizing mixed-use buildings based on multi-source big data. *International Journal of Geographical Information Science*, 32, 738-756.

Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C.H., Zhi, Y., Chi, G.H., & Shi, L. (2015). Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers*, 105, 512-530.

Liu, Y., Wang, F.H., Xiao, Y.F., & Gao, S. (2012). Urban land uses and traffic ‘source-sink areas’: Evidence from GPS-enabled taxi data in Shanghai. *Landscape and Urban Planning*, 106, 73-87.

Liu, Z. (2021). Identifying urban land use social functional units: a case study using OSM data. *International Journal of Digital Earth*, 14, 1798–1817.

Long, Y., & Liu, L. (2017). How green are the streets? An analysis for central areas of Chinese cities using Tencent Street View. *Plos One*, *12*, 1-18.

Long, Y., Zhai, W.X., Shen, Y., & Ye, X.Y. (2018). Understanding uneven urban expansion with natural cities using open data. *Landscape and Urban Planning*, *177*, 281-293.

Lou, G., Chen, Q.X., He, K., Zhou, Y., & Shi, Z. (2019). Using nighttime light data and POI big data to detect the urban centers of Hangzhou. *Remote Sensing*, *11*, 1821-1841.

Lu, D.S., & Weng, Q.H. (2006). Use of impervious surface in urban land-use classification. *Remote Sensing of Environment*, *102*, 146-160.

Luan, X.L., Yu, Z.W., Zhang, Y.T., Wei, S., Miao, X.Y., Huang, Z.Y.X., Teng, S.Q.N., & Xu, C. (2020). Remote sensing and social sensing data reveal scale-dependent and system-specific strengths of urban heat island determinants. *Remote Sensing*, *12*, 391.

Ludwig, C., Hecht, R., Lautenbach, S., Schorcht, M., & Zipf, A. (2021). Mapping public urban green spaces based on OpenStreetMap and Sentinel-2 imagery using belief functions. *ISPRS International Journal of Geo-Information*, *10*, 251.

M'Ikiugu, M.M., Kinoshita, I., & Tashiro, Y. (2012). Urban green space analysis and identification of its potential expansion areas. *Procedia-Social and Behavioral Sciences*, *35*, 449-458.

Ma, L., Liu, Y., Zhang, X.L., Ye, Y.X., Yin, G.F., & Johnson, B.A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *152*, 166-177.

MacLachlan, A., Roberts, G., Biggs, E., & Boruff, B. (2017). Subpixel land-cover classification for improved urban area estimates using Landsat. *International Journal of Remote Sensing*, *38*, 5763-5792.

Mao, W.L., Lu, D.B., Hou, L., Liu, X., & Yue, W.Z. (2020). Comparison of machine-learning methods for urban land-use mapping in Hangzhou City, China. *Remote Sensing*, *12*, 2817-2834.

Martí, P., Serrano-Estrada, L., & Nolasco-Cirugeda, A. (2019). Social Media data: Challenges, opportunities and limitations in urban studies. *Computers, Environment and Urban Systems*, *74*, 161-174.

Masoudi, M., & Tan, P.Y. (2019). Multi-year comparison of the effects of spatial pattern of urban green spaces on urban land surface temperature. *Landscape and Urban Planning*, *184*, 44-58.

Masoudi, M., Tan, P.Y., & Fadaei, M. (2021). The effects of land use on spatial pattern of urban green spaces and their cooling ability. *Urban Climate*, *35*, 100743-100760.

Matos, P., Vieira, J., Rocha, B., Branquinho, C., & Pinho, P. (2019). Modeling the provision of air-quality regulation ecosystem service provided by urban green spaces using lichens as ecological indicators. *Science of the Total Environment*, *665*, 521-530.

McKenzie, D.P., Mackinnon, A.J., Péladeau, N., Onghena, P., Bruce, P.C., Clarke, D.M., Harrigan, S., & McGorry, P.D. (1996). Comparing correlated kappas by resampling: is one level of agreement significantly different from another? *Journal of psychiatric research*, *30*, 483-492.

Meng, Q.Y., Zhang, L.L., Sun, Z.H., Meng, F., Wang, L., & Sun, Y.X. (2018). Characterizing spatial and temporal trends of surface urban heat island effect in an urban main built-up area: A 12-year case study in Beijing, China. *Remote Sensing of Environment*, *204*, 826-837.

Meng, Y., Hou, D.Y., & Xing, H.F. (2017). Rapid detection of land cover changes using crowdsourced geographic information: A case study of Beijing, China. *Sustainability*, *9*, 1547-1563.

Miller, J.D., Stewart, E., Hess, T., & Brewer, T. (2020). Evaluating landscape metrics for characterising hydrological response to storm events in urbanised catchments. *Urban Water Journal*, *17*, 247-258.

Ministry of housing and urban-rural development (MoHURD) (2012). Code for design of urban road engineering CJJ37. In. Beijing, China: China Architecture & Building Press.

Miralles-Guasch, C., Dopico, J., Delclòs-Alió, X., Knobel, P., Marquet, O., Maneja-Zaragoza, R., Schipperijn, J., & Vich, G. (2019). Natural landscape, infrastructure, and health: The physical activity implications of urban green space composition among the elderly. *International Journal of Environmental Research and Public Health*, *16*, 3986.

Mitchell, L., Frank, M.R., Harris, K.D., Dodds, P.S., & Danforth, C.M. (2013). The geography of happiness: connecting twitter sentiment and expression, demographics, and objective characteristics of place. *Plos One*, 8, 1-15.

Momeni, R., Aplin, P., & Boyd, D.S. (2016). Mapping complex urban land cover from spaceborne imagery: the influence of spatial resolution, spectral band set and classification approach. *Remote Sensing*, 8, 88.

Morell-Monzó, S., Sebastián-Frasquet, M.T., & Estornell, J. (2021). Land use classification of VHR images for mapping small-sized abandoned citrus plots by using spectral and textural information. *Remote Sensing*, 13, 681.

Muller, S.V., Racoviteanu, A.E., & Walker, D.A. (2010). Landsat MSS-derived land-cover map of northern Alaska: Extrapolation methods and a comparison with photo-interpreted and AVHRR-derived maps. *International Journal of Remote Sensing*, 20, 2921-2946.

Munoz, C.A., Brovelli, M.A., Corti, S., & Zamboni, G. (2016). Big geo data management: An exploration with social media and telecommunications open data. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41, 595-601.

Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q.H. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment*, 115, 1145-1161.

Niu, N., Liu, X.P., Jin, H., Ye, X.Y., Liu, Y., Li, X., Chen, Y.M., & Li, S.Y. (2017). Integrating multi-source big data to infer building functions. *International Journal of Geographical Information Science*, 31, 1871-1890.

Okujeni, A., Canters, F., Cooper, S.D., Degerickx, J., Heiden, U., Hostert, P., Priem, F., Roberts, D.A., Somers, B., & van der Linden, S. (2018). Generalizing machine learning regression models using multi-site spectral libraries for mapping vegetation-impervious-soil fractions across multiple cities. *Remote Sensing of Environment*, 216, 482-496.

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., & Wulder, M.A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, 148, 42-57.

Olofsson, P., Foody, G.M., Stehman, S.V., & Woodcock, C.E. (2013). Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment*, 129, 122-131.

Pacifici, F., Chini, M., & Emery, W.J. (2009). A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. *Remote Sensing of Environment*, 113, 1276-1292.

Pan, G., Qi, G.D., Wu, Z.H., Zhang, D.Q., & Li, S.J. (2013). Land-use classification using Taxi GPS traces. *IEEE Transactions on Intelligent Transportation Systems*, 14, 113-123.

Pan, H.Y., Tong, X.H., Xu, X., Luo, X., Jin, Y.M., Xie, H., & Li, B.B. (2020). Updating of land cover maps and change analysis using GlobeLand30 product: A case study in Shanghai metropolitan area, China. *Remote Sensing*, *12*, 3147-3172.

Patino, J.E., & Duque, J.C. (2013). A review of regional science applications of satellite remote sensing in urban settings. *Computers, Environment and Urban Systems*, *37*, 1-17.

Pei, T., Sobolevsky, S., Ratti, C., Shaw, S.L., Li, T., & Zhou, C.H. (2014). A new insight into land use classification based on aggregated mobile phone data. *International Journal of Geographical Information Science*, *28*, 1988-2007.

Pflugmacher, D., Rabe, A., Peters, M., & Hostert, P. (2019). Mapping pan-European land cover using Landsat spectral-temporal metrics and the European LUCAS survey. *Remote Sensing of Environment*, *221*, 583-595.

Pindral, S., Kot, R., Hulisz, P., & Charzyński, P. (2020). Landscape metrics as a tool for analysis of urban pedodiversity. *Land Degradation & Development*, *31*, 2281-2294.

Puissant, A., Hirsch, J., & Weber, C. (2006). The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *International Journal of Remote Sensing*, *26*, 733-745.

Qi, L., Li, J., Wang, Y., & Gao, X.B. (2019). Urban observation: Integration of remote sensing and social media data. In, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (pp. 1-13)

Qian, Y.G., Zhou, W.Q., Li, W.F., & Han, L.J. (2015). Understanding the dynamic of greenspace in the urbanized area of Beijing based on high resolution satellite images. *Urban Forestry & Urban Greening*, *14*, 39-47.

Qian, Y.G., Zhou, W.Q., Pickett, S.T.A., Yu, W.J., Xiong, D.P., Wang, W.M., & Jing, C.B. (2020). Integrating structure and function: mapping the hierarchical spatial heterogeneity of urban landscapes. *Ecological Processes*, *9*, 1-11.

Qin, Y.Q., Chi, M.M., Liu, X., Zhang, Y.F., Zeng, Y.J., & Zhao, Z.M. (2018). Classification of high resolution urban remote sensing images using deep networks by integration of social media photos. In, *IEEE International Geoscience and Remote Sensing Symposium* (pp. 7243-7246)

Ratnasari, N., Candra, E.D., Saputra, D.H., Perdana, A.P., & Iop (2017). Urban spatial pattern and interaction based on analysis of nighttime remote sensing data and geo-social media information. In, *2nd International Conference of Indonesian Society for Remote Sensing* (p. 012038)

Ratti, C., Frenchman, D., Pulselli, R.M., & Williams, S. (2016). Mobile landscapes: Using location data from cell phones for urban analysis. *Environment and Planning B: Planning and Design*, *33*, 727-748.

Reba, M., & Seto, K.C. (2020). A systematic review and assessment of algorithms to detect, characterize, and monitor urban land change. *Remote Sensing of Environment*, *242*, 111739.

Reda, K., & Kedzierski, M. (2020). Detection, classification and boundary regularization of buildings in satellite imagery using faster edge region convolutional neural networks. *Remote Sensing*, *12*, 2240-2273.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*, 195-204.

Reyes-Riveros, R., Altamirano, A., De La Barrera, F., Rozas, D., Vieli, L., & Meli, P. (2021). Linking public urban green spaces and human well-being: A systematic review. *Urban Forestry & Urban Greening*, *61*, 127105.

Ricci, A., Guasco, M., Caboni, F., Orlanno, M., Giachetta, A., & Repetto, M.P. (2022). Impact of surrounding environments and vegetation on wind comfort assessment of a new tower with vertical green park. *Building and Environment*, *207*, 108409.

Rogan, J., & Chen, D.M. (2004). Remote sensing technology for mapping and monitoring land-cover and land-use change. *Progress in planning*, *61*, 301-325.

Rozenfeld, H.D., Rybski, D., Gabaix, X., & Makse, H.A. (2011). The area and population of cities: New insights from a different perspective on cities. *American Economic Review*, *101*, 2205-2225.

Sarmin, N.S., & Ismail, M.H. (2016). A review of potentialities and challenges of integrating remote sensing and GIS with socioeconomic data. *Pertanika Journal of Scholarly Research Reviews*, *2*, 129-141.

Schetke, S., Haase, D., & Breuste, J. (2010). Green space functionality under conditions of uneven urban land use development. *Journal of Land Use Science*, 5, 143-158.

Schipperijn, J., Bentsen, P., Troelsen, J., Toftager, M., & Stigsdotter, U.K. (2013). Associations between physical activity and characteristics of urban green space. *Urban Forestry & Urban Greening*, 12, 109-116.

Schneider, A., Friedl, M.A., & Potere, D. (2009). A new map of global urban extent from MODIS satellite data. *Environmental Research Letters*, 4, 044003.

Schneider, A., Friedl, M.A., & Potere, D. (2010). Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions'. *Remote Sensing of Environment*, 114, 1733-1746.

Schneider, A., McIver, D.K., Friedl, M.A., & Strahler, A.H. (2001). Classification of urban areas at continental scales using remotely sensed data. In, *IEEE 2001 International Geoscience and Remote Sensing Symposium* (pp. 2146-2148)

See, L., Mooney, P., Foody, G.M., Bastin, L., Comber, A., Estima, J., Fritz, S., Kerle, N., Jiang, B., Laakso, M., Liu, H.Y., Milčinski, G., Nikšič, M., Painho, M., Pödör, A., Olteanu-Raimond, A.M., & Rutzinger, M. (2016). Crowdsourcing, citizen science or volunteered geographic information? The current state of crowdsourced geographic information. *ISPRS International Journal of Geo-Information*, 5, 55-78.

Semeraro, T., Scarano, A., Buccolieri, R., Santino, A., & Aarrevaara, E. (2021). Planning of urban green spaces: An ecological perspective on human benefits. *Land, 10*, 105.

Seto, K.C., & Fragkias, M. (2005). Quantifying spatiotemporal patterns of urban land-use change in four cities of China with time series landscape metrics. *Landscape Ecology, 20*, 871-888.

Seto, K.C., Fragkias, M., Gueneralp, B., & Reilly, M.K. (2011). A meta-analysis of global urban land expansion. *Plos One, 6*, e23777.

Seto, K.C., & Kaufmann, R.K. (2003). Modeling the drivers of urban land use change in the Pearl River Delta, China: Integrating remote sensing with socioeconomic data. *Land Economics, 79*, 106-121.

Seto, K.C., & Shepherd, J.M. (2009). Global urban land-use trends and climate impacts. *Current Opinion in Environmental Sustainability, 1*, 89-95.

Shahtahmassebi, A.R., Li, C.L., Fan, Y.F., Wu, Y.N., Gan, M.Y., Wang, K., Malik, A.H., & Blackburn, G.A. (2021). Remote sensing of urban green spaces: A review. *Urban Forestry & Urban Greening, 57*, 126946.

Shi, H.C., Zhao, M.X., Simth, D.A., & Chi, B. (2022). Behind the land use mix: measuring the functional compatibility in urban and sub-urban areas of China. *Land, 11*, 2.

Shi, Y., Qi, Z.X., Liu, X.P., Niu, N., & Zhang, H. (2019). Urban land use and land cover classification using multisource remote sensing images and social media data. *Remote Sensing*, *11*, 2719-2743.

Song, J.C., Lin, T., Li, X.H., & Prishchepov, A.V. (2018a). Mapping urban functional zones by integrating very high spatial resolution remote sensing imagery and points of interest: A case study of Xiamen, China. *Remote Sensing*, *10*, 1737-1752.

Song, J.C., Zhao, C.L., Lin, T., Li, X.H., & Prishchepov, A.V. (2019). Spatio-temporal patterns of traffic-related air pollutant emissions in different urban functional zones estimated by real-time video and deep learning technique. *Journal of Cleaner Production*, *238*, 117881-117891.

Song, X.P., Sexton, J.O., Huang, C.Q., Channan, S., & Townshend, J.R. (2016). Characterizing the magnitude, timing and duration of urban growth from time series of Landsat-based estimates of impervious cover. *Remote Sensing of Environment*, *175*, 1-13.

Song, X.P., Tan, P.Y., Edwards, P., & Richards, D. (2018b). The economic benefits and costs of trees in urban forest stewardship: A systematic review. *Urban Forestry & Urban Greening*, *29*, 162-170.

Song, X.Q., Feng, Q., Xia, F.Z., Li, X.Y., & Scheffran, J. (2021). Impacts of changing urban land-use structure on sustainable city growth in China: A population-density dynamics perspective. *Habitat International*, *107*, 102296-102314.

Sperandelli, D.I., Dupas, F.A., & Dias Pons, N.A. (2013). Dynamics of urban sprawl, vacant land, and green spaces on the metropolitan fringe of São Paulo, Brazil. *Journal of Urban Planning and Development*, 139, 274-279.

Srivastava, S., Vargas Muñoz, J.E., Lobry, S., & Tuia, D. (2018). Fine-grained landuse characterization using ground-based pictures: a deep learning solution based on globally available data. *International Journal of Geographical Information Science*, 34, 1117-1136.

Steffen, W., Grinevald, J., Crutzen, P., & McNeill, J. (2011). The Anthropocene: conceptual and historical perspectives. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369, 842-867.

Stehman, S.V., & Foody, G.M. (2019). Key issues in rigorous accuracy assessment of land cover products. *Remote Sensing of Environment*, 231, 111199-111222.

Stehman, S.V., & Wickham, J.D. (2011). Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy assessment. *Remote Sensing of Environment*, 115, 3044-3055.

Su, M., Guo, R.Z., Chen, B., Hong, W.Y., Wang, J.Q., Feng, Y.M., & Xu, B. (2020). Sampling strategy for detailed urban land use classification: A systematic analysis in Shenzhen. *Remote Sensing*, 12, 1497-1515.

Sun, B., Zhang, Y., Zhou, Q.M., & Zhang, X.C. (2022). Effectiveness of semi-supervised learning and multi-source data in detailed urban landuse mapping with a few labeled samples. *Remote Sensing*, 14, 648.

Sun, J., Wang, H., Song, Z.L., Lu, J.B., Meng, P.Y., & Qin, S.H. (2020). Mapping essential urban land use categories in Nanjing by integrating multi-source big data. *Remote Sensing*, *12*, 2386-2404.

Tang, H.Z., Liu, W.P., & Yun, W.J. (2018). Spatiotemporal dynamics of green spaces in the Beijing–Tianjin–Hebei region in the past 20 years. *Sustainability*, *10*, 2949-2964.

Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., & Dech, S. (2012). Monitoring urbanization in mega cities from space. *Remote Sensing of Environment*, *117*, 162-176.

Taubenböck, H., Wurm, M., Setiadi, N., Gebert, N., Roth, A., Strunz, G., Birkmann, J., & Dech, S. (2009). Integrating remote sensing and social science—the correlation of urban morphology with socioeconomic parameters. In, *IEEE Urban Remote Sensing Joint Event* (pp. 1-7)

Teixeira, C.F.B. (2021). Green space configuration and its impact on human behavior and URBAN environments. *Urban Climate*, *35*, 100746-100757.

Threlfall, C.G., Ossola, A., Hahs, A.K., Williams, N.S.G., Wilson, L., & Livesley, S.J. (2016). Variation in vegetation structure and composition across urban green space types. *Frontiers in Ecology and Evolution*, *4*, 66-78.

Tong, X.Y., Lu, Q.K., Xia, G.S., & Zhang, L.P. (2018). Large-scale land cover classification in Gaofen-2 satellite imagery. In, *IEEE International Geoscience and Remote Sensing Symposium* (pp. 3599-3602): IEEE.

Tu, W., Hu, Z.W., Li, L.F., Cao, J.Z., Jiang, J.C., Li, Q.P., & Li, Q.Q. (2018). Portraying urban functional zones by coupling remote sensing imagery and human sensing data. *Remote Sensing*, *10*, 141-161.

Tu, Y., Chen, B., Lang, W., Chen, T.T., Li, M., Zhang, T., & Xu, B. (2021). Uncovering the nature of urban land use composition using multi-source open big data with ensemble learning. *Remote Sensing*, *13*, 4241.

Uuemaa, E., Antrop, M., Roosaare, J., Marja, R., & Mander, Ü. (2009). Landscape metrics and indices: an overview of their use in landscape research. *Living reviews in landscape research*, *3*, 1-28.

Van Zanten, B.T., Van Berkel, D.B., Meentemeyer, R.K., Smith, J.W., Tieskens, K.F., & Verburg, P.H. (2016). Continental-scale quantification of landscape values using social media data. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 12974-12979.

Vargas Muñoz, J.E., Tuia, D., & Falcão, A.X. (2021). Deploying machine learning to assist digital humanitarians: making image annotation in OpenStreetMap more efficient. *International Journal of Geographical Information Science*, *35*, 1725-1745.

Villeneuve, P.J., Jerrett, M., Su, J.G., Burnett, R.T., Chen, H., Wheeler, A.J., & Goldberg, M.S. (2012). A cohort study relating urban green space with mortality in Ontario, Canada. *Environmental research*, *115*, 51-58.

Vu, T.T., Vu, N.V.A., Phung, H.P., & Nguyen, L.D. (2021). Enhanced urban functional land use map with free and open-source data. *International Journal of Digital Earth*, 14, 1744-1757.

Wan, T.L., Lu, H.Y., Lu, Q.K., & Luo, N.X. (2017). Classification of high-resolution remote-sensing image using OpenStreetMap information. *IEEE Geoscience and Remote Sensing Letters*, 14, 2305-2309.

Wang, J., Zhou, W.Q., Wang, J., & Yu, W.J. (2020). Spatial distribution of urban greenspace in response to urban development from a multi-scale perspective. *Environmental Research Letters*, 15, 064031.

Wang, R.H., Zhao, J.W., Meitner, M.J., Hu, Y., & Xu, X.L. (2019a). Characteristics of urban green spaces in relation to aesthetic preference and stress recovery. *Urban Forestry & Urban Greening*, 41, 6-13.

Wang, S.S., Zhao, Z.L., Xu, Y.Q., & Li, X.L. (2019b). Mathematical analysis of urban land use change in Xi'an city wall area by using parcel-level data. *Science China Technological Sciences*, 62, 687-697.

Wang, X.J. (2009). Analysis of problems in urban green space system planning in China. *Journal of Forestry Research*, 20, 79-82.

Wang, X.J., Zhang, Y., Zhang, J., Fu, C.L., & Zhang, X.L. (2021a). Progress in urban metabolism research and hotspot analysis based on CiteSpace analysis. *Journal of Cleaner Production*, 281, 125224.

Wang, Y.D., Wang, T., Tsou, M.H., Li, H., Jiang, W., & Guo, F.Q. (2016). Mapping dynamic urban land use patterns with crowdsourced geo-tagged social media (Sina-Weibo) and commercial points of interest collections in Beijing, China. *Sustainability*, 8, 1202-1221.

Wang, Y.J., Chen, X.P., Zhang, Z.L., Xue, B., & Lu, C.Y. (2019c). Cross-city convergence in urban green space coverage in China. *Sustainability*, 11, 4707-4718.

Wang, Z.Q., Gang, C.C., Li, X.L., Chen, Y.Z., & Li, J.L. (2015). Application of a normalized difference impervious index (NDII) to extract urban impervious surface features based on Landsat TM images. *International Journal of Remote Sensing*, 36, 1055-1069.

Wang, Z.Y., Ma, D.B., Sun, D.Q., & Zhang, J.X. (2021b). Identification and analysis of urban functional area in Hangzhou based on OSM and POI data. *Plos One*, 16, e0251988.

Weiss, D.J., Nelson, A., Gibson, H.S., Temperley, W., Peedell, S., Lieber, A., Hancher, M., Poyart, E., Belchior, S., Fullman, N., Mappin, B., Dalrymple, U., Rozier, J., Lucas, T.C.D., Howes, R.E., Tusting, L.S., Kang, S.Y., Cameron, E., Bisanzio, D., Battle, K.E., Bhatt, S., & Gething, P.W. (2018). A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature*, 553, 333-336.

Weng, Q.H. (2012). Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sensing of Environment*, 117, 34-49.

Weng, Q.H., Hu, X.F., & Liu, H. (2009). Estimating impervious surfaces using linear spectral mixture analysis with multitemporal ASTER images. *International Journal of Remote Sensing*, 30, 4807-4830.

Wetherley, E.B., Roberts, D.A., & McFadden, J.P. (2017). Mapping spectrally similar urban materials at sub-pixel scales. *Remote Sensing of Environment*, 195, 170-183.

Wolch, J.R., Byrne, J., & Newell, J.P. (2014). Urban green space, public health, and environmental justice: The challenge of making cities ‘just green enough’. *Landscape and Urban Planning*, 125, 234-244.

Woldesemayat, E.M., & Genovese, P.V. (2021). Urban green space composition and configuration in functional land use areas in Addis Ababa, Ethiopia, and their relationship with urban form. *Land*, 10, 85-106.

Wu, C.S., & Murray, A.T. (2003). Estimating impervious surface distribution by spectral mixture analysis. *Remote Sensing of Environment*, 84, 493-505.

Wu, H., Lin, A.Q., Xing, X.D., Song, D.X., & Li, Y. (2021a). Identifying core driving factors of urban land use change from global land cover products and POI data using the random forest method. *International Journal of Applied Earth Observation and Geoinformation*, 103, 102475.

Wu, H.Y., Zhang, T., & Gong, J.Y. (2015). GeoComputation for geospatial big data. *Transactions in GIS*, 18, 1-2.

Wu, L., Cheng, X.M., Kang, C.G., Zhu, D., Huang, Z., & Liu, Y. (2020). A framework for mixed-use decomposition based on temporal activity signatures extracted from big geo-data. *International Journal of Digital Earth*, 13, 708-726.

Wu, R., Wang, J.Y., Zhang, D.C., & Wang, S.J. (2021b). Identifying different types of urban land use dynamics using Point-of-interest (POI) and Random Forest algorithm: The case of Huizhou, China. *Cities*, 114, 103202.

Wu, Y.Z., Fan, P.L., Li, B., Ouyang, Z.T., Liu, Y., & You, H.Y. (2017). The effectiveness of planning control on urban growth: Evidence from Hangzhou, China. *Sustainability*, 9, 855.

Wu, Z.F., Chen, R.S., Meadows, M.E., Sengupta, D., & Xu, D. (2019). Changing urban green spaces in Shanghai: trends, drivers and policy implications. *Land Use Policy*, 87, 104080.

Wüstemann, H., Kalisch, D., & Kolbe, J. (2017). Access to urban green space and environmental inequalities in Germany. *Landscape and Urban Planning*, 164, 124-131.

Xiao, X.D., Dong, L., Yan, H.N., Yang, N., & Xiong, Y.M. (2018). The influence of the spatial characteristics of urban green space on the urban heat island effect in Suzhou Industrial Park. *Sustainable Cities and Society*, 40, 428-439.

Xing, H.F., & Meng, Y. (2018). Integrating landscape metrics and socioeconomic features for urban functional region classification. *Computers, Environment and Urban Systems*, 72, 134-145.

Xing, H.F., Meng, Y., Hou, D.Y., Song, J., & Xu, H.B. (2017). Employing crowdsourced geographic information to classify land cover with spatial clustering and topic model. *Remote Sensing*, *9*, 602-622.

Xing, H.F., Meng, Y., & Shi, Y. (2018). A dynamic human activity-driven model for mixed land use evaluation using social media data. *Transactions in GIS*, *22*, 1130-1151.

Xiong, G.P., Cao, X., Hamm, N.A.S., Lin, T., Zhang, G.Q., & Chen, B.H. (2021). Unbalanced development characteristics and driving mechanisms of regional urban spatial form: A case study of Jiangsu Province, China. *Sustainability*, *13*, 3121.

Xu, C., Haase, D., Pribadi, D.O., & Pauleit, S. (2018). Spatial variation of green space equity and its relation with urban dynamics: A case study in the region of Munich. *Ecological Indicators*, *93*, 512-523.

Xu, G., Zhu, X.X., Fu, D.J., Dong, J.W., & Xiao, X.M. (2017). Automatic land cover classification of geo-tagged field photos by deep learning. *Environmental Modelling & Software*, *91*, 127-134.

Xu, H.Q. (2010). Analysis of impervious surface and its impact on urban heat environment using the normalized difference impervious surface index (NDISI). *Photogrammetric Engineering and Remote Sensing*, *76*, 557-565.

Xu, N., Luo, J.C., Wu, T.J., Dong, W., Liu, W., & Zhou, N. (2021a). Identification and portrait of urban functional zones based on multisource heterogeneous data and ensemble learning. *Remote Sensing*, *13*, 373.

Xu, S.Y., Qing, L.B., Han, L.M., Liu, M., Peng, Y.H., & Shen, L.F. (2020). A new remote sensing images and point-of-interest fused (RPF) model for sensing urban functional regions. *Remote Sensing*, *12*, 1032-1057.

Xu, Z.Y., Su, C., & Zhang, X.C. (2021b). A semantic segmentation method with category boundary for Land Use and Land Cover (LULC) mapping of Very-High Resolution (VHR) remote sensing image. *International Journal of Remote Sensing*, *42*, 3146-3165.

Yammine, S.Z., Jarreau, P.B., Liu, C., & Coe, I.R. (2018). Social media for social change in science. *Science*, *360*, 162-163.

Yan, W.Y., Shaker, A., & El-Ashmawy, N. (2015). Urban land cover classification using airborne LiDAR data: A review. *Remote Sensing of Environment*, *158*, 295-310.

Yang, B., Lee, D.K., Heo, H.K., & Biging, G. (2019). The effects of tree characteristics on rainfall interception in urban areas. *Landscape and Ecological Engineering*, *15*, 289-296.

Yang, J., Guan, Y.Y., Xia, J.C., Jin, C., & Li, X.M. (2018). Spatiotemporal variation characteristics of green space ecosystem service value at urban fringes: A case study on Ganjingzi District in Dalian, China. *Science of the Total Environment*, *639*, 1453-1461.

Yang, L.M., Huang, C.Q., Homer, C.G., Wylie, B.K., & Coan, M.J. (2003a). An approach for mapping large-area impervious surfaces: synergistic use of Landsat-7

ETM+ and high spatial resolution imagery. *Canadian Journal of Remote Sensing*, 29, 230-240.

Yang, L.M., Xian, G., Klaver, J.M., & Deal, B. (2003b). Urban land-cover change detection through sub-pixel imperviousness mapping using remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 69, 1003-1010.

Yang, W., & Mu, L. (2015). GIS analysis of depression among Twitter users. *Applied Geography*, 60, 217-223.

Yang, X., & Lo, C.P. (2010). Using a time series of satellite imagery to detect land use and land cover changes in the Atlanta, Georgia metropolitan area. *International Journal of Remote Sensing*, 23, 1775-1798.

Yao, Y., Yan, X.Q., Luo, P., Liang, Y.Y., Ren, S.L., Hu, Y., Han, J., & Guan, Q.F. (2022). Classifying land-use patterns by integrating time-series electricity data and high-spatial resolution remote sensing imagery. *International Journal of Applied Earth Observation and Geoinformation*, 106, 102664.

Yao, Y., Zhang, J.B., Hong, Y., Liang, H.L., & He, J.L. (2018). Mapping fine-scale urban housing prices by fusing remotely sensed imagery and social media data. *Transactions in GIS*, 22, 561-581.

Ye, X.Y., Huang, Q.Y., & Li, W.W. (2016). Integrating big social data, computing and modeling for spatial social science. *Cartography and Geographic Information Science*, 43, 377-378.

Ye, Y.Q., An, Y., Chen, B., Wang, J.J., & Zhong, Y.Q. (2020). Land use classification from social media data and satellite imagery. *Journal of Supercomputing*, 76, 777-792.

Yin, J.D., Dong, J.W., Hamm, N.A.S., Li, Z.C., Wang, J.H., Xing, H.F., & Fu, P. (2021a). Integrating remote sensing and geospatial big data for urban land use mapping: A review. *International Journal of Applied Earth Observation and Geoinformation*, 103, 102514-102525.

Yin, J.D., Fu, P., Hamm, N.A.S., Li, Z.C., You, N.S., He, Y.L., Cheshmehzangi, A., & Dong, J.W. (2021b). Decision-level and feature-level integration of remote sensing and geospatial big data for urban land use mapping. *Remote Sensing*, 13, 1579-1592.

Yu, B.L., Lian, T., Huang, Y.X., Yao, S.J., Ye, X.Y., Chen, Z.Q., Yang, C.S., & Wu, J.P. (2019). Integration of nighttime light remote sensing images and taxi GPS tracking data for population surface enhancement. *International Journal of Geographical Information Science*, 33, 687-706.

Yu, Y., Li, J., Zhu, C.Y., & Plaza, A. (2018). Urban impervious surface estimation from remote sensing and social data. *Photogrammetric Engineering and Remote Sensing*, 84, 771-780.

Yu, Z., Wang, Y., Deng, J., Shen, Z., Wang, K., Zhu, J., & Gan, M. (2017). Dynamics of hierarchical urban green space patches and implications for management policy. *Sensors*, 17, 1304-1319.

Zanella, A., Bui, N., Castellani, A., Vangelista, L., & Zorzi, M. (2014). Internet of things for smart cities. *Ieee Internet of Things Journal*, *1*, 22-32.

Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, *24*, 583-594.

Zhan, X.Y., Ukkusuri, S.V., & Zhu, F. (2014). Inferring urban land use using large-scale social media check-in data. *Networks and Spatial Economics*, *14*, 647-667.

Zhang, C., Sargent, I., Pan, X., Li, H.P., Gardiner, A., Hare, J., & Atkinson, P.M. (2018). An object-based convolutional neural network (OCNN) for urban land use classification. *Remote Sensing of Environment*, *216*, 57-70.

Zhang, H.S., & Xu, R. (2018). Exploring the optimal integration levels between SAR and optical data for better urban land cover mapping in the Pearl River Delta. *International Journal of Applied Earth Observation and Geoinformation*, *64*, 87-95.

Zhang, J., Yu, L., Li, X.C., Zhang, C.C., Shi, T.Z., Wu, X.Y., Yang, C., Gao, W.X., Li, Q.Q., & Wu, G.F. (2020a). Exploring annual urban expansions in the Guangdong-Hong Kong-Macau Greater Bay Area: Spatiotemporal features and driving factors in 1986–2017. *Remote Sensing*, *12*, 2615-2631.

Zhang, J.B., Li, X., Yao, Y., Hong, Y., He, J.Y., Jiang, Z.W., & Sun, J.C. (2021). The Traj2Vec model to quantify residents' spatial trajectories and estimate the proportions of urban land-use types. *International Journal of Geographical Information Science*, *35*, 193-211.

Zhang, Q., Chen, C.L., Wang, J.Z., Yang, D.Y., Zhang, Y., Wang, Z.F., & Gao, M. (2020b). The spatial granularity effect, changing landscape patterns, and suitable landscape metrics in the Three Gorges Reservoir Area, 1995–2015. *Ecological Indicators*, *114*, 106259.

Zhang, Q.L., & Seto, K.C. (2011). Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sensing of Environment*, *115*, 2320-2329.

Zhang, X.Y., Du, S.H., & Wang, Q. (2017a). Hierarchical semantic cognition for urban functional zones with VHR satellite images and POI data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *132*, 170-184.

Zhang, X.Y., Du, S.H., & Zheng, Z.J. (2020c). Heuristic sample learning for complex urban scenes: Application to urban functional-zone mapping with VHR images and POI data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *161*, 1-12.

Zhang, Y., Li, Q.Z., Huang, H.P., Wu, W., Du, X., & Wang, H.Y. (2017b). The combined use of remote sensing and social sensing data in fine-grained urban land use mapping: A case study in Beijing, China. *Remote Sensing*, *9*, 865-888.

Zhang, Y.T., Li, Q.Q., Tu, W., Mai, K., Yao, Y., & Chen, Y.Y. (2019). Functional urban land use recognition integrating multi-source geospatial data and cross-correlations. *Computers, Environment and Urban Systems*, *78*, 101374-101400.

Zhao, J., Chen, S., Jiang, B., Ren, Y., Wang, H., Vause, J., & Yu, H. (2013). Temporal trend of green space coverage in China and its relationship with urbanization over the last two decades. *Science of the Total Environment*, 442, 455-465.

Zhao, W.F., Li, Q.Q., & Li, B.J. (2011). Extracting hierarchical landmarks from urban POI data. *Journal of Remote Sensing*, 15, 973-988.

Zhao, W.Z., Bo, Y.C., Chen, J.G., Tiede, D., Blaschke, T., & Emery, W.J. (2019). Exploring semantic elements for urban scene recognition: Deep integration of high-resolution imagery and OpenStreetMap (OSM). *ISPRS Journal of Photogrammetry and Remote Sensing*, 151, 237-250.

Zheng, Q.M., Weng, Q.H., Huang, L.Y., Wang, K., Deng, J.S., Jiang, R.W., Ye, Z.R., & Gan, M.Y. (2018). A new source of multi-spectral high spatial resolution night-time light imagery—JL1-3B. *Remote Sensing of Environment*, 215, 300-312.

Zheng, S., Wang, J., Sun, C., Zhang, X., & Kahn, M.E. (2019). Air pollution lowers Chinese urbanites' expressed happiness on social media. *Nature Human Behaviour*, 3, 237-243.

Zhong, C., Huang, X.F., Müller Arisona, S., Schmitt, G., & Batty, M. (2014). Inferring building functions from a probabilistic model using public transportation data. *Computers, Environment and Urban Systems*, 48, 124-137.

Zhong, Y.F., Su, Y., Wu, S.Q., Zheng, Z.D., Zhao, J., Ma, A.L., Zhu, Q.Q., Ye, R.C., Li, X.M., Pellikka, P., & Zhang, L.P. (2020). Open-source data-driven urban land-use

mapping integrating point-line-polygon semantic objects: A case study of Chinese cities. *Remote Sensing of Environment*, 247, 111838-111858.

Zhou, W., Wang, J., Qian, Y., Pickett, S.T.A., Li, W., & Han, L. (2018a). The rapid but "invisible" changes in urban greenspace: A comparative study of nine Chinese cities. *Science of the Total Environment*, 627, 1572-1584.

Zhou, X.L., & Wang, Y.C. (2011). Spatial-temporal dynamics of urban green space in response to rapid urbanization and greening policies. *Landscape and Urban Planning*, 100, 268-277.

Zhou, X.L., & Zhang, L. (2016). Crowdsourcing functions of the living city from Twitter and Foursquare data. *Cartography and Geographic Information Science*, 43, 393-404.

Zhou, Y.Y., Li, X.C., Asrar, G.R., Smith, S.J., & Imhoff, M. (2018b). A global record of annual urban dynamics (1992–2013) from nighttime lights. *Remote Sensing of Environment*, 219, 206-220.

Zhu, X.X., Tuia, D., Mou, L.C., Xia, G.S., Zhang, L.P., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, 5, 8-36.

Zhu, Z., Wang, S.X., & Woodcock, C.E. (2015). Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, 159, 269-277.

Zhu, Z., Woodcock, C.E., Rogan, J., & Kellndorfer, J. (2012). Assessment of spectral, polarimetric, temporal, and spatial dimensions for urban and peri-urban land cover classification using Landsat and SAR data. *Remote Sensing of Environment*, *117*, 72-82.

Zhu, Z., Zhou, Y.Y., Seto, K.C., Stokes, E.C., Deng, C.B., Pickett, S.T.A., & Taubenböck, H. (2019). Understanding an urbanizing planet: Strategic directions for remote sensing. *Remote Sensing of Environment*, *228*, 164-182.

Zhuo, Y.F., Zheng, H.Y., Wu, C.F., Xu, Z.G., Li, G., & Yu, Z.N. (2019). Compatibility mix degree index: A novel measure to characterize urban land use mix pattern. *Computers, Environment and Urban Systems*, *75*, 49-60.

Zong, L.L., He, S.J., Lian, J.T., Bie, Q., Wang, X.Y., Dong, J.R., & Xie, Y.W. (2020). Detailed mapping of urban land use based on multi-source data: A case study of Lanzhou. *Remote Sensing*, *12*, 1987-2006.

**Appendix 1 : Integrating remote sensing and
geospatial big data for urban land use mapping:
A review**



Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

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

Integrating remote sensing and geospatial big data for urban land use mapping: A review

Jiadi Yin^{a,b}, Jinwei Dong^{a,*}, Nicholas A.S. Hamm^b, Zhichao Li^a, Jianghao Wang^c, Hanfa Xing^{d,e}, Ping Fu^b

^a Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

^b School of Geographical Sciences and Geospatial Research Group, Faculty of Science and Engineering, University of Nottingham, Ningbo 315100, China

^c State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

^d College of Geography and Environment, Shandong Normal University, Jinan 250300, China

^e School of Geography, South China Normal University, Guangzhou 510631, China

ARTICLE INFO

Keywords:

Integration methods
Urban functional zone classification
Urban management
Land use

ABSTRACT

Remote Sensing (RS) has been used in urban mapping for a long time; however, the complexity and diversity of urban functional patterns are difficult to be captured by RS only. Emerging Geospatial Big Data (GBD) are considered as the supplement to RS data, and help to contribute to our understanding of urban lands from physical aspects (i.e., urban land cover) to socioeconomic aspects (i.e., urban land use). Integrating RS and GBD could be an effective way to combine physical and socioeconomic aspects with great potential for high-quality urban land use classification. In this study, we reviewed the existing literature and focused on the state-of-the-art and perspective of the urban land use categorization by integrating RS and GBD. Specifically, the commonly used RS features (e.g., spectral, textural, temporal, and spatial features) and GBD features (e.g., spatial, temporal, semantic, and sequence features) were identified and analyzed in urban land use classification. The integration strategies for RS and GBD features were categorized into feature-level integration (FI) and decision-level integration (DI). To be more specific, the FI method integrates the RS and GBD features and classifies urban land use types using the integrated feature sets; the DI method processes RS and GBD independently and then merges the classification results based on decision rules. We also discussed other critical issues, including analysis unit setting, parcel segmentation, parcel labeling of land use types, and data integration. Our findings provide a retrospect of different features from RS and GBD, strategies of RS and GBD integration, and their pros and cons, which could help to define the framework for future urban land use mapping and better support urban planning, urban environment assessment, urban disaster monitoring and urban traffic analysis.

1. Introduction

With the advent of the Anthropocene (Ellis and Ramankutty, 2008; Steffen et al., 2011), urbanization is accelerating and the urban population is predicted to grow from 4.2 billion (57.5% of the world population) in 2018 to about 6.7 billion (69.1% in 2050 (Seto et al., 2011)). Such increasing human-induced influences are changing urban land in different dimensions from physical aspects (urban land cover) to socioeconomic aspects (urban land use) (Elmqvist et al., 2019; Hersperger et al., 2018). A large number of high-accuracy urban land cover products (mainly physical characteristics) at the annual level with relatively high

spatial resolution have been developed worldwide (Li et al., 2020a; Liu et al., 2018; Zhou et al., 2018). However, urban land governance and planning need more information on urban land use, which is particularly complex and includes both physical aspects and socioeconomic aspects. Unfortunately, high-quality urban land use products with timely and accurate information related to human activities are still limited (Gong et al., 2020). Understanding the state-of-the-art of existing urban land use mapping efforts, considering both physical and socioeconomic functions, would enable better urban land management and monitoring (Martí et al., 2019; Yammine et al., 2018).

A wide range of satellite remote sensing (RS) data (e.g., Moderate-

* Corresponding author.
E-mail address: dongjw@igsrr.ac.cn (J. Dong).

<https://doi.org/10.1016/j.jag.2021.102514>

Received 2 June 2021; Received in revised form 8 August 2021; Accepted 20 August 2021

0303-2434/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

resolution Imaging Spectroradiometer (MODIS), Landsat TM/ETM/OLS, Defense Meteorological Satellite Program Operation Linescan System (DMSP-OLS)) have been used to study the structures, boundaries, and areas of cities (Huang et al., 2021; Gong et al., 2019a; Schneider et al., 2010). Nevertheless, the complexity and diversity of functional patterns in urban areas cannot be captured well by using RS only due to limited information (e.g. spectral, textural, and temporal information) from RS techniques (Cao et al., 2020). Advances in information and communication technologies make it possible to get access to geospatial big data (GBD) (Li et al., 2016; Li et al., 2021; Yin et al., 2021). Fixed and mobile sensors such as environmental sensors, cameras, webcams, social media, or even urban residents through their regular activities (Wu et al., 2015) create tremendous GBD every day. These data such as mobile phone data (Gong et al., 2020), traffic trajectories (Yu et al., 2019), geo-tagged photos (Cadavid Restrepo et al., 2017; Krylov et al., 2018), Points of interest (POIs) (Yin et al., 2021) and social media data (Huang et al., 2018) provide an alternative approach to uncover how cities function (Ye et al., 2016). It is possible for examining the physical and socio-economic characteristics of the urban land system by taking both the advantages of RS and GBD (Qi et al., 2019; Song et al., 2021; Zhang et al., 2021a; Xiong et al., 2021; Zhang et al., 2021b).

Despite the great potential of integrating RS and GBD for providing better insights into urban land use, it is challenging to combine them due to the differences in the spatial data quality (e.g., semantic, timestamp, and scale), technical format, and data structure (Liu et al., 2015). Summarizing the features of RS and GBD and integration strategies in the literature are needed for guiding future studies and help to understand more detailed urban functional patterns.

In this context, this study examined the literature on the nature of RS and GBD, as well as their integration strategies in urban land use classification, and identified the opportunities and challenges for synthesizing RS and GBD (Table S1). The primary objective of this paper is to review the state-of-the-art in this field by considering (1) the key characteristics of RS and GBD and (2) the methods for integrating RS and GBD. We consider only satellite-based RS and do not consider RS data obtained from airborne platforms. This review is organized into six sections. Section 2 summarized the transformation from urban land cover to urban land use. In section 3, we summarized the commonly used RS and GBD features for urban land use categorization. In Section 4, the integration strategies were analyzed systematically. In Section 5, we discussed the challenges and potential applications of the integration of RS and GBD on urban land use maps. Section 6 concluded the main

findings and implications.

2. Evolution from urban land cover to urban land use

Using satellite data to map urban land cover has a long history (Howarth and Boasson, 1983; Patino and Duque, 2013; Reba and Seto, 2020). Examples of existing efforts for global urban land cover products that have been derived from RS are shown in Table 1. The data source, the nomenclature of urban land, spatial resolution, time period, and reference for each global urban land cover map were summarized. Most urban land cover maps were obtained from coarse spatial resolution images (100 m-10 km), such as MODIS and Advanced Very High-Resolution Radiometer (AVHRR) (Friedl et al., 2002; Schneider et al., 2010). With the substantial progress of RS techniques, recent maps were derived from moderate spatial resolution images (10-100 m), including Satellite Pour l' Observation de la Terre (SPOT) and Landsat images (Deng et al., 2019; Li et al., 2020a). The time period of these global urban land cover products has transformed from a single period to repeated observations, which could provide better quality and time series urban land cover information (Gong et al., 2019a; Li et al., 2020b). Overall, these global urban RS studies focused on the identification of physical urban attributes (e.g., impervious surface, built-up areas, artificial surfaces, and urban extent) that have provided opportunities for a better understanding of global urbanization's effects on human civilization and the environment (Zhu et al., 2019). Despite the aforementioned extensive applications of RS data for mapping urban land cover, more specific information of inner-urban functions cannot be retrieved by using RS only (Li et al., 2017a; Liu et al., 2015).

The demands for urban land products have changed gradually, with increasing information needs on socioeconomic properties, emphasizing a transformation from urban land cover to urban land use (Fig. 1). The multi-sourced GBD can contribute to the understanding of socioeconomic characteristics of urban land use, and identify how people use lands (Srivastava et al., 2018; Zhang et al., 2019a; Zhao et al., 2018). Recently, GBD has been used in conjunction with RS data to extract urban land use information (Liu et al., 2020a; Shi et al., 2019). Therefore, understanding characteristics derived from RS and GBD and their integration methods are necessary for urban land use mapping (Li et al., 2017a; Qi et al., 2019).

Table 1
Comparison of existing global urban land cover products.

Products	Data	Nomenclature of urban land	Spatial resolution	Time period	Reference
GLCC	AVHRR	Built-up areas	1 km	1992,1993	Loveland et al., 2000
UMD1km	AVHRR	Urban and built	1 km	1992,1993	Hansen et al., 2000
GRUMP	VMAP, Census data, DMSP-OLS, Maps	Urban extent	1 km	1995	CIESIN et al., 2011
GLC2000	SPOT-Vegetation, DMSP-OLS	Artificial surfaces and associated areas	1 km	2000	Bartholomé and Belward, 2005
IMPSA	DMSP-OLS	Impervious surface	1 km	2000	Elvidge et al., 2007
NTL-Urban	DMSP-OLS	Urban extent	1 km	1992-2013	Zhou et al., 2018
MOD500	MODIS	Non vegetated, human-constructed elements	500 m	2001-2017	Friedl et al., 2002
GHSL	Fine-scale satellite imagery, census data, and OSM	Built-up areas	500 m	1975, 1990, 2000, 2015	Pesaresi et al., 2013
GlobCover	SPOT-Vegetation	Artificial surfaces and associated areas	300 m	2005, 2009	Ariño et al., 2007
CCI-LC	SPOT-Vegetation	Urban extent	300 m	1992-2015	Defourny et al., 2018
GlobalLand30	Landsat TM/ETM+	Artificial surfaces	30 m	2000, 2010	Chen et al., 2015
HBASE	Landsat TM/ETM+	Built-up and settlement extent	30 m	2010	Wang et al., 2017
GMSI	Landsat TM/ETM+	Impervious surface	30 m	2010	Colston et al., 2017
FROM-GLC	Landsat TM/ETM+/OLI	Impervious surface	30 m	2010, 2015, 2017	Gong et al., 2013
Global Urban Land	Landsat TM/ETM+	Impervious surface	30 m	1990, 1995, 2000, 2005, 2010	Liu et al., 2018
GUF	TerraSAR-X, TanDEM-X	Built-up areas	12 m	2011	Esch et al., 2012
FROM-GLC10	Sentinel	Impervious surface	10 m	2017	Gong et al., 2019b

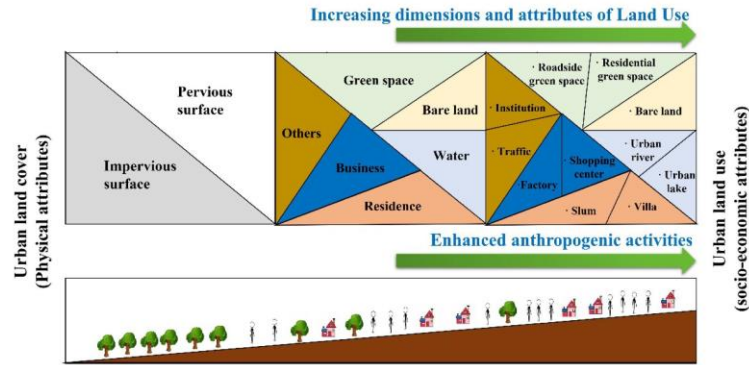


Fig. 1. Transformation of the need for urban land products from physical attributes to socioeconomic attributes due to enhanced anthropogenic activities. The three boxes above represent the transformation from urban land cover (e.g., impervious surface and pervious surface) to urban land use (e.g., traffic, institution, urban lake, residential green space, and others). The box below refers to the enhanced human impacts on urban areas. It should be noted that the change of area proportions of different urban land use types is not reflected in Fig. 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. RS and GBD features for urban land use mapping

The extraction of RS and GBD features is the most essential

procedure for urban land use recognition because the performance of urban land use maps relies heavily on these features. The purpose of this section is thus to introduce several commonly used RS and GBD features

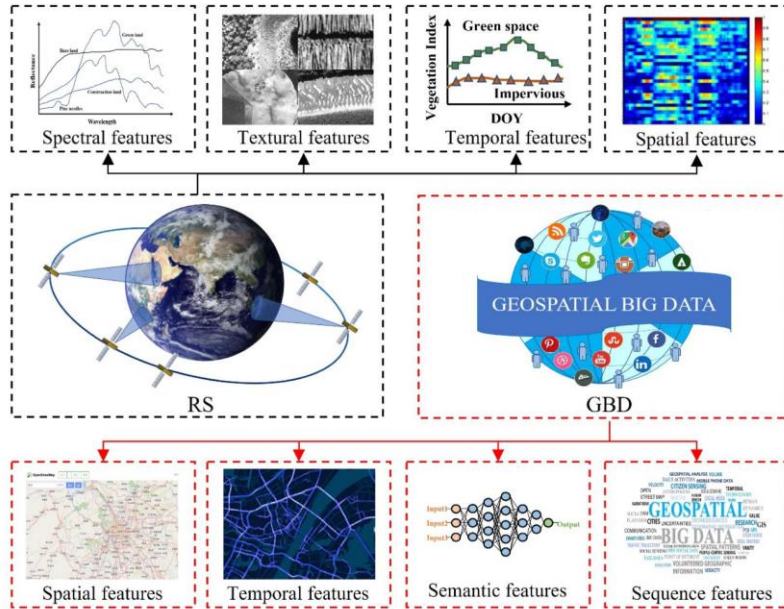


Fig. 2. Summary of the features of RS and GBD. The dotted black box and red box in the middle show the commonly used RS data (e.g., MODIS, Landsat, Sentinel) and GBD (e.g., traffic, social media data, geo-tagged photos), respectively. The upper dotted black boxes represent the features (spectral, textural, temporal, and spatial) extracted from RS, while the dotted red boxes below represent the features (spatial, temporal, semantic, and sequence) extracted from GBD. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

for categorizing urban land use types (Fig. 2).

3.1. RS-based features

The features derived from RS images used in urban land use classification could be categorized into spectral, textural, temporal, and spatial features (Gong and Howarth, 1989; Zhu et al., 2017). Among them, spectral and textural features are common characteristics of RS data to extract urban land use information because different textures and spectra could reflect different urban land use types (Zhu et al., 2019). The temporal features have proven beneficial for improving urban land use maps by providing valuable information (e.g., time series information) on urban land use types (Zhu et al., 2012). Recently, deep learning techniques provide new possibilities for extracting spatial features automatically from the very-high-resolution (VHR) satellite imageries such as WorldView-3, Gaofen-2, and SPOT-5 (Zhang et al., 2020). Spatial features could help classify urban land use at a very detailed level (Zhao et al., 2019). The details of the RS features are specified as follows.

- (1) Spectral features: Generally, the spectral features of urban land show lower reflectance in the near-infrared region (NIR), comparing to vegetation, which has higher reflectance in NIR (Herold et al., 2003). Additionally, the spectra for the visible, short-wave infrared region (SWIR) and microwave regions were also found to be suitable for characterizing urban objects (Heiden et al., 2007). Recently, spectral features from the increased number of bands (e.g., from Landsat to Sentinel-2) provide an opportunity for the acquisition of detailed information on the physical attributes of urban land use, but it also leads to data redundancy due to the high correlation between adjacent bands (Okujeni et al., 2018). Liu et al. (2020b) and Zhang et al. (2017b) extracted spectral features from RS imagery for classifying urban land use. Both the methods calculated the mean and standard deviation of each band by using a certain window.
- (2) Textural features: Textural features contain rich information of the spatial distribution of tonal changes, as well as the structural arrangement of surfaces and their relationships to the surrounding environment (Gong and Howarth, 1989; Haralick et al., 1973). Different textures (e.g., coarse, smooth, rippled, irregular, and lineated) show different image characteristics, such as homogeneity, linear structure, and contrast (Kuffer et al., 2016; Wurm et al., 2017). Therefore, textural features could help to increase the accuracy of land use categorization in heterogeneous landscapes (Jin et al., 2014; Pacifici et al., 2009), where ground objects with different sizes, patterns, structures, and shapes co-exist (Lu and Weng, 2006).
- (3) Temporal features: Temporal features refer to the differences caused by the changes in the spectral and textural features of urban surfaces over time. Due to the seasonality of vegetation growth, it has proven to be effective in improving vegetation and other land cover mapping accuracy (Dong and Xiao, 2016; Zurita-Milla et al., 2013). The extraction of urban land usually tends to be less accurate in the autumn and winter due to more bare land (Weng et al., 2009). However, it is still a challenge to distinguish the variety of processes that generate different time series, for example, due to climate, topography, and terrestrial vegetation (Pflugmacher et al., 2019).
- (4) Spatial features: Along with spectral, textural, and temporal features, the most commonly used feature extracted from RS data is the spatial feature. Recent studies used deep learning techniques such as the supervised convolutional neural network (CNN) models and unsupervised autoencoders (AE) models to extract spatial information automatically from RS images (Reichstein et al., 2019). Deep learning algorithms, which extract high-level spatial information provided by hierarchical

structures, demonstrate remarkable capacity in image representation and understanding in these studies. Traditional approaches such as Random forest models (RF), Support vector machine (SVM), and Decision tree (DT) can only process basic features (e.g., spectral, textural, and temporal features) from RS images. Due to the fine structural information (i.e., spatial details) of urban land use in VHR RS images, VHR RS images were found to be commonly used by deep learning techniques for obtaining urban land use information (Ma et al., 2019).

3.2. GBD-based features

The development of mobile positioning, wireless communication, and the Internet of Things (IoT) provides opportunities for the rapid growth of big data (Kitchin, 2013). According to Kitchin and McArdle (2016), big data is defined in part by its large size and in part by its characteristics, such as volume, variety, and velocity. Liu et al. (2016) further defined the characteristics of big data as exhaustivity, relation-ality, veracity, value, and variability. In the IBM Annual Report, 2.5 terabytes of data are generated every day, with 80% of these data (pictures, texts, and videos) being geo-referenced or capable of being geo-referenced. Therefore, a large proportion of big data is likely to be the GBD (i.e., big data with geographical reference). GBD is generated every day mostly by fixed and mobile sensors such as environmental sensors, cameras, webcams, social media, or even residents' daily activities (Brovelli et al., 2015).

The most commonly used GBD for land use mapping are social sensing (SS) (Liu et al., 2015), citizen sensing (CS) (Jiang et al., 2016), social media data (SMD) (Ilieva and McPhearson, 2018), and volunteered geographic information (VGI) (Goodchild and Glennon, 2010). The descriptions of SS, CS, SMD, and VGI are provided in Table 2. The contents and concepts of SS and CS are much broader than GBD. SS data is used in a variety of applications besides urban land use mapping. CS does not contain the data produced by companies and institutions and VGI focuses on user-generated data. However, the term GBD encompasses all of these above-mentioned geospatial data.

Each record of GBD contains spatial, temporal, semantic, and/or sequence information associated with individuals reflecting human behavior, although the quality of this information may vary in space and time (Fig. 2 and Table 3). Due to the high correlation between human spatiotemporal activities and urban dynamic socioeconomic attributes, these emerging GBD can help to capture the growing complexity of urban functional patterns (Wu et al., 2018). Therefore, in order to better classify and understand urban land use, we add GBD to the identification of socioeconomic and human activities.

- (1) Spatial features: Almost all GBD can provide spatial information (Table 3). For example, the OSM has proven to be a useful spatial

Table 2
Comparison of concepts related to GBD in previous studies.

Concepts	Refs.	Description
Social Sensing (SS)	Liu et al., 2015	A series of data sources with spatiotemporal information which record human activities, as well as the methods and applications based on such data source.
Citizen Sensing (CS)	Jiang et al., 2016	Datasets contributed by citizens provide benefits for themselves and policymakers.
Social Media Data (SMD)	Ilieva and McPhearson, 2018	Information of 'big data' from social media such as Facebook, Flickr, etc.
Volunteered Geographic Information (VGI)	Goodchild and Glennon, 2010	Geographic data provided voluntarily by people use technologies to generate, assemble, and disseminate information.

Table 3
Summary of the GBD and relevant features used in urban land use mapping.

Data sources	Spatial features	Temporal features	Semantic features	Sequence features
Mobile phone data	✓	✓		
Traffic data	✓	✓		
Social media data	✓	✓	✓	✓
Geo-tagged photos	✓		✓	
Maps	✓			✓
Search engine data	✓		✓	✓
Smart card data	✓	✓		

data source including land use information like buildings, roads, and parks (Helbich et al., 2012), which could be used to extract training pixels from RS images for classifying urban land use types (Johnson and Iizuka, 2016; Wan et al., 2017). Google maps, Gaode maps, and other maps have also been successfully applied to urban land use mapping (Xu et al., 2020; Zhang et al., 2020). Furthermore, social media data can provide indirect spatial location data (Long et al., 2018; Yao et al., 2018), for example, Zhan et al. (2014) inferred the urban land use types in New York city by using check-in social media data on Twitter.

- (2) Temporal features: There are many examples of data sources (e.g., mobile phone data, traffic data, social media data, and smart card data) with temporal features that could reveal the mobility patterns of human activities (Pan et al., 2013; Wu et al., 2018). For instance, Gong et al. (2015) analyzed the spatiotemporal characteristics of nine daily activity types referring to the trip purposes of taxi trajectory data. Shi et al. (2019) extracted the temporal variation in WeChat (i.e., WeChat is China's most popular messaging app) user density from different land use categories for urban land use classification combined with RS data. The activities of human beings with temporal features can be determined to indicate the social function and patterns of urban land use (Chen et al., 2017; Frias-Martinez and Frias-Martinez, 2014; Pei et al., 2014).
- (3) Semantic features: Photographs are an important element of GBD. Examples include street view photographs, crowd-sourced geo-tagged photos, and social media photos (Kang et al., 2018; Xu et al., 2017). Semantic features obtained from photographs have much in common with those obtained from RS data. However, there are also important distinctions that present challenges for analysis. RS is usually undertaken by national and international organizations following established scientific and engineering principles, including regular acquisition cycles (Ursula et al., 2004). While crowdsourced photographs may be acquired by a range of organizations (e.g., Google Street View) and private individuals. They provide fine-resolution data, but the spatial and temporal sampling may be ad hoc, and the quality can be highly variable (Hu et al., 2015). Recently, with the development of image recognition and deep learning technology, extracting semantic features from photographs and applying them to the perception of places have become possible (Xu et al., 2017).
- (4) Sequence features: Social media data and search engine data have become an important source of GBD in current research (Zheng et al., 2019). Studies utilizing sequence features mainly include the following three aspects (Long and Liu, 2017): (a) to obtain the evaluation index or topic of a place; (b) to obtain information on emotions related to the place, such as happiness or depression (Yang and Mu, 2015; Zheng et al., 2019); (c) to identify public attention to hot events, such as disasters, diseases, and accidents. Specific examples of sequence information include sentiment,

opinions, locations, time, and places. For example, Mitchell et al. (2013) generated a method for analyzing the correlations between human being's real-time expressions and others like emotional, geographic, and demographic characteristics.

4. Integration of RS and GBD for urban land use mapping

The RS and GBD features can be then combined using different approaches for urban land use categorization. According to the fusion mode between RS and GBD features, RS and GBD integration techniques in the literature could be divided into feature-level integration (FI) and decision-level integration (DI).

4.1. Feature-level integration

In FI-based classification, RS features (e.g., spectral, textural, temporal, and spatial features), and GBD features (e.g., spatial, temporal, semantic, sequence features) are first extracted. These features are fused into integrated feature sets for urban land use classification (Fig. 3).

Several efforts have been made for extracting urban land use information using the FI method (Bao et al., 2020; Chang et al., 2020; Hu et al., 2016; Zhang et al., 2017a). For the study unit of the FI-based classification, most studies utilized parcel-level or object-level as the study unit since object-level or parcel-level units are compatible with both RS features and GBD features (Liu et al., 2017). To be more specific, the format of the RS features is usually grid, while for the GBD features is various. It is thus necessary to unify the two kinds of features. Parcel-level classification units can be generated by using the OSM road network or other road data (Huang et al., 2020). The urban parcels are obtained by removing the road buffers from the study site. Sometimes, the study site should also exclude the rivers according to the actual situation. Furthermore, the elevated road in cities would interfere with the segmentation results and need to be considered. Pixel-based and grid-based units are also used for FI-based urban land use classification (Dong et al., 2020).

For the feature extraction stage, the FI-based classification method extracts RS and GBD features respectively. For example, Zhang et al. (2019b) delineated physical features including spectral features (e.g., mean and standard deviation for each band), textural features (e.g., contrast, entropy, correlation, and homogeneity for each band) from RS for urban land use categorization by integrating GBD features (e.g., POI words and real-time Tencent users words). Sun et al. (2020) extracted RS features (spectral and textural features), GBD features (POI frequency, POI spatial distribution), and other features to train the RF classifier for recognizing urban functions (e.g., residence, business, and industries).

In the feature integration stage, the FI method usually classifies RS and GBD features by using machine learning techniques such as RF classifier, SVM, and DT. For example, Gong et al. (2020) proposed a research method that is extracting several features from RS and GBD for training the RF classifier, which has proven to be a new way for mapping urban land use over large areas. Du et al. (2020) generated urban functional zones by removing road buffers, and then the functional zones were classified by coupling Latent Dirichlet Allocation (LDA) and SVM. The proposed method is promising for urban land use mapping over large areas. Recently, deep learning techniques such as CNN and AE models are also used for urban land use mapping, which could help improve the classification performance (Mao et al., 2020).

4.2. Decision-level integration

For categorizing urban land use, DI-based classification combines RS-based classification results with GBD-based classification results. To be more specific, RS and GBD features are evaluated independently before being integrated for urban land use mapping utilizing various modes and methods (Fig. 4).

Compared to the FI method, the basic unit for the DI method is more

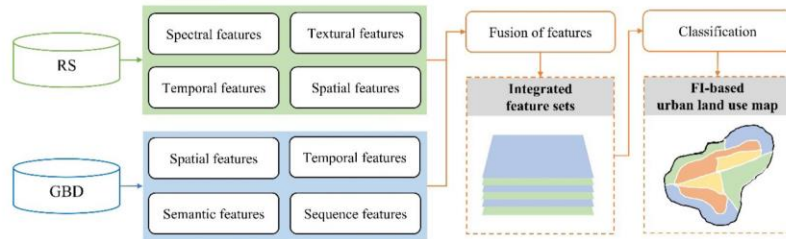


Fig. 3. FI-based categorization strategies.

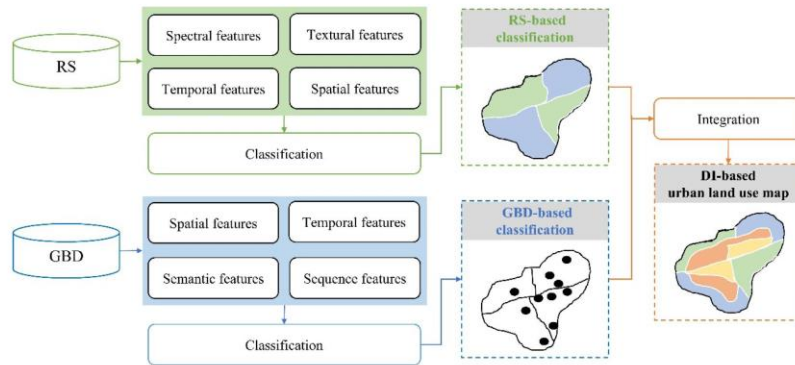


Fig. 4. DI-based categorization strategies.

diverse (Chang et al., 2015). The characteristic of DI-based classification is to combine the RS-based classification and GBD-based classification for extracting urban land use information. It is thus not necessary to unify the spatial units for RS-based classification and GBD-based classification, and the spatial units for the two classification methods will be different. For example, Jia et al. (2018) first classified Gaofen images and mobile phone positioning data by using a support vector machine, and then the two classification results were fused for mapping urban land use based on grid-level units. Tu et al. (2018) utilized the hierarchical cluster analysis to combine RS-based landscape metrics and GBD-based human activity metrics for investigating urban functional patterns in terms of object-based units. In addition, Zhao et al. (2019) delineated the geographical object by using OSM data to train the GNN model. Chen et al. (2018) identified urban green space by utilizing parcels generated by the OSM road networks as the basic units.

There are several methods for the integration of the RS-based classification and GBD-based classification in the DI method based on certain decision fusion strategies such as hierarchical clustering, overlaying, and labeling (Anugraha et al., 2020; Liu et al., 2020b; Xu et al., 2020). For example, Xu et al. (2020) proposed a framework that extracts spatial geographic characteristics from RS images by using deep neural networks and functional distribution characteristics from POIs, and further normalized the two results to identify urban functional regions. Song et al. (2018) used an object-based approach to generate urban objects by using RS data. Then, the objects were further classified and aggregated using POIs. Furthermore, Zhong et al. (2020) presented a method by using the rule-based category mapping (RCM) model to integrate RS-oriented results and GBD-oriented results for extracting

urban functional zones. In Zhong's work, POIs data and RS images were classified through different machine learning methods based on the parcel-level unit, then the two results with different classification systems were combined by weighting the word frequency within each parcel.

5. Discussion

5.1. Advantages and disadvantages of FI-based and DI-based classification

The FI method lies in the realization of considerable information compression of RS data and GBD, which is conducive to real-time processing. This method has the advantages of a low level of human intervention and short processing time. In addition, feature optimization and deep interactions between RS and GBD features can be achieved. Several studies have been analyzed to quantify the relative importance of independent features in the FI method (Zhang et al., 2019b). Furthermore, heterogeneity may occur from variations in data quality, time periods, data formats, and data scales, between RS and GBD, resulting in different representations, descriptions, and interpretations of the goal.

The DI-based classification method has been a fundamental contribution of integrating RS and GBD to urban knowledge (Cai et al., 2017; Zhao et al., 2019). RS and GBD features could be calculated and processed respectively in the DI method, which avoids the feature integrating and conflicting issues. Specifically, RS and GBD with various features could be processed by different methods, and then integrated by

certain models. However, the accuracies of DI-based classification are affected by the two kinds of processing procedures because the mapping results are obtained by overlapping RS-based classification and GBD-based classification. It is thus important to control the procedures for both RS and GBD classification for high accuracy and better performance.

5.2. Limitations and future consideration for RS and GBD integration

The main limitation is that the modality gap (data structure, data format, and data quality) between RS data and GBD, which brings difficulty for the integration. GBD has different sources (modern sensors, geo-tagged web, ground surveying, mobile mapping, and social media platforms), spatiotemporal resolution, and structures from that of RS data (Ali et al., 2017). For example, the data formats of GBD include the image, geo-tagged text, video, and vector. Comparatively, the most commonly used RS data is a raster (Zhu et al., 2019). Furthermore, GBD is not collected evenly across space because the composition of participation (i.e., sensors, platforms, and social habits) in GBD varies across political, cultural, demographics, and commercial factors, while RS data usually has spatially consistent observation frequencies (Chen et al., 2020; Yu et al., 2018). Some areas might have the data sparsity issue of GBD, which might present problems for mapping based on the integration of RS and GBD owing to the lack of GBD. More specifically, social networking platforms such as WeChat, Sina Weibo, and Tencent are widely used in China, while Twitter, Facebook, and Instagram are more popular in Western countries (Li et al., 2018; Wang et al., 2016; Zhou and Zhang, 2016). The emergence of deep learning technologies provides an opportunity to bridge the gap between different data modalities.

The advances in classification algorithms, computing platforms, and data sources are beneficial for mapping urban land use by integrating RS and GBD. Deep learning, as a novel branch of machine learning, establishes fundamental parameters about the data and trains the computer to learn on its own by detecting patterns using a multi-layered approach (Reichstein et al., 2019; Zhu et al., 2017). Several studies have shown that deep learning is particularly effective in integrating RS and GBD for urban region function recognition (Ma et al., 2019; Qin et al., 2018).

Furthermore, the fast growth of cloud computing platforms offers a promising solution for processing large amounts of RS data and GBD. For example, Yin et al. (2021) processed the Sentinel images and POIs data for urban land use classification by using the Google Earth Engine (GEE). GEE could provide a range of data processing methods as well as RS images at various temporal and spatial scales, which is beneficial for improving data computing difficulties (Gorelick et al., 2017). In addition to GEE, other platforms such as Earth Observation Data Center and the Amazon Web Services have also been used for analyzing RS data and GBD on urban land use classification. Furthermore, it is necessary to consider auxiliary data (e.g., census data, statistical data, weather data, hydrological data, and digital elevation data) for more space- and time-referenced information on urban land use classification that integrates RS and GBD (Taubenböck et al., 2009).

5.3. Potential applications of urban land use maps derived from RS and GBD integration

Urban land use maps integrating RS and emerging GBD provide more potential in urban management such as urban planning, urban environment assessment, urban disaster monitoring, and urban traffic analysis (Fig. 5).

A more scientific and efficient urban planning system will benefit from the urban land use map integrating RS and GBD. For example, Xing and Meng (2018) extracted urban land use information by integrating landscape metrics from RS images and semantic features from GBD, which plays as an indicator in urban planning and management. Chen et al. (2018) identified urban green space (e.g., municipal park, community park, etc.) by extracting land cover features from RS data and land use features from POIs for the urban green space planning, which could assist government departments in urban green space planning. In general, integrating RS and GBD would help to improve the city function from the "hard" physical environment and the "soft" services. The addition of GBD for urban land use maps with adequate and timely information could enhance the feedback loops of urban insights for urban governors and planners.

The current advances of urban land use maps that integrate RS data and GBD make it capable of monitoring urban environments such as

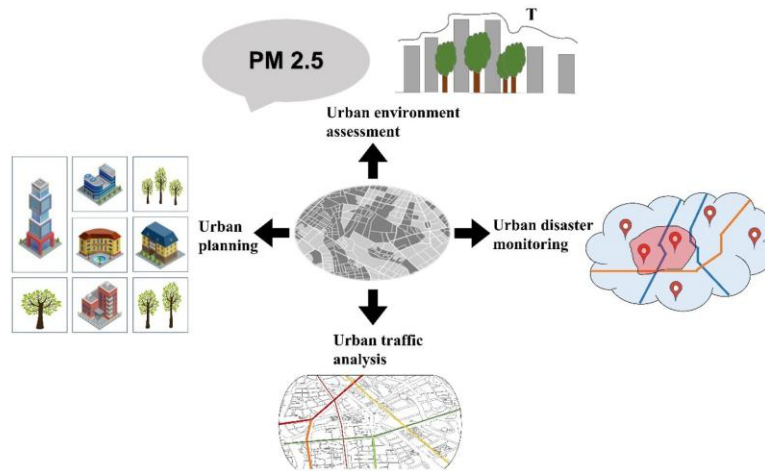


Fig. 5. Urban land use map integrating RS and GBD for urban management.

urban heat island and air pollution (Halim et al., 2020; Masoudi et al., 2021; Venter et al., 2020; Wang et al., 2021). For example, Song et al. (2019) analyzed the relationship between urban functional regions and air pollutant emissions and presented a cost-effective way of mapping spatiotemporal patterns of air pollution by utilizing the urban land use map that integrates RS images and POIs. Luan et al. (2020) quantified the impacts of urban natural surfaces and non-surface human activities on urban heat islands by using RS data and GBD, and explored the relationships between urban heat islands and urban land use patterns. This evidence demonstrated that air pollution concentrations are associated with RS-based urban land cover (e.g., industrial layout) as well as GBD-based urban land use (e.g., travel behavior). Other applications have also proved that the spatial patterns of the urban environments have strong relationships with the urban land use patterns (Pan et al., 2013).

Massive information from urban land use maps that integrate RS and quick-updated GBD is generated continuously and dynamically, providing resources to aid in disaster analysis of historical and future occurrences (earthquakes, fires, or floods) (Huang et al., 2017; Li et al., 2017b). For instance, Li et al. (2019) reviewed recent research that utilized RS and GBD for urban disaster information detection including the suffering area, suffering location, and suffering pattern, which provided a useful method for disaster management. Furthermore, Cervone et al. (2016) proposed a novel framework for urban damage assessment under severe weather by using RS images and real-time Twitter data, which were then combined with other GBD in order to get abundant information in disaster areas. Quickly updated urban land use maps that integrate RS and GBD with fine spatial resolution could support urban disaster management by providing unprecedented reference data.

Up to date, urban traffic conditions have become serious problems that threatening human's daily living quality. Several studies have been made on the comprehensive analysis of urban traffic by taking advantage of RS data and GBD on urban land. Applications in traffic quality analysis, for example, classifying Shanghai city into six traffic "source-sink" areas according to the pick-ups and drop-offs of traffic data and LandScan product (Liu et al., 2012). Improved urban land use maps could also provide technical support for the transportation of the Smart Cities (Zanella et al., 2014). A thorough perception of urban traffic conditions could be achieved through the integration of RS and GBD.

6. Conclusions

This study examined the applications of RS and GBD features in urban land use categorization, as well as methods for RS and GBD integration. The analysis of the existing literature concludes that the emerging GBD provides opportunities for the transformation from urban land cover (physical environment) to urban land use (living environment). Applications on the urban land use maps integrating RS and GBD for urban management mainly include urban planning, urban environment assessment, urban disaster monitoring, and urban traffic analysis. Deeper understandings of the urban surface can be acquired by adding GBD values to the traditional urban RS works. As the integration of RS and GBD has become more generalized, significant progress can be already seen for urban management. Also, integrating RS and GBD on urban land use provides an opportunity for putting people at the center of processes of knowledge and management of the urban planet.

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.

Acknowledgements

This study was supported by the Key Research Program of Frontier Sciences (QYZDB-SSW-DQC005) and the Strategic Priority Research

Program (XDA19040301) of the Chinese Academy of Sciences (CAS), the National Natural Science Foundation of China (41871349, 41801336, 41971078). Jiadi Yin is supported by a Ph.D. scholarship from the University of Nottingham Ningbo China as part of a doctoral training partnership with the Institute for Geographic Science and Natural Resources Research (IGSNRR), CAS.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijag.2021.102514>.

References

- Ali, A.L., Falomir, Z., Schmid, F., Freksa, C., 2017. Rule-guided human classification of Volunteered Geographic Information. *ISPRS J. Photogram. Rem. Sens.* 127, 3–15.
- Anugraha, A., Chu, H.J., Ali, M., 2020. Social sensing for urban land use identification. *ISPRS Int. J. Geo-Inf.* 9 (9), 550. <https://doi.org/10.3390/ijgi9090550>.
- Arlino, O., Gross, D., Ramera, F., Leroy, M., Bicheron, P., Brockmann, C., Defourny, P., Vancutsem, C., Achard, F., Durieux, L., 2007. GlobCover: ESA service for global land cover from MERIS. *IGARSS 2007*, 2412–2415.
- Bao, H., Ming, D., Guo, Y.A., Zhang, K., Zhou, K., Du, S., 2020. DFCNN based semantic recognition of urban functional zones by integrating remote sensing data and POI data. *Remote Sens.* 12 (7), 1088. <https://doi.org/10.3390/rs12071088>.
- Bartholomé, E., Belward, A.S., 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. *Int. J. Remote Sens.* 26 (9), 1959–1977.
- Brovelli, M.A., Zamboni, G., Arias Muñoz, C., 2015. From paper maps to the Digital Earth and the Internet of Places. *Rendiconti Lincei*. 26 (S1), 97–103.
- Cadavid Restrepo, A.M., Yang, Y.R., Hanani, N.A.S., Gray, D.J., Barnes, T.S., Williams, G. M., Soares Magalhães, R.J., McManus, D.P., Guo, D., Clements, A.C.A., 2017. Land cover change during a period of extensive landscape restoration in Ningxia Hui Autonomous Region, China. *Sci. Total Environ.* 598, 669–679.
- Cai, J., Huang, B.O., Song, Y., 2017. Using multi-source geospatial big data to identify the structure of polycentric cities. *Rem. Sens. Environ.* 202, 210–221.
- Cao, R., Tu, W., Yang, C., Li, Q., Liu, J., Zhu, J., Zhang, Q., Li, Q., Qiu, G., 2020. Deep learning based remote and social sensing data fusion for urban region function recognition. *ISPRS J. Photogram. Rem. Sens.* 163, 82–97.
- Cervone, G., Sava, E., Huang, Q., Schnebele, E., Harrison, J., Waters, N., 2016. Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study. *Int. J. Remote Sens.* 37 (1), 100–124.
- Chang, C., Ye, Z., Huang, Q., Wang, C., 2015. An integrative method for mapping urban land use change using "geo-sensor" data. In: *UrbanGIS'15: Proceedings of the 1st International ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics*, pp. 47–54.
- Chang, S., Wang, Z., Mao, D., Guan, K., Jia, M., Chen, C., 2020. Mapping the essential urban land use in Changchun by applying random forest and multi-source geospatial data. *Remote Sens.* 12 (15), 2488. <https://doi.org/10.3390/rs12152488>.
- Chen, B., Song, Y., Huang, B.O., Xu, B., 2020. A novel method to extract urban human settlements by integrating remote sensing and mobile phone locations. *Science of Remote Sensing*, 1, 100003. <https://doi.org/10.1016/j.srs.2020.100003>.
- Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., Mills, J., 2015. Global land cover mapping at 30m resolution: A POK based operational approach. *ISPRS J. Photogram. Rem. Sens.* 103, 7–27.
- Chen, W., Huang, H., Dong, J., Zhang, Y., Tian, Y., Yang, Z., 2018. Social functional mapping of urban green space using remote sensing and social sensing data. *ISPRS J. Photogram. Rem. Sens.* 146, 436–452.
- Chen, Y., Liu, X., Li, X., Liu, X., Yao, Y., Hu, G., Xu, X., Pei, F., 2017. Delineating urban functional areas with building-level social media data: A dynamic time warping (DTW) distance based k-medoids method. *Landscape Urban Plan.* 160, 48–60.
- CIESIN, IEPRI, CIAT, 2011. *Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Land and Geographic Unit Area Grids*. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).
- Colston, B.D., C. E., Huang, C., Wang, P., Tilton, J.C., Tan, B., Phillips, J., Niemczura, S., Ling, P.Y., Wolfe, R.E., 2017. *Global Man-made Impervious Surface (GMIS) Dataset From Landsat*. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).
- Defourny, P., Kirches, G., Brockmann, C., Boettcher, M., Peters, M., Bontemps, S., Lamarche, C., Schlerf, M., M., S., 2018. *Land Cover CCI: Product User Guide Version 2*.
- Deng, Jinsong, Huang, Yibo, Chen, Binjie, Tong, Cheng, Liu, Pengbo, Wang, Hongquan, Hong, Yang, 2019. A methodology to monitor urban expansion and green space change using a time series of multi-sensor SPOT and Sentinel-2A images. *Remote Sens.* 11 (10), 1230. <https://doi.org/10.3390/rs11101230>.
- Dong, J., Xiao, X., 2016. Evolution of regional to global paddy rice mapping methods: A review. *ISPRS J. Photogram. Rem. Sens.* 119, 214–227.
- Dong, X., Xu, Y., Huang, L., Liu, Z., Xu, Y., Zhang, K., Hu, Z., Wu, G., 2020. Exploring impact of spatial unit on urban land use mapping with multisource data. *Remote Sens.* 12.
- Du, S.J., Du, S.H., Liu, B., Zhang, X.Y., Zheng, Z.J., 2020. Large-scale urban functional zone mapping by integrating remote sensing images and open social data. *Gisci Remote Sens.* 57, 411–430.

- Ellis, E.C., Ramankutty, N., 2008. Putting people in the map: anthropogenic biomes of the world. *Front Ecol Environ.* 6, 439–447.
- Elmqvist, T., Andersson, E., Fraunteszkaki, N., McPhearson, T., Olsson, P., Gaffney, O., Takeuchi, K., Folke, C., 2019. Sustainability and resilience for transformation in the urban century. *Nat. Sustain.* 2, 267–273.
- Elvidge, C.D., Tuttle, B.T., Sutton, P.C., Baugh, K.E., Howard, A.T., Milesi, C., Bhaduri, B. L., Nemani, R., 2007. Global distribution and density of constructed impervious surfaces. *Sensors.* 1962–1979.
- Esch, T., Taubenböck, H., Roth, A., Heldens, W., Felber, A., Thiel, M., Schmidt, M., Mueller, A., Dech, S., 2012. TanDEM-X mission new perspectives for the inventory and monitoring of global settlement patterns. *J. Appl. Remote Sens.* 6, 6.
- Frias-Martinez, V., Frias-Martinez, E., 2014. Spectral clustering for sensing urban land use using Twitter activity. *Eng Appl Artif Intell.* 35, 237–245.
- Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schneider, A., Cooper, A., Bacchini, A., Gao, F., Schaaf, C., 2002. Global land cover mapping from MODIS: algorithms and early results. *Rem. Sens. Environ.* 83, 287–302.
- Gong, L., Liu, X., Wu, L., Liu, Y., 2015. Inferring trip purposes and uncovering travel patterns from taxi trajectory data. *Cartogr Geogr Inf Sci.* 43, 103–114.
- Gong, P., Chen, B., Li, X., Liu, H., Wang, J., Bai, Y., Chen, J., Chen, X., Fang, L., Feng, S., Feng, Y., Gong, Y., Gu, H., Huang, H., Huang, X., Jiao, H., Kang, Y., Lei, G., Li, A., Li, X., Li, X., Li, Y., Li, Z., Li, Z., Liu, C., Liu, M., Liu, S., Mao, W., Miao, C., Ni, H., Pan, Q., Qi, S., Ren, Z., Shen, Z., Shen, S., Shi, M., Song, Y., Su, M., Sun, H.P., Sun, B., Sun, F., Sun, J., Sun, L., Sun, W., Tian, T., Tong, X., Tseng, Y., Tu, Y., Wang, H., Wang, L., Wang, X., Wang, Z., Wu, T., Xie, Y., Yang, J., Yang, J., Yuan, M., Yue, W., Zeng, H., Zhang, K., Zhang, N., Zhang, T., Zhang, Y., Zhao, F., Zheng, Y., Zhou, Q., Clinton, N., Zhu, Z., Xu, B., 2020. Mapping essential urban land use categories in China (EULUC China): preliminary results for 2018. *Sci. Bull.* 65, 182–187.
- Gong, P., Howarth, P.J., 1989. Performance analyses of probabilistic relaxation methods for land-cover classification. *Rem. Sens. Environ.* 30, 33–42.
- Gong, P., Li, X., Zhang, W., 2019a. 40-Year (1978–2017) human settlement changes in China reflected by impervious surfaces from satellite remote sensing. *Sci. Bull.* 64, 756–763.
- Gong, P., Liu, H., Zhang, M., Li, C., Wang, J., Huang, H., Clinton, N., Ji, L., Li, W., Bai, Y., Chen, B., Xu, B., Zhu, Z., Yuan, C., Ping Suen, H., Guo, J., Xu, N., Li, W., Zhao, Y., Yang, J., Yu, C., Wang, X., Fu, H., Yu, L., Dronova, I., Hui, F., Cheng, X., Shi, X., Xiao, F., Liu, Q., Song, L., 2019b. Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017. *Sci. Bull.* 64, 370–373.
- Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., 2013. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. *Int. J. Remote Sens.* 34, 2607–2654.
- Goodchild, M.F., Glennon, J.A., 2010. Crowdsourcing geographic information for disaster response: a research frontier. *Int. J. Digit Earth.* 3, 231–241.
- Gorelick, N., Haucher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Rem. Sens. Environ.* 202, 18–27.
- Halim, N.D.A., Latif, M.T., Mohamed, A.F., Maulud, K.N.A., Idrus, S., Azhari, A., Othman, M., Sofwan, N.M., 2020. Spatial assessment of land use impact on air quality in mega urban regions. *Malaysia. Sustain. Cities Soc.* 63, 102436.
- Hansen, M.C., Defries, R.S., Townshend, J.R.G., Sohlberg, R., 2000. Global land cover classification at 1km spatial resolution using a classification tree approach. *Int. J. Remote Sens.* 21, 1331–1364.
- Haralick, R.M., Shanmugam, K., Dinstein, I.H., 1973. Textural features for image classification. *Studies in Media and Communication.* 3, 610–621.
- Heiden, U., Segl, K., Roessner, S., Kaufmann, H., 2007. Determination of robust spectral features for identification of urban surface materials in hyperspectral remote sensing data. *Rem. Sens. Environ.* 111, 537–552.
- Hellich, M., Amelunxen, C., Neis, P., Zipf, A., 2012. In: Comparative spatial analysis of positional accuracy of OpenStreetMap and proprietary Geodata. *Society and Learning, Salzburg, Germany*, pp. 4–6.
- Herold, M., Gardner, M.E., Roberts, D.A., 2003. Spectral resolution requirements for mapping urban areas. *IEEE Trans Geosci Remote Sens.* 41, 1907–1919.
- Hersperger, A.M., Oliveira, E., Pagliarin, S., Palka, G., Verburg, P., Bolliger, J., Grdinariu, S., 2018. Urban land-use change: The role of strategic spatial planning. *Glob Environ Change.* 51, 32–42.
- Howarth, P.J., Boasson, E., 1983. Landsat digital enhancements for change detection in urban environments. *Rem. Sens. Environ.* 13, 149–160.
- Hu, T., Yang, J., Li, X., Gong, P., 2016. Mapping urban land use by using Landsat images and open social data. *Remote Sens.* 8, 151.
- Hu, Y., Gao, S., Janowicz, K., Yu, B., Li, W., Prasad, S., 2015. Extracting and understanding urban areas of interest using geotagged photos. *Comput Environ Urban Syst.* 54, 240–254.
- Huang, Q., Cervone, G., Zhang, G., 2017. A cloud-enabled automatic disaster analysis system of multi-sourced data streams: An example synthesizing social media, remote sensing and Wikipedia data. *Comput Environ Urban Syst.* 66, 23–37.
- Huang, R., Taubenböck, H., Mou, L., Zhu, X., 2018. Classification of settlement types from Tweets using LDA and LSTM. *IGARSS 2018*, 6408–6411.
- Huang, Z., Qi, H., Kang, C., Su, Y., Liu, Y., 2020. An ensemble learning approach for urban land use mapping based on remote sensing imagery and social sensing data. *Remote Sens.* 12.
- Huang, X., Huang, J., Wen, D., Li, J., 2021. An updated MODIS global urban extent product (MGUP) from 2001 to 2018 based on an automated mapping approach. *Int. J. Appl Earth Obs Geoinf.* 95.
- Ilieva, R.T., McPhearson, T., 2018. Social-media data for urban sustainability. *Nat. Sustain.* 1, 553–565.
- Jia, Y., Ge, Y., Ling, F., Guo, X., Wang, J., Wang, L., Chen, Y., Li, X., 2018. Urban land use mapping by combining remote sensing imagery and mobile phone positioning data. *Remote Sens.* 10.
- Jiang, Q., Kresin, F., Arnold K. Bregt, Kooistra, L., Pareschi, E., Putten, E.v., Volten, H., Wesseling, J., 2016. Citizen sensing for improved urban environmental monitoring. *J. Sensors.* 1–9.
- Jin, H., Mountrakis, G., Stehman, S.V., 2014. Assessing integration of intensity, polarimetric scattering, interferometric coherence and spatial texture metrics in PALSAR-derived land cover classification. *ISPRS J. Photogramm. Rem. Sens.* 98, 70–84.
- Johnson, B.A., Iizuka, K., 2016. Integrating OpenStreetMap crowdsourced data and Landsat time-series imagery for rapid land use/land cover (LULC) mapping: Case study of the Laguna de Bay area of the Philippines. *Appl Geogr.* 67, 140–149.
- Kang, J., Körner, M., Wang, Y., Taubenböck, H., Zhu, X.X., 2018. Building instance classification using street view images. *ISPRS J. Photogramm. Rem. Sens.* 145, 44–59.
- Kitchin, R., 2013. Big data and human geography. *Dialogues in Human Geography.* 3, 262–267.
- Kitchin, R., McArdle, G., 2016. What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society.* 3.
- Krylov, V., Kenny, E., Dalrymple, R., 2018. Automatic discovery and geotagging of objects from street view imagery. *Remote Sens.* 10, 661.
- Kuffer, M., Pfeffer, K., Sluzas, R., 2016. Slums from space-15 years of slum mapping using remote sensing. *Remote Sens.* 8.
- Li, J., Benediktsson, J.A., Zhang, B., Yang, T., Plaza, A., 2017a. Spatial technology and social media in remote sensing: challenges and opportunities. *P IEEE.* 105, 1583–1585.
- Li, J., He, Z., Plaza, J., Li, S., Chen, J., Wu, H., Wang, Y., Liu, Y., 2017b. Social media: New perspectives to improve remote sensing for emergency response. *P IEEE.* 105, 1900–1912.
- Li, J., Ying, L., Dang, A., 2018. Live-Work-Play Centers of Chinese cities: Identification and temporal evolution with emerging data. *Comput Environ Urban Syst.* 71, 58–66.
- Li, S., Dragicevic, S., Castro, F.A., Sester, M., Winter, S., Coltekin, A., Pettit, C., Jiang, B., Howarth, J., Stein, A., Cheng, T., 2016. Geospatial big data handling theory and methods: A review and research challenges. *ISPRS J. Photogramm. Rem. Sens.* 115, 119–133.
- Li, W., Dong, R., Fu, H., Wang, J., Yu, L., Gong, P., 2020a. Integrating Google Earth imagery with Landsat data to improve 30-m resolution land cover mapping. *Rem. Sens. Environ.* 237, 111563.
- Li, X., Hu, T., Gong, P., Du, S., Chen, B., Li, X., Dai, Q., 2021. Mapping Essential Urban Land Use Categories in Beijing with a Fast Area of Interest (AOI)-Based Method. *Remote Sens.* 13.
- Li, X., Zhou, Y., Zhu, Z., Cao, W., 2020b. A national dataset of 30 m annual urban extent dynamics (1985–2015) in the conterminous United States. *Earth Syst. Sci. Data.* 12, 357–371.
- Li, Z., Huang, Q., Enrich, C.T., 2019. Introduction to social sensing and big data computing for disaster management. *Int. J. Digit Earth.* 12, 1198–1204.
- Liu, D., Chen, N., Zhang, X., Wang, C., Du, W., 2020a. Annual large-scale urban land mapping based on Landsat time series in Google Earth Engine and OpenStreetMap data: A case study in the middle Yangtze River basin. *ISPRS J. Photogramm. Rem. Sens.* 159, 337–351.
- Liu, H., Xu, Y., Tang, J., Deng, M., Huang, J., Yang, W., Wu, F., 2020b. Recognizing urban functional zones by a hierarchical fusion method considering landscape features and human activities. *Trans GIS.* 24, 1359–1381.
- Liu, J., Li, J., Li, W., Wu, J., 2016. Rethinking big data: A review on the data quality and usage issues. *ISPRS J. Photogramm. Rem. Sens.* 115, 134–142.
- Liu, X., He, J., Yao, Y., Zhang, J., Liang, H., Wang, H., Hong, Y., 2017. Classifying urban land use by integrating remote sensing and social media data. *Int. J. Geogr. Inf. Sci.* 31, 1675–1696.
- Liu, X., Hu, G., Chen, Y., Li, X., Xu, X., Li, S., Pei, F., Wang, S., 2018. High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform. *Rem. Sens. Environ.* 209, 227–239.
- Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., Chi, G., Shi, L., 2015. Social sensing: A new approach to understanding our socioeconomic environments. *Ann Assoc Am Geogr.* 105, 512–530.
- Liu, Y., Wang, F., Xiao, Y., Gao, S., 2012. Urban land uses and traffic 'source-sink areas': Evidence from GPS-enabled taxi data in Shanghai. *Landscape Urban Plan.* 106, 73–87.
- Long, Y., Liu, L., 2017. How green are the streets? An analysis for central areas of Chinese cities using Tencent Street View. *PLoS One.* 12.
- Long, Y., Zhai, W., Shen, Y., Ye, X., 2018. Understanding uneven urban expansion with natural cities using open data. *Landscape Urban Plan.* 177, 281–293.
- Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L., Merchant, J.W., 2000. Development of a global land cover characteristics database and IGBP DIScover from 1 km AVHRR data. *Int. J. Remote Sens.* 21, 1303–1330.
- Liu, D., Weng, Q., 2006. Use of impervious surface in urban land use classification. *Rem. Sens. Environ.* 102, 146–160.
- Luan, X.L., Yu, Z.W., Zhang, Y.T., Wei, S., Miao, X.Y., Huang, Z.Y.X., Teng, S.Q.N., Xu, C., 2020. Remote sensing and social sensing data reveal scale dependent and system-specific strengths of urban heat island determinants. *Remote Sens.* 12.
- Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B.A., 2019. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Rem. Sens.* 152, 166–177.
- Mao, W., Lu, D., Hou, L., Liu, X., Yue, W., 2020. Comparison of machine-learning methods for urban land-use mapping in Hangzhou City, China. *Remote Sens.* p. 12.

- Marti, P., Serrano Estrada, L., Nolasco Cirugeda, A., 2019. Social Media data: Challenges, opportunities and limitations in urban studies. *Comput Environ Urban Syst.* 74, 161–174.
- Masoudi, M., Tan, P.Y., Fadaei, M., 2021. The effects of land use on spatial pattern of urban green spaces and their cooling ability. *Urban Clim.* 35, 100743.
- Mitchell, L., Frank, M.R., Harris, K.D., Dodds, P.S., Danforth, C.M., 2013. The geography of happiness: connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS One.* 8, 1–15.
- Okujeni, A., Gamers, F., Cooper, S.D., Degeerickx, J., Heiden, U., Hostert, P., Friem, F., Roberts, D.A., Somers, B., van der Linden, S., 2018. Generalizing machine learning regression models using multi-site spectral libraries for mapping vegetation impervious soil fractions across multiple cities. *Rem. Sens. Environ.* 216, 482–496.
- Pacifici, F., Chini, M., Emery, W.J., 2009. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land use classification. *Rem. Sens. Environ.* 113, 1276–1292.
- Pan, G., Qi, G., Wu, Z., Zhang, D., Li, S., 2013. Land-Use Classification Using Taxi GPS Traces. *IEEE Trans. Intell. Transp. Syst.* 14, 113–123.
- Patino, J.E., Duque, J.C., 2013. A review of regional science applications of satellite remote sensing in urban settings. *Comput Environ Urban Syst.* 37, 1–17.
- Pei, T., Sobolevsky, S., Ratti, C., Shaw, S.-L., Li, T., Zhou, C., 2014. A new insight into land use classification based on aggregated mobile phone data. *Int J Geogr Inf Sci.* 28, 1988–2007.
- Pesaresi, M., Gao, H., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., Halkia, M., Kuitmann, M., Kemper, T., Lu, L., Marin-Herrera, M.A., Ouzounis, G.K., Scavazzon, M., Soille, P., Syrris, V., Zanchetta, L., 2013. A global human settlement layer from optical HR/VHR RS data: Concept and first results. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* 6, 2102–2131.
- Plugnacher, D., Rabe, A., Peters, M., Hostert, P., 2019. Mapping pan-European land cover using Landsat spectral-temporal metrics and the European LUCAS survey. *Remote Sens. Environ.* 221, 583–595.
- Qi, L., Li, J., Wang, Y., Gao, X., 2019. Urban Observation: Integration of remote sensing and social media data. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* 1–13.
- Qin, Y., Chi, M., Liu, X., Zhang, Y., Zeng, Y., Zhao, Z., 2018. Classification of high resolution urban remote sensing images using deep networks by integration of social media photos. *IGARSS 2018*, 7243–7246.
- Rebu, M., Seto, K.C., 2020. A systematic review and assessment of algorithms to detect, characterize, and monitor urban land change. *Rem. Sens. Environ.* 242, 111739.
- Reichstein, M., Camps Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven Earth system science. *Nature.* 566, 195–204.
- Schneider, A., Friedl, M.A., Potere, D., 2010. Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions'. *Remote Sens. Environ.* 114, 1733–1746.
- Seto, K.C., Fragkias, M., Gueneralp, B., Reilly, M.K., 2011. A meta-analysis of global urban land expansion. *PLoS One.* 6.
- Shi, Y., Qi, Z., Liu, X., Niu, N., Zhang, H., 2019. Urban land use and land cover classification using multisource remote sensing images and social media data. *Remote Sens.* 11, 2719.
- Song, J., Liu, T., Li, X., Prishchepov, A.V., 2018. Mapping urban functional zones by integrating very high spatial resolution remote sensing imagery and points of interest: A case study of Xiamen, China. *Remote Sens.* 10, 1737.
- Song, J., Zhao, C., Lin, T., Li, X., Prishchepov, A.V., 2019. Spatio-temporal patterns of traffic-related air pollutant emissions in different urban functional zones estimated by real-time video and deep learning technique. *J. Clean. Prod.* 238, 117881.
- Song, X., Feng, Q., Xia, F., Li, X., Scheffran, J., 2021. Impacts of changing urban land-use structure on sustainable city growth in China: A population-density dynamics perspective. *Habitat Int.* 107, 102296.
- Srivastava, S., Vargas Muñoz, J.E., Lobry, S., Tuia, D., 2018. Fine grained landuse characterization using ground-based pictures: a deep learning solution based on globally available data. *Int J Geogr Inf Sci.* 1–20.
- Steffen, W., Grinevald, J., Crutzen, P., McNeill, J., 2011. The Anthropocene: conceptual and historical perspectives. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences.* 369, 842–867.
- Sun, J., Wang, H., Song, Z., Liu, J., Meng, P., Qin, S., 2020. Mapping essential urban land use categories in Nanjing by integrating multi-source Big Data. *Remote Sens.* 12, Taubenböck, H., Wurm, M., Setiadi, N., Gebert, N., Roth, A., Strunz, G., Birkmann, J., Dech, S., 2009. Integrating remote sensing and social science: The correlation of urban morphology with socioeconomic parameters. *Joint Urban Remote Sensing Event.* 2009, 1–7.
- Tu, W., Hu, Z., Li, L., Cao, J., Jiang, J., Li, Q., Li, Q., 2018. Portraying urban functional zones by coupling remote sensing imagery and human sensing data. *Remote Sens.* 10.
- Ursula, C., Benz, and, Peter, Hofmann, and, Gregor, Willhauck, and, 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS J. Photogram. Rem. Sens.* 58, 239–258.
- Venter, Z.S., Brasseur, O., Esau, I., Mester, F., 2020. Hypertical mapping of urban air temperature using remote sensing and crowdsourced weather data. *Rem. Sens. Environ.* p. 242.
- Wan, T., Lu, H., Lu, Q., Luo, N., 2017. Classification of high resolution remote sensing image using OpenStreetMap information. *IEEE Geosci. Remote. Sens. Lett.* 14, 2305–2309.
- Wang, G., Han, Q., de vries, B., 2021. The multi-objective spatial optimization of urban land use based on low-carbon city planning. *Ecol Indic* 125, 107540.
- Wang, P., Huang, C., Brown de Colstoun, E.C., Tilton, J.C., Tan, B., 2017. Global human built-up and settlement extent (HBASE) dataset from Landsat. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY.
- Wang, Y., Wang, T., Tsou, M.-H., Li, H., Jiang, W., Guo, F., 2016. Mapping dynamic urban land use patterns with crowdsourced geo-tagged social media (Sina-Weibo) and commercial Points of interest collections in Beijing, China. *Sustainability.* 8, 1202.
- Weng, Q., Hu, X., Liu, H., 2009. Estimating impervious surfaces using linear spectral mixture analysis with multitemporal ASTER images. *Int. J. Remote Sens.* 30, 4807–4830.
- Wu, H., Zhang, T., Gong, J., 2015. GeoComputation for geospatial big data. *Trans GIS* 18, 1–2.
- Wu, L., Cheng, X., Kang, C., Zhu, D., Huang, Z., Liu, Y., 2018. A framework for mixed use decomposition based on temporal activity signatures extracted from big geo-data. *Int J Digit Earth.* 1–19.
- Wurm, M., Taubenböck, H., Weigand, M., Schmitt, A., 2017. Slum mapping in polarimetric SAR data using spatial features. *Rem. Sens. Environ.* 194, 190–204.
- Xing, H., Meng, Y., 2018. Integrating landscape metrics and socioeconomic features for urban functional region classification. *Comput Environ Urban Syst.* 72, 134–145.
- Xiong, G., Cao, X., Hamm, N.A.S., Lin, T., Zhang, G., Chen, B., 2021. Unbalanced Development Characteristics and Driving Mechanisms of Regional Urban Spatial Form: A Case Study of Jiangsu Province, China. *Sustainability.* p. 13.
- Xu, G., Zhu, X., Fu, D., Dong, J., Xiao, X., 2017. Automatic land cover classification of geo-tagged field photos by deep learning. *Environ Model Softw.* 91, 127–134.
- Xu, S., Qing, L., Han, L., Liu, M., Peng, Y., Shen, L., 2020. A new remote sensing images and point-of-interest fused (RPF) model for sensing urban functional regions. *Remote Sens.* 12.
- Yammine, S.Z., Jarreau, P.B., Liu, C., Coe, I.R., 2018. Social media for social change in science. *Science.* 360, 163.
- Yang, W., Mu, L., 2015. GIS analysis of depression among Twitter users. *Appl Geogr.* 60, 217–223.
- Yao, Y., Zhang, J., Hong, Y., Liang, H., He, J., 2018. Mapping fine-scale urban housing prices by fusing remotely sensed imagery and social media data. *Trans GIS.* 22, 561–581.
- Ye, X., Huang, Q., Li, W., 2016. Integrating big social data, computing and modeling for spatial social science. *Cartogr Geogr Inf Sci.* 43, 377–378.
- Yin, J., Fu, P., Hamm, N.A.S., Li, Z., You, N., He, Y., Cheshmehzangi, A., Dong, J., 2021. Decision-level and feature-level integration of remote sensing and geospatial big data for urban land use mapping. *Remote Sens.* 13.
- Yu, B., Lian, T., Huang, Y., Yao, S., Ye, X., Chen, Z., Yang, C., Wu, J., 2019. Integration of nighttime light remote sensing images and taxi GPS tracking data for population surface enhancement. *Int J Geogr Inf Sci.* 33, 687–706.
- Yu, Y., Li, J., Zhu, C., Plaza, A., 2018. Urban impervious surface estimation from remote sensing and social data. *Photogramm Eng Remote Sensing.* 84, 771–780.
- Zanella, A., Bui, N., Castellani, A., Vangelista, L., Zorzi, M., 2014. Internet of Things for Smart Cities. *IEEE Internet Things J.* 1, 22–32.
- Zhan, X., Ukusuri, S.V., Zhu, F., 2014. Inferring urban land use using large-scale social media check-in data. *Netw. Spat. Econ.* 14, 647–667.
- Zhang, C., Xu, L., Yan, Z., Wu, S., 2021a. A GloVe-Based POI Type Embedding Model for Extracting and Identifying Urban Functional Regions. *ISPRS Int. J. Geoinf.* p. 10.
- Zhang, F., Wu, L., Zhu, D., Liu, Y., 2019a. Social sensing from street-level imagery: A case study in learning spatio-temporal urban mobility patterns. *ISPRS J. Photogram. Rem. Sens.* 153, 48–58.
- Zhang, J., Li, X., Yao, Y., Hong, Y., He, J., Jiang, Z., Sun, J., 2021b. The Traj2Vec model to quantify residents' spatial trajectories and estimate the proportions of urban land-use types. *Int J Geogr Inf Sci.* 35, 193–211.
- Zhang, X., Du, S., Wang, Q., 2017a. Hierarchical semantic cognition for urban functional zones with VHR satellite images and POI data. *ISPRS J. Photogram. Rem. Sens.* 132, 170–184.
- Zhang, X., Du, S., Zheng, Z., 2020. Heuristic sample learning for complex urban scenes: Application to urban functional-zone mapping with VHR images and POI data. *ISPRS J. Photogram. Rem. Sens.* 161, 1–12.
- Zhang, Y., Li, Q., Huang, H., Wu, W., Du, X., Wang, H., 2017b. The combined use of remote sensing and social sensing data in fine-grained urban land use mapping: A case study in Beijing, China. *Remote Sens.* 9, 865.
- Zhang, Y., Li, Q., Tu, W., Mai, K., Yao, Y., Chen, Y., 2019b. Functional urban land use recognition integrating multi-source geospatial data and cross-correlations. *Comput Environ Urban Syst.* 78, 101374.
- Zhao, N., Cao, G., Zhang, W., Samson, E.L., 2018. Tweets or nighttime lights: Comparison for preeminence in estimating socioeconomic factors. *ISPRS J. Photogram. Rem. Sens.* 146, 1–10.
- Zhao, W., Bo, Y., Chen, J., Tiede, D., Blaschke, T., Emery, W.J., 2019. Exploring semantic elements for urban scene recognition: Deep integration of high-resolution imagery and OpenStreetMap (OSM). *ISPRS J. Photogram. Rem. Sens.* 151, 237–250.
- Zheng, S., Wang, J., Sun, C., Zhang, X., Kahn, M.E., 2019. Air pollution lowers Chinese urbanites' expressed happiness on social media. *Nat. Hum. Behav.* 3, 237–243.
- Zhong, Y., Su, Y., Wu, S., Zheng, Z., Zhao, J., Ma, A., Zhu, Q., Ye, R., Li, X., Pellikka, P., Zhang, L., 2020. Open-source data-driven urban land-use mapping integrating point-line polygon semantic objects: A case study of Chinese cities. *Rem. Sens. Environ.* p. 247.
- Zhou, X., Zhang, L., 2016. Crowdsourcing functions of the living city from Twitter and Foursquare data. *Cartogr Geogr Inf Sci.* 43, 393–404.
- Zhou, Y., Li, X., Asrar, G.R., Smith, S.J., Imhoff, M., 2018. A global record of annual urban dynamics (1992–2013) from nighttime lights. *Rem. Sens. Environ.* 219, 206–220.
- Zhu, X., Tuia, D., Moti, L., Xia, G., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geosci. Remote Sens. Mag.* 5, 8–36.

J. Yin et al.

Zhu, Z., Woodcock, C.E., Rogan, J., Kellndorfer, J., 2012. Assessment of spectral, polarimetric, temporal, and spatial dimensions for urban and peri-urban land cover classification using Landsat and SAR data. *Rem. Sens. Environ.* 117, 72–82.

Zhu, Z., Zhou, Y., Seto, K.C., Stokes, E.C., Deng, C., Pickett, S.T.A., Taubenböck, H., 2019. Understanding an urbanizing planet: Strategic directions for remote sensing. *Rem. Sens. Environ.* 228, 164–182.




International Journal of Applied Earth Observation and Geoinformation 103 (2021) 102514

Zurita-Milla, R., Gijzel, J.A.E.v., Hamu, N.A.S., Augustijn, P.W.M., Vrieling, A., 2013. Exploring Spatiotemporal Phenological Patterns and Trajectories Using Self-Organizing Maps. *IEEE Trans Geosci Remote Sens.* 51, 1914–1921.

Appendix 2 : Decision-level integration and feature-level integration of remote sensing and geospatial big data for urban land use mapping

Article

Decision-Level and Feature-Level Integration of Remote Sensing and Geospatial Big Data for Urban Land Use Mapping

Jiadi Yin ^{1,2}, Ping Fu ^{1,*}, Nicholas A. S. Hamm ¹ , Zhichao Li ², Nanshan You ^{2,3}, Yingli He ^{2,3}, Ali Cheshmehzangi ⁴  and Jinwei Dong ² 

¹ School of Geographical Sciences and Geospatial and Geohazards Research Group, Faculty of Science and Engineering, University of Nottingham Ningbo China, Ningbo 315100, China; Jiadi.Yin@nottingham.edu.cn (J.Y.); Nicholas.Hamm@nottingham.edu.cn (N.A.S.H.)

² Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; lizc@igsnrr.ac.cn (Z.L.); youns.18b@igsnrr.ac.cn (N.Y.); heyl.19b@igsnrr.ac.cn (Y.H.); dongjw@igsnrr.ac.cn (J.D.)

³ University of Chinese Academy of Sciences, Beijing 100049, China

⁴ Department of Architecture and Built Environment, University of Nottingham Ningbo China, Ningbo 315100, China; Ali.Cheshmehzangi@nottingham.edu.cn

* Correspondence: Ping.Fu@nottingham.edu.cn



Citation: Yin, J.; Fu, P.; Hamm, N.A.S.; Li, Z.; You, N.; He, Y.; Cheshmehzangi, A.; Dong, J. Decision-Level and Feature-Level Integration of Remote Sensing and Geospatial Big Data for Urban Land Use Mapping. *Remote Sens.* **2021**, *13*, 1579. <https://doi.org/10.3390/rs13081579>

Academic Editor: Jochen Hack

Received: 23 February 2021

Accepted: 15 April 2021

Published: 19 April 2021

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



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Information about urban land use is important for urban planning and sustainable development. The emergence of geospatial big data (GBD), increased the availability of remotely sensed (RS) data and the development of new methods for data integration to provide new opportunities for mapping types of urban land use. However, the modes of RS and GBD integration are diverse due to the differences in data, study areas, classifiers, etc. In this context, this study aims to summarize the main methods of data integration and evaluate them via a case study of urban land use mapping in Hangzhou, China. We first categorized the RS and GBD integration methods into decision-level integration (DI) and feature-level integration (FI) and analyzed their main differences by reviewing the existing literature. The two methods were then applied for mapping urban land use types in Hangzhou city, based on urban parcels derived from the OpenStreetMap (OSM) road network, 10 m Sentinel-2A images, and points of interest (POI). The corresponding classification results were validated quantitatively and qualitatively using the same testing dataset. Finally, we illustrated the advantages and disadvantages of both approaches via bibliographic evidence and quantitative analysis. The results showed that: (1) The visual comparison indicates a generally better performance of DI-based classification than FI-based classification; (2) DI-based urban land use mapping is easy to implement, while FI-based land use mapping enables the mixture of features; (3) DI-based and FI-based methods can be used together to improve urban land use mapping, as they have different performances when classifying different types of land use. This study provides an improved understanding of urban land use mapping in terms of the RS and GBD integration strategy.

Keywords: urban land use; remote sensing; geospatial big data; decision-level integration; feature-level integration; Hangzhou

1. Introduction

China has undergone rapid urbanization since the early 1980s, shown as substantial urban expansion and dynamics in urban land use and structure [1]. The timely and accurate urban land use information is important for guiding urban planning and land use management [2]. Remote sensing (RS) techniques were widely used to update urban land use information over the past few decades, by referring to the differences in aspects of texture, spectrum, and context among urban land use categories [3,4]. However, due to the high similarity among urban land use categories in physical attributes, it is hard to identify

the complexity and diversity of urban internal structures [5,6], especially in cities with high-density populations and buildings, such as Hangzhou, Beijing, Shanghai, and Shenzhen [7]. Mixed-use land and shadows from high-rise buildings also pose great challenges for the RS data classification. Big data are defined partly by their large size and their characteristics, in particular their volume, variety, and velocity [8]. Much of this big data are geo-referenced, or can be geo-referenced, leading to geospatial big data (GBD). The emergence of GBD [9], such as mobile phone positioning data [10,11], points-of-interest (POI) [12], social media data [13–15], traffic trajectory data [16], and geotagged photographs [17–19], provides new opportunities to delineate human dimensions in an urban environment [20]. These multi-sourced data contain abundant human activity information, compensating for the lack of socioeconomic attributes of the RS data [21].

The integration of the space-based RS data and the time-based GBD can improve the existing urban land use maps by providing more detailed socioeconomic information and finer spatio-temporal resolution [22–24]. Promising progress was made on the applications of integrated RS and GBD on urban land use mapping at different scales and regions [25–29]. For example, Liu et al. [30] collected six features (spatial, texture, spatial envelope, rotation-invariant from RS images, Tencent real-time user density, and POI) to identify the urban land use type within each land use parcel. Jia et al. [31] classified RS data and mobile phone positioning data separately and then fused the two results by using a decision fusion strategy for categorizing urban land use. However, it is still challenging to integrate RS and GBD for urban land use mapping because of the modality gap (i.e., the differences in spatial data quality, technical format, and data structure) and heterogeneity in the data [32].

In this context, the main goal of this study is to propose the general framework by summarizing the RS and GBD integration approaches used in urban land use mapping and illustrate the advantages and disadvantages by applying them to map the urban land use situation in Hangzhou city, China. To our knowledge, this is the first study to summarize the existing efforts of integrating RS and GBD for urban land use mapping and analyzing their differences.

2. Related Work of RS and GBD Integration Used in Urban Land Use Mapping

Based on the literature review of previous studies, we categorize the various methods of RS and GBD integration used in urban land use mapping into decision-level integration (DI) and feature-level integration (FI) by considering the data integration modes (e.g., integration timing or integration method). A general framework is proposed, as shown in Figure 1, to demonstrate and differentiate the two approaches. Briefly, the main difference between DI-based and FI-based urban land use mapping is that DI combines the classification results of the RS and GBD, whereas FI integrates the features extracted from the RS and GBD for classification.

2.1. DI-Based Urban Land Use Mapping

The DI-based urban land use mapping processes the RS and GBD independently with different models and methods and combines the RS-based and GBD-based results for further generating the urban land use map. The DI-based method was first introduced by Chang et al. [33] and provides an efficient way for mapping urban land use in Kunming City by integrating POI data and Landsat images. Since then, a series of studies have followed that method. For example, Song et al. [34] further developed the method by fully exploring GBD (Weibo POI, Baidu POI) for more in-depth information and combining them with RS data (Landsat images, Gaofen). Tu et al. [35] fused landscape metrics (i.e., total class area, patch density, number of patches, and Shannon's diversity index) taken from RS and human activity metrics (i.e., the density of human activities in different functional zones) extracted from mobile phone positioning data for classifying urban functional zones by hierarchical clustering. Jia et al. [31] integrated RS-based urban land cover maps and GBD-based urban land use maps for mapping urban land use through decision fusion strategy (i.e., certain fusion rules for classifying different urban land use categories). Zhao

et al. [28] generated land cover types by training RS images with the semantic elements derived from OpenStreetMap (OSM) data, then identified each building through semantic classification by using POI. Xu et al. [36] extracted geographic information from RS and functional distribution from GBD (Gaode POI), then combined them by assigning different weights for urban land use mapping in China. In the DI-based method, the RS and GBD features are calculated and processed separately, avoiding the feature conflicting issues. Furthermore, it has the advantage of easy implementation. In summary, the RS and GBD features reflecting disparate attributes and dimensions are processed through independent methods, then fused by certain models (e.g., cross-correlation, decision fusion strategy).

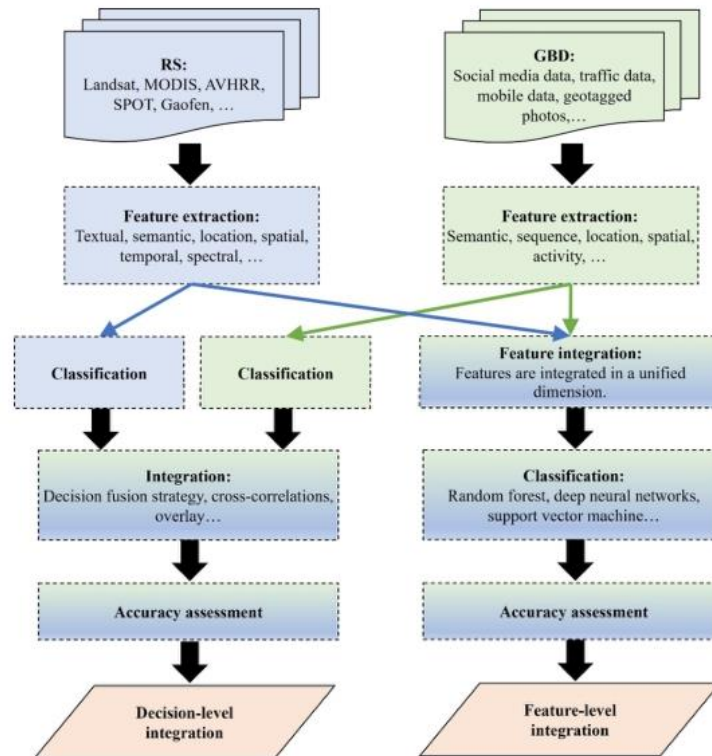


Figure 1. A general framework of the two integration strategies: Decision-level integration (DI), and Feature-level integration (FI).

2.2. FI-Based Urban Land Use Mapping

In the FI-based method, the features are first extracted from the RS and GBD and are then fused into the integrated feature sets and put into a classifier for urban land use classification. For example, Zhang et al. [37] extracted the features from the RS and GBD (i.e., spectral features, texture features, landscape metrics, density, spatial patterns of POI, and the density of geotagged Weibo posts), based on which urban land use types were determined and fused to train the classifiers for final mapping. Zhang et al. [27] evaluated the performances of the RS and GBD features and their cross-correlations for classifying urban functions by using a Random Forest (RF) model. Shi et al. [29] extracted

RS features from ZY-3, Landsat8, Sentinel-1A, and GBD (e.g., Wechat user density) for identifying urban land use. Cao et al. [32] proposed a novel framework that extracted the RS-based and GBD-based features (e.g., temporal features from user-visit data) for recognizing urban land use through deep learning. Zhang et al. [25] extracted the RS and GBD features to construct a sparse topic model for identifying representative zones with distinct patterns, then trained these zones to recognize urban functional zones using the deep forest model. Compared to the DI-based approach, the FI-based approach provides a more integrative way of processing the features derived from both RS data and GBD. The RS-based and GBD-based features were fully used for the mining of the urban land use information. However, the modality gap (i.e., differences in data quality, data formats, scales, and timestamps between RS data and GBD) between the RS and GBD features is still a challenging issue.

3. Case Study

3.1. Study Site

Based on the general framework, a detailed methodological framework was designed for testing the two approaches (DI and FI) by using Hangzhou city as a case study. The case study was carried out in Hangzhou city, the capital of Zhejiang Province in China, which is one of the representative cities of urban expansion, population growth, economic development, and land use changes [38,39]. We selected the area surrounded by the third ring (i.e., the Hangzhou belt highway) as our study site, which is located in the northeast of Hangzhou city (Figure 2c). Besides, the green spaces in the West Lake ecological zone and the croplands in the northeast region were removed to ensure that our study site had the main functional areas related to urban human activities, including institutions, residences, businesses, and open space.

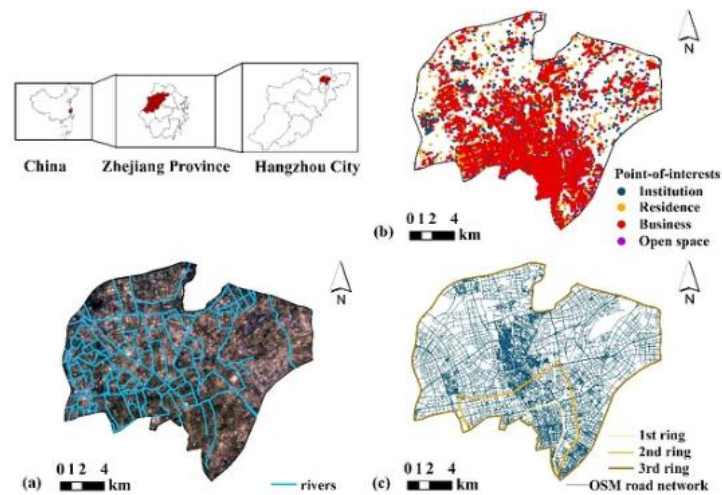


Figure 2. The study site (a) Sentinel-2 true-color composition image with OSM river network; (b) Spatial distribution of Point-of-interests (POI); (c) Road network extracted from OpenStreetMap (OSM).

3.2. Data Source and Preprocessing

Sentinel-2 is a high-resolution (10 m), multi-spectral, remote-sensing satellite [40,41]. In this paper, the data were integrated from 1 January to 31 December 2019 as mean composites after removing clouds using the Sentinel-2 Quality Assurance (QA) band [42].

OpenStreetMap (OSM) is a collaborative project providing free editable maps across the world (<https://www.openstreetmap.org>, accessed on 14 December 2019). The creation and growth of OSM data have mainly been promoted by users and GIS experts [43]. The high accuracy and free accessibility of OSM in urban areas have been confirmed by many researchers so far [44]. The weekly updated OSM data is very important for capturing the rapid development of urban areas. The OSM road network data are in vector format and contain different classes of roads with different road sizes [45,46]. It was proposed as a promising candidate for quick and robust delineations of urban structures and socioeconomic patterns since it is gradually being used to understand Chinese cities [47]. Thus, it is reliable to generate urban land use parcels using the OSM road network data.

As a prominent example of the GBD that attracts the most users, POI contain abundant information including land use category, geographic location, and other features (e.g., address, telephone, and postcode) [36]. A total number of 28,898 available records of POI within Hangzhou city in 2019 was collected via application programming interfaces (APIs) provided by Gaode Map Services. The urban land use classification system used for aggregating initial POI types were shown in Table 1 (<https://lbs.amap.com>, accessed on 14 December 2019). Moreover, according to previous research, the urban area could be divided into six categories, including institution, residence, commercial land, industrial land, open space, and transportation [25,36]. Since we have excluded the road and river layers for generating urban parcels, there are five categories left. We further combined the commercial and industrial land into business as both have the nature of the business, and the business level meets the classification requirements of this paper. Therefore, the urban classification system used in this study consists of four categories: institution, residence, business, and open space.

Table 1. The framework used for aggregating initial Gaode POI types.

Gaode POI Classification	Urban Land Use Classification
Governmental organization Medical service	Institution
Finance and Insurance service	
Sports and Recreation	
Culture and Education	
Daily life service Commercial house	Residence
Commercial service Shopping Food and Beverages Enterprises Accommodation service	Business
Tourist attraction	Open space

3.3. Methods

The case study was carried out with the assumption that a parcel divided by urban road networks is homogeneous in terms of urban land use function [44,48]. Quantitatively illustrating the differences between DI-based and FI-based methods requires the following steps: (1) extracting urban land parcels based on the road network and river data (Figure 3a) and preparing training and validation parcels (Figure 3b); (2) generating urban land use maps by using DI-based (Figure 3c) and FI-based methods, respectively, (Figure 3d) and evaluating the two based on validation datasets; (3) analyzing the differences between the

two methods in terms of quantitative and qualitative assessment of resulting maps and their data processing modes. The detailed procedures are presented below.

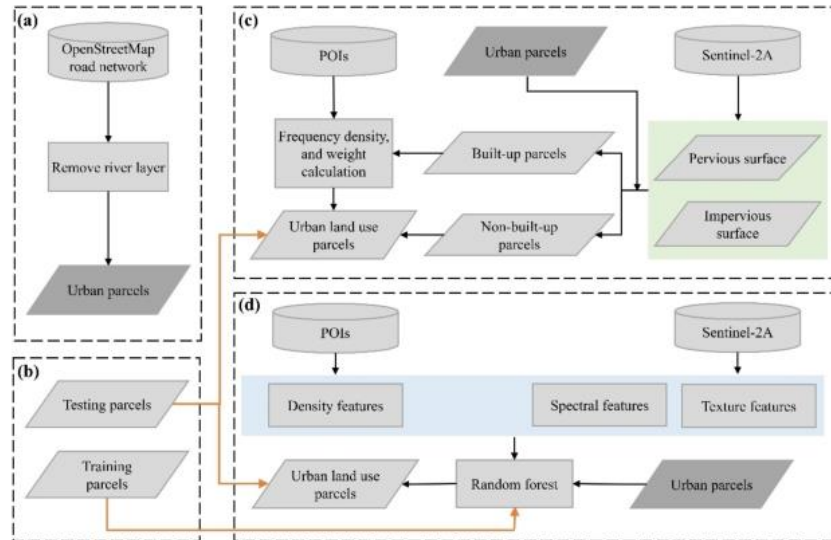


Figure 3. This is a figure. Schemes follow the same formatting. Research methodological framework of mapping urban land use with decision-level integration (DI) and feature-level integration (FI). (a) Urban parcel generation based on the road network and river data; (b) Collection of testing and training parcels; (c) DI-based urban land use mapping (DI); (d) FI-based urban land use mapping.

3.3.1. Urban Parcel Generation

The study area was segmented into urban parcels based on the method of the automated identification and characterization of parcels (AICP) [47]. Specifically, roads less than 100 m and roads inside communities were considered as redundant information and manually removed. The road network and river layer were then used for generating road buffers and river buffers. The buffer widths of the road and river were determined based on their respective classes according to the Ministry of Housing and Urban-Rural Development (MoHURD) [49] and the actual situation in Hangzhou city (Table 2). The river layer was delineated by visual interpretation based on Google Earth, and the preprocessing of the road network and buffer generation were conducted by using ArcMap 10.2.

Table 2. The road and river levels and their corresponding buffer sizes used in this study.

Classes	Road Descriptions	Road Widths (m)	River Descriptions	River Widths (m)
Level 1	Trunk, primary, motorway, railway	40	Main rivers	50
Level 2	Secondary	20	Intermediate rivers	20
Level 3	Tertiary, unclassified, residential, service, others	10	Small rivers	10

3.3.2. Training and Testing Parcel Collection

We selected parcels for preparing the training and testing parcels, of which the land use types were identified by using the Baidu map, Baidu street view, and a field survey (Figures 4 and 5). The hybrid satellite map of Baidu map presents a top view of the urban landscape with high-resolution satellite imagery and names of ground features, which were further inspected or confirmed using Baidu street view as it contains high-resolution photos and allows multiple perspectives [50,51]. An example was presented in Figure 4 to illustrate the top view of an institution and residence in the Baidu map and their corresponding Baidu street views. We determined the parcel type while different types of street views were observed for the same parcel, and there is no predominant characteristic referring to the rules in essential urban land use categories in China (EULUC-China) [52]. To be more specific, we recorded a list of items, including the geolocations, classification types, landmark buildings and facilities, and mixed land use situation and their estimated proportions for each urban parcel. Finally, a total of 700 training parcels were identified, including 175 institution parcels, 175 residence parcels, 175 business parcels, and 175 open space parcels. These parcels were selected using stratified random sampling, with an equal sample size (175) for each land use class. Following the approach discussed by Stehmen and Foody [53], we selected 550 parcels for testing. These 550 parcels were selected using simple random sampling from the list of parcels. For each parcel, the land use class was determined by visual interpretation based on the Baidu map. The 550 parcels included 85 institution parcels, 225 residence parcels, 156 business parcels, and 84 open space parcels. The training parcels were used for training the RF classifiers in FI-based classification, while testing parcels were used for accuracy assessment in both DI-based and FI-based classification, respectively.

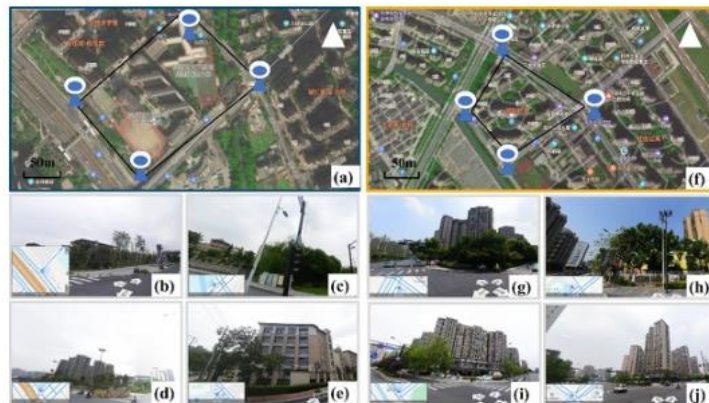


Figure 4. Illustration of the institution and residence parcels in Baidu map and Baidu street view: (a) represents the institution parcel in Baidu map; (b–e) represent the four perspectives of institution parcel in Baidu street view; (f) represents the residence parcel in Baidu map; and (g–j) represent the four perspectives of residence parcel in Baidu street view. The Chinese labels for (a,f) represent the name of POIs.

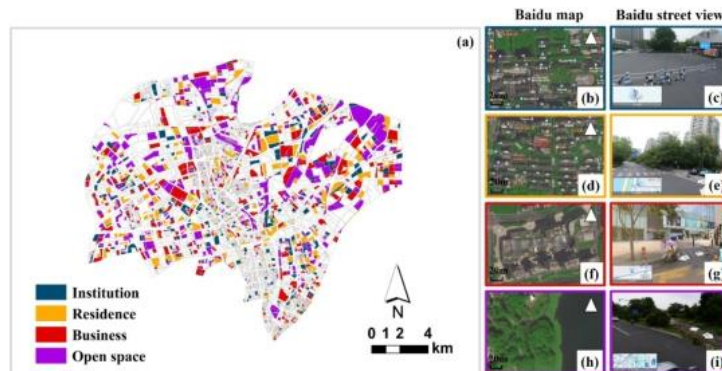


Figure 5. Spatial distribution and illustration of training and testing parcels for the four urban land use types. (a) represents the spatial distribution of training and testing parcels; (b,d,f,h) represents the Baidu map image for institution, residence, business, and open space, respectively; and (c,e,g,i) represent the Baidu street view for the institution, residence, business, and open space, respectively. The Chinese labels for (b,d,f,h) represent the name of POIs.

3.3.3. DI-Based Urban Land Use Mapping

The DI-based urban land use mapping integrated the results of the RS and GBD classification based on integration rules (Figure 3c). First, the Sentinel-2 image was classified for producing the impervious and pervious classes (Figure 6a). Specifically, the training samples identified by visual interpretation based on Google Earth and a field survey were put into the RF classifier for the classification of the impervious and pervious surface. The impervious surface is characterized by the artificial structures that are covered by water-resistant materials, while the pervious surface is the opposite [54]. The classification result was then overlaid with the urban parcels to generate the built-up and non-built-up parcels (Figure 6b). In this study, urban parcels with impervious surfaces were defined as built-up, otherwise, the parcels were labeled as non-built-up. This step was carried out under the assumption that built-up parcels require at least partially impervious surfaces [33]. The non-built-up parcels were further classified as open space, while built-up parcels were used for further analysis.

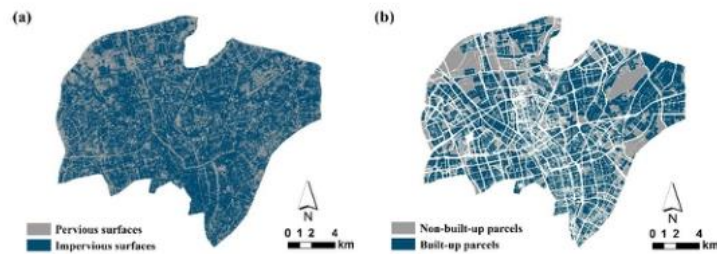


Figure 6. (a) Distribution of the pervious and impervious classes; (b) Distribution of the built-up and non-built-up parcels.

Finally, the type of each built-up parcel was determined according to a modified method [55]. The method consists of two parts: the frequency density calculation, and the weight calculation. In general, the main type in each parcel was determined after comparing the final frequency density ratios of four types of the POI, and the highest ratio of POI is considered as the final type. The final frequency density ratios were determined

by multiplying the initial frequency density ratio and the weight of each types' POI. Specifically, the initial frequency density ratios of the four types' POI were calculated by using Equations (1)–(3).

$$F_i = \frac{n_i}{N_i}, i = 1, 2, 3, 4 \quad (1)$$

where i represents the type of POI, which are institution, residence, business, or open space. n_i demonstrates the quantity of type i in a parcel. N_i represents the total quantity of type i in all parcels. F_i demonstrates the frequency density which equals the ratio of n_i to N_i .

$$C_i = \frac{F_i}{\sum_1^4 F_i} \times 100\%, i = 1, 2, 3, 4 \quad (2)$$

where C_i represents the initial ratio of type i 's frequency density to the sum of all four types' frequency density.

The initial frequency density ratios of the four types' POI were then respectively multiplied with their weights since various types of POI have different recognition degrees. The weight of institution, residence, business, and open space POI was defined as 2, 1, 1.5, and 2.5, referring to the construction area and the actual situation. The type in each parcel was determined as the one with the highest ratio and the final urban land use map was obtained. The open space parcels produced by the calculation of POI were further combined with the non-built-up parcels as the final open space parcels.

$$D_i = C_i \times W_i \times 100\%, i = 1, 2, 3, 4 \quad (3)$$

where D_i represents the final frequency density ratio which equals the initial ratio of type i 's frequency density multiplying the weight of each types' POI. $W_1 = 2$, $W_2 = 1$, $W_3 = 1.5$, $W_4 = 2.5$.

3.3.4. FI-Based Urban Land Use Mapping

The FI-based classification used 18 features derived from the RS and GBD, including eight spectral features, four textural features, and six density features (Table 3). Textural features indicate the adjacency relationships between the gray levels of pixels to characterize the texture of an image [37,56]. In this study, the textural features (i.e., angular second moment, contrast, dissimilarity, and entropy) were extracted by computing Gray Level Co-Occurrence Matrix (GLCM) that contain rich information on spatial structure and landscapes [57]. The near-infrared (NIR) band from the Sentinel image was used for computing the spectral and textural features, as it contains a negative correlation between built-up and vegetation [58–60]. Moreover, density features from POI help to present the differences in spatial patterns among different types of parcels. The Kernel Density Estimation (KDE) tool in ArcMap 10.2 was utilized to generate four layers of kernel density of the four land use types of POI, respectively. In FI-based classification, the RF model was trained using all the above-mentioned features 100 times to attenuate the uncertainties and get more stable accuracies. All of the operations in the FI-based classification were realized using the Google Earth Engine (GEE) platform.

Table 3. Spectral and textural features from RS, and density features from GBD used in FI-based classification.

Feature Types	Indices
Spectral features	Enhanced Vegetation Index (EVI), Normal Difference Built-up Index (NDBI), Normal Difference Vegetation Index (NDVI), Normal Difference Water Index (NDWI), mean, standard deviation, kurtosis, skewness
Textural features	Angular second moment, contrast, dissimilarity, and entropy
Density features	Minimum, maximum, range, sum, mean and standard deviation

3.3.5. Analysis of the DI-Based and FI-Based Classification

To understand the differences between DI-based and FI-based urban land use mapping and describe their advantages and disadvantages, we first evaluated the results derived from the two methods by computing the overall accuracy (OA), user’s accuracy (UA), and producer’s accuracy (PA) based on the confusion matrix [61]. Approximately 95% confidence intervals of OA, UA, and PA were also calculated based on the equations proposed by Card [62] (Supplementary Materials). In the estimation procedure for DI and FI methods, we used a proportion of the number of parcels in each category, relative to all categories in the DI and FI classification results, and denoted this as w_j , following Olofsson, et al. [63], and Stehmen and Foody [53]. The column refers to the classified map category while the row refers to the true category based on the reference data (see the Supplementary Materials). For DI, No Data parcels refer to the parcels without POI values that were not involved in the confusion matrix. We then implemented the visual analysis of the DI-based and FI-based urban land use maps. Furthermore, we evaluated the implementation of the two methods by discussing their advantages and disadvantages.

4. Results

4.1. Quantitative Performance of DI-Based and FI-Based Urban Land Use Mapping

Table 4 showed the confusion matrix of the pervious and impervious map. The estimated OA ($\pm 95\%$ confidence interval) is 0.971 ± 0.007 . The estimated UA and PA of the impervious surface are 0.975 ± 0.008 and 0.986 ± 0.005 . Table 5 presents the confusion matrix of DI-based classification. The estimated OA for DI-based urban land use map is 0.635 ± 0.049 . Note that the residence and open space were classified relatively well, with the estimated UA of 0.728 ± 0.069 and 0.714 ± 0.140 , respectively. For institution and business, they have relatively lower UA (0.512 ± 0.140 and 0.522 ± 0.095). Table 6 presents the confusion matrix of FI-based classification. The estimated OA for FI-based classification results is 0.569 ± 0.041 (Table 6). It can be noted that the residence and open space in FI-based classification can be classified relatively well, with the UA of 0.631 ± 0.059 and 0.640 ± 0.111 , respectively.

Table 4. Confusion matrix of the impervious and pervious surface extraction results. UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map.

		Map Category					
		Class	Pervious	Impervious	Total	PA	OA
True Category	Pervious		0.240	0.018	0.258	0.928 ± 0.022	0.971 ± 0.007
	Impervious		0.010	0.732	0.742	0.986 ± 0.005	
	Total (π_j)		0.250	0.750	1.000		
	UA		0.958 ± 0.018	0.975 ± 0.008			

Table 5. Confusion matrix of DI-based classification results (I: Institution; R: Residence; B: Business; O: Open Space). UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map.

		Map Category							
		Class	I	R	B	O	Total	PA	OA
True Category	I		0.071	0.029	0.042	0.014	0.156	0.453 ± 0.107	0.635 ± 0.049
	R		0.015	0.326	0.072	0.008	0.421	0.774 ± 0.055	
	B		0.034	0.065	0.157	0.010	0.265	0.591 ± 0.082	
	O		0.019	0.029	0.029	0.081	0.158	0.514 ± 0.102	
	Total (π_j)		0.138	0.448	0.300	0.113	1.000		
	UA		0.512 ± 0.140	0.728 ± 0.069	0.522 ± 0.095	0.714 ± 0.140			

Table 6. Confusion matrix of FI-based classification results (I: Institution; R: Residence; B: Business; O: Open Space). UA: users accuracy; PA: producers accuracy; OA: overall accuracy; π_j is the class proportion according to the classified map.

Class		Map Category				Total	PA	OA
		I	R	B	O			
True Category	I	0.071	0.030	0.025	0.016	0.142	0.496 ± 0.092	0.569 ± 0.041
	R	0.020	0.291	0.108	0.019	0.439	0.663 ± 0.045	
	B	0.036	0.112	0.124	0.012	0.284	0.436 ± 0.062	
	O	0.008	0.028	0.015	0.084	0.135	0.620 ± 0.094	
	Total (π_j)	0.135	0.462	0.272	0.130	1.000		
UA		0.523 ± 0.114	0.631 ± 0.059	0.453 ± 0.080	0.640 ± 0.111			

4.2. Qualitative Performance of DI-Based and FI-Based Urban Land Use Mapping

Figure 7 presents the urban land use maps derived from DI-based and FI-based methods in Hangzhou city. Overall, there is a difference in the number of each urban land use type, and the spatial distribution of the four land use types is not similar. In the DI-based map, the residence parcels were mainly distributed in the city center, while the open space parcels were distributed around Hangzhou city. The distribution of urban land uses in the DI-based map was visually consistent with the actual urban land use referring to the Baidu map. In the FI-based resulting map, the distribution of each land use type was more widely spread out across the study area than that in the DI resulting map. No-value parcels only exist in the resultant map from the DI-based classification, where the built-up parcels determined from the RS images are further labeled based on POI (see Section 3.3.3). No-value parcels arise where there is no POI data within the built-up parcel to support this labeling. On the contrary, the FI-based classification integrates features from both RS and GBD before land use classification could be classified, as each parcel contains features at least from RS data.

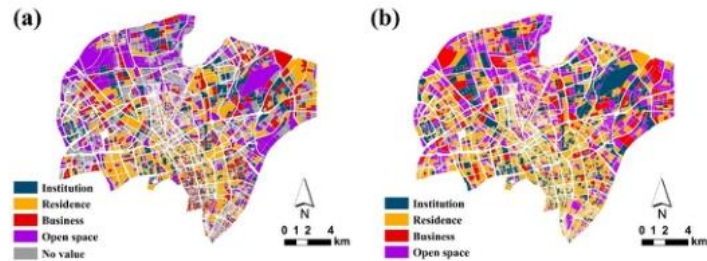


Figure 7. Illustration of the (a) DI-based and (b) FI-based urban land use maps in Hangzhou city.

Figure 8 presents a visual comparison of the urban land use maps derived from DI-based and FI-based methods and the corresponding Baidu map in four subareas. The four subareas A, B, C, and D are dominated by institution, residence, business and opens space, respectively. Overall, the visual comparison indicates a generally better performance of the DI-based classification than the FI-based classification. For example, some institution parcels were classified as residence for the FI-based classification in subarea A. In subarea B, some residence parcels were misclassified as business by the FI-based classification. The FI-based result map misclassified business parcels as residence parcels in subarea C. For subarea D, the open space parcel was wrongly classified as a residence parcel.

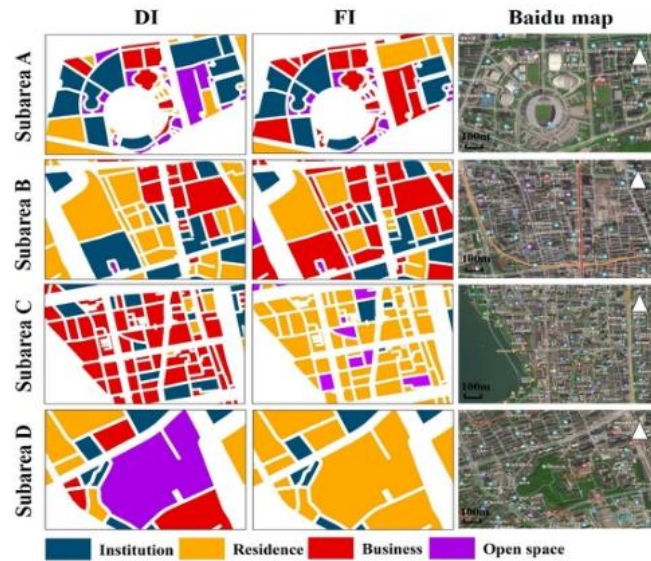


Figure 8. Classification results of DI and FI methods and testing samples in the test regions. Subareas A, B, C, and D were selected randomly in Hangzhou city, and each region was shown with DI-based classification, FI-based classification, and Baidu map. The Chinese labels for the subfigure of right column represent the name of POIs.

5. Discussion

This study implemented the DI-based and FI-based methods for mapping urban land use types in Hangzhou city, based on the same urban land use classification system (Table 1), the same data sources (i.e., Sentinel-2A, OSM road networks, and POI), and the same training and validation urban parcels (Figure 5). In our study, the only difference between these two approaches is the manner and timing for integrating the RS features and GBD features (Figures 1 and 3). It should be noted that the objective of this study is not to quantitatively compare the performance of the two methods but to illustrate their advantages and disadvantages by applying them to the mapping of the urban land use situation in Hangzhou city. The results in this study, which show that DI-based classification performs better than FI-based classification, may not apply to other circumstances [36,64]. Given that our study area is located in eastern China, where the diversity of urban land use types has a higher level of complexity than other regions in China, our mapping result is probably not representative of other regions. It should be noted that different data sources, selected features, classifiers, training samples, or land use types may also lead to different results. According to the process of the integration methods, these differences can be summarized into four types: (1) different urban road networks in different regions might lead to different urban parcel unit; (2) The variety of GBD in different regions lead to different mapping results; (3) The number of mixed parcels in different regions varies depending on the tiers of cities; (4) The availability of RS data greatly varies in different cities due to the coverage of clouds and cloud shadows that might be another factor affecting the application of the proposed methods.

5.1. Summary of the Advantages and Disadvantages of the Two Methods

In our study, for the DI-based method, the pervious and impervious surfaces map was overlaid with urban parcels for generating built-up and non-built-up parcels, and then they were classified into the four urban land use types by calculating the frequency density and weight of the four types of POI. For the FI-based classification, the RS and GBD features were combined to train the RF model for urban land use mapping based on the training parcels. Some advantages and disadvantages could be identified through the analysis of classification results and the above-mentioned process:

- DI-based urban land use mapping is easy to implement, avoids feature integration, and accompanies conflicting issues. However, it depends largely on the quality and quantity of the GBD in each urban parcel (it depended on POI data in our case study) which might cause the missing value and misclassification for some urban parcels (Figures 7 and 8). For example, Zhao et al. [28] indicated that inaccurate POI will produce incorrect labels, as the classification results are directly generated from POI. Furthermore, the classes used in this paper might not match well with the POI classes considering that DI-based classification is based on labeling parcels based on POI. Reclassifying POI classes according to the nomenclature of land use types can result in some uncertainties of urban land use classification. For example, certain POI could be associated with more than one type of urban land use;
- The FI-based land use mapping enables the mixture of features from RS and GBD, however, the implementation has challenges due to the modality gap [20,32] between the RS and GBD, such as the spatial data quality, technical format, and data structure. Moreover, both feature selection and feature integration in the FI-based classification can contribute to different mapping results. In this study, spectral features and textural features derived from the Sentinel-2 image were integrated with the density features derived from POI for mapping urban land use (Table 3). Accordingly, these features have multiple backgrounds and thus can have various understandings of urban land use mapping, leading to different classification results [65]. The performance of the FI-based classification is probably related to the complexity of urban parcels. To be more specific, a single urban parcel can comprise different urban land uses, such as office buildings, residential buildings, and shopping centers. Therefore, the differences in features among urban parcels might be hard to distinguish.

5.2. The Improved Method for Urban Land Use Mapping

Given the framework (Figure 1), quantitative and qualitative assessments (Tables 5 and 6, Figures 7 and 8) of the two methods, it is possible to highlight some points for improving the DI-based and FI-based urban land use mapping. First, very-high-resolution (VHR) satellite images provide abundant information about geographical objects in terms of geometrical features and spatial patterns [28], which can provide more detailed RS information (e.g., textual, contextual, and spectral information) [37]. Second, the generation of urban parcels relied on the OSM road network, which contains unbalanced classes that are unevenly distributed, leading to heterogeneous urban land use parcels in terms of parcel area and spatial pattern [66]. The accuracy of the urban parcels created based on OSM can be further improved by combining different data sources, and the veracity of urban parcels should be validated. Third, different GBD features have different contributions to the classification of different land use types according to Zhang et al. [37]. Evaluating the optimal size of the GBD set is a topic for future research. Fourth, the training and testing parcels used in this study were identified through the Baidu map, Baidu street view, and a field survey, which are promising to validate urban land use parcels [67]. However, it is inevitable for parcels to contain multiple land use types. A better classification performance would be expected from the adoption of an independent land use map from the local government and urban planners [68]. Fifth, the RF model was used for mapping built-up in DI-based classification and training urban land use parcels in the FI-based classification. However, we need to improve the RF model with transfer learning capability for mapping

the urban land use of any region with high accuracy and a different data source. To be more specific, the transfer learning model can be used as a pre-trained model, which might help increase the classification accuracy. Furthermore, how to integrate GBD into the RF model is a worthwhile question to discuss, since there is a significant difference between RS and GBD [39]. Sixth, the two methods can be used together to improve urban land use maps as they have different performances in classifying different types of land use. For example, the FI-based method can be used for classifying very different urban land use types, such as open space land and industrial land, while the DI-based method can be utilized for classifying urban land use types with very little difference in the RS features but have a significant difference in GBD features, such as commercial land and industrial land.

6. Conclusions

With the increase in RS and GBD, integrating the two types of data was widely used in urban land use mapping in recent years as it provides an opportunity to characterize both physical and socioeconomic attributes of urban land. For the first time, this study summarized the diverse methods of RS and GBD integration used in essential urban land use mapping and categorized them into two types, including the DI-based and FI-based methods. A general framework was proposed for explaining the main differences in the process of the two methods. Taking Hangzhou city, China as a case study, we then applied the two methods to produce the urban land use maps consisting of the institution, residence, business, and open space, based on the OSM road network, Sentinel-2 data, and Gaode POI, and highlighted the differences in classification results and the process of the DI-based and FI-based methods. On this basis, future studies should pay more attention to the integration methods, and further apply the methods to more specific classification-type scenarios, data sources, and other regions.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/rs13081579/s1>, Table S1: Contingency table for accuracy assessment, Table S2: Contingency table for accuracy assessment in terms of proportion of numbers, Table S3: Contingency table for evaluating map accuracy of DI-based classification results (I: Institution; R: Residence; B: Business; O: Open Space), Table S4: Contingency table for evaluating map accuracy of FI-based classification results (I: Institution; R: Residence; B: Business; O: Open Space).

Author Contributions: J.Y. and N.A.S.H. conducted the data analysis and wrote the manuscript; J.Y., N.Y., and Y.H. developed the methodology; Z.L., P.F., N.A.S.H., A.C., and J.D. provided valuable insights and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (Grant nos. 41971078, 41871349, 41801336). The Chinese Academy of Sciences the Strategic Priority Research Program (XDA19040301), the Key Research Program of Frontier Sciences (QYZDB-SSW-DQC005).

Acknowledgments: J.Y. acknowledges a Ph.D. scholarship provided by the University of Nottingham Ningbo China and the Institute for Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences.

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

References

1. Gong, P.; Li, X.; Zhang, W. 40-Year (1978–2017) human settlement changes in China reflected by impervious surfaces from satellite remote sensing. *Sci. Bull.* **2019**, *64*, 756–763. [CrossRef]
2. Zhu, Z.; Zhou, Y.; Seto, K.C.; Stokes, E.C.; Deng, C.; Pickett, S.T.A.; Taubenböck, H. Understanding an urbanizing planet: Strategic directions for remote sensing. *Remote Sens. Environ.* **2019**, *228*, 164–182. [CrossRef]
3. Liu, X.; Hu, G.; Chen, Y.; Li, X.; Xu, X.; Li, S.; Pei, F.; Wang, S. High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform. *Remote Sens. Environ.* **2018**, *209*, 227–239. [CrossRef]
4. Reba, M.; Seto, K.C. A systematic review and assessment of algorithms to detect, characterize, and monitor urban land change. *Remote Sens. Environ.* **2020**, *242*, 111739. [CrossRef]

5. Qi, L.; Li, J.; Wang, Y.; Gao, X. Urban observation: Integration of remote sensing and social media data. *IEEE J. Select. Top. Appl. Earth Observat. Remote Sens.* **2019**, 1–13. [\[CrossRef\]](#)
6. Li, J.; He, Z.; Plaza, J.; Li, S.; Chen, J.; Wu, H.; Wang, Y.; Liu, Y. Social media: New perspectives to improve remote sensing for emergency response. *Proc. IEEE* **2017**, 105, 1900–1912. [\[CrossRef\]](#)
7. Cai, J.; Huang, B.; Song, Y. Using multi-source geospatial big data to identify the structure of polycentric cities. *Remote Sens. Environ.* **2017**, 202, 210–221. [\[CrossRef\]](#)
8. Kitchin, R.; McArdle, G. What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data Soc.* **2016**, 3, 205395171663113. [\[CrossRef\]](#)
9. Li, S.; Dragicevic, S.; Castro, F.A.; Sester, M.; Winter, S.; Coltekin, A.; Pettit, C.; Jiang, B.; Haworth, J.; Stein, A.; et al. Geospatial big data handling theory and methods: A review and research challenges. *ISPRS J. Photogramm. Remote Sens.* **2016**, 115, 119–133. [\[CrossRef\]](#)
10. Ratti, C.; Frenchman, D.; Pulselli, R.M.; Williams, S. Mobile landscapes: Using location data from cell phones for urban analysis. *Environ. Plann. B Plann. Des.* **2016**, 33, 727–748. [\[CrossRef\]](#)
11. Wu, L.; Cheng, X.; Kang, C.; Zhu, D.; Huang, Z.; Liu, Y. A framework for mixed-use decomposition based on temporal activity signatures extracted from big geo-data. *Int. J. Digit. Earth* **2020**, 13, 708–726. [\[CrossRef\]](#)
12. Andrade, R.; Alves, A.; Bento, C. POI mining for land use classification: A case study. *ISPRS Int. J. Geo-Inf.* **2020**, 9, 493. [\[CrossRef\]](#)
13. Ilieva, R.T.; McPhearson, T. Social-media data for urban sustainability. *Nat. Sustain.* **2018**, 1, 553–565. [\[CrossRef\]](#)
14. Yammine, S.Z.; Jarreau, P.B.; Liu, C.; Coe, I.R. Social media for social change in science. *Science* **2018**, 360, 163.
15. Ye, X.; Huang, Q.; Li, W. Integrating big social data, computing and modeling for spatial social science. *Cartogr. Geogr. Inf. Sci.* **2016**, 43, 377–378. [\[CrossRef\]](#)
16. Niu, N.; Liu, X.; Jin, H.; Ye, X.; Liu, Y.; Li, X.; Chen, Y.; Li, S. Integrating multi-source big data to infer building functions. *Int. J. Geogr. Inf. Sci.* **2017**, 31, 1871–1890. [\[CrossRef\]](#)
17. Srivastava, S.; Vargas Muñoz, J.E.; Lobry, S.; Tuia, D. Fine-grained landuse characterization using ground-based pictures: A deep learning solution based on globally available data. *Int. J. Geogr. Inf. Sci.* **2018**, 34, 1117–1136. [\[CrossRef\]](#)
18. Krylov, V.; Kenny, E.; Dahyot, R. Automatic discovery and geotagging of objects from street view imagery. *Remote Sens.* **2018**, 10, 661. [\[CrossRef\]](#)
19. Cadavid Restrepo, A.M.; Yang, Y.R.; Hamm, N.A.S.; Gray, D.J.; Barnes, T.S.; Williams, G.M.; Soares Magalhaes, R.J.; McManus, D.P.; Guo, D.; Clements, A.C.A. Land cover change during a period of extensive landscape restoration in Ningxia Hui Autonomous Region, China. *Sci. Total Environ.* **2017**, 598, 669–679. [\[CrossRef\]](#)
20. Liu, Y.; Liu, X.; Gao, S.; Gong, L.; Kang, C.; Zhi, Y.; Chi, G.; Shi, L. Social sensing: A new approach to understanding our socioeconomic environments. *Ann. Assoc. Am. Geogr.* **2015**, 105, 512–530. [\[CrossRef\]](#)
21. Li, J.; Benediktsson, J.A.; Zhang, B.; Yang, T.; Plaza, A. Spatial technology and social media in remote sensing: Challenges and opportunities. *Proc. IEEE* **2017**, 105, 1583–1585. [\[CrossRef\]](#)
22. Sarmin, N.S.; Ismail, M.H. A review of potentialities and challenges of integrating remote sensing and GIS with socioeconomic data. *PJSRR* **2016**, 2, 129–141.
23. Dong, J.; Metternicht, G.; Hostert, P.; Fensholt, R.; Chowdhury, R.R. Remote sensing and geospatial technologies in support of a normative land system science: Status and prospects. *Curr. Opin. Environ. Sustain.* **2019**, 38, 44–52. [\[CrossRef\]](#)
24. Xiong, G.; Cao, X.; Hamm, N.A.S.; Lin, T.; Zhang, G.; Chen, B. Unbalanced development characteristics and driving mechanisms of regional urban spatial form: A case study of Jiangsu Province, China. *Sustainability* **2021**, 13, 3121. [\[CrossRef\]](#)
25. Zhang, X.; Du, S.; Zheng, Z. Heuristic sample learning for complex urban scenes: Application to urban functional-zone mapping with VHR images and POI data. *ISPRS J. Photogramm. Remote Sens.* **2020**, 161, 1–12. [\[CrossRef\]](#)
26. Goffi, A.; Bordogna, G.; Stroppiana, D.; Boschetti, M.; Brivio, P.A. Knowledge and data-driven mapping of environmental status indicators from remote sensing and VGI. *Remote Sens.* **2020**, 12, 495. [\[CrossRef\]](#)
27. Zhang, Y.; Li, Q.; Tu, W.; Mai, K.; Yao, Y.; Chen, Y. Functional urban land use recognition integrating multi-source geospatial data and cross-correlations. *Comput. Environ. Urban Syst.* **2019**, 78, 101374. [\[CrossRef\]](#)
28. Zhao, W.; Bo, Y.; Chen, J.; Tiede, D.; Blaschke, T.; Emery, W.J. Exploring semantic elements for urban scene recognition: Deep integration of high-resolution imagery and OpenStreetMap (OSM). *ISPRS J. Photogramm. Remote Sens.* **2019**, 151, 237–250. [\[CrossRef\]](#)
29. Shi, Y.; Qi, Z.; Liu, X.; Niu, N.; Zhang, H. Urban land use and land cover classification using multisource remote sensing images and social media data. *Remote Sens.* **2019**, 11, 2719. [\[CrossRef\]](#)
30. Liu, X.; He, J.; Yao, Y.; Zhang, J.; Liang, H.; Wang, H.; Hong, Y. Classifying urban land use by integrating remote sensing and social media data. *Int. J. Geogr. Inf. Sci.* **2017**, 31, 1675–1696. [\[CrossRef\]](#)
31. Jia, Y.; Ge, Y.; Ling, F.; Guo, X.; Wang, J.; Wang, L.; Chen, Y.; Li, X. Urban land use mapping by combining remote sensing imagery and mobile phone positioning data. *Remote Sens.* **2018**, 10, 446. [\[CrossRef\]](#)
32. Cao, R.; Tu, W.; Yang, C.; Li, Q.; Qiu, G. Deep learning-based remote and social sensing data fusion for urban region function recognition. *ISPRS J. Photogramm. Remote Sens.* **2020**, 163, 82–97. [\[CrossRef\]](#)
33. Chang, C.; Ye, Z.; Huang, Q.; Wang, C. An integrative method for mapping urban land use change using “geo-sensor” data. In Proceedings of the 1st International ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics, UrbanGIS’15, Bellevue, WA, USA, 3–6 November 2015; pp. 47–54.

34. Song, J.; Lin, T.; Li, X.; Prishchepov, A.V. Mapping urban functional zones by integrating very high spatial resolution remote sensing imagery and points of interest: A case study of Xiamen, China. *Remote Sens.* **2018**, *10*, 1737. [[CrossRef](#)]
35. Tu, W.; Hu, Z.; Li, L.; Cao, J.; Jiang, J.; Li, Q.; Li, Q. Portraying urban functional zones by coupling remote sensing imagery and human sensing data. *Remote Sens.* **2018**, *10*, 141. [[CrossRef](#)]
36. Xu, S.; Qing, L.; Han, L.; Liu, M.; Peng, Y.; Shen, L. A new remote sensing images and point-of-interest fused (RPF) model for sensing urban functional regions. *Remote Sens.* **2020**, *12*, 1032. [[CrossRef](#)]
37. Zhang, Y.; Li, Q.; Huang, H.; Wu, W.; Du, X.; Wang, H. The combined use of remote sensing and social sensing data in fine-grained urban land use mapping: A case study in Beijing, China. *Remote Sens.* **2017**, *9*, 865. [[CrossRef](#)]
38. Li, D.; Lu, D.; Wu, M.; Shao, X.; Wei, J. Examining land cover and greenness dynamics in Hangzhou bay in 1985–2016 using Landsat time-series data. *Remote Sens.* **2017**, *10*, 32. [[CrossRef](#)]
39. Mao, W.; Lu, D.; Hou, L.; Liu, X.; Yue, W. Comparison of machine-learning methods for urban land-use mapping in Hangzhou City, China. *Remote Sens.* **2020**, *12*, 2817. [[CrossRef](#)]
40. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [[CrossRef](#)]
41. Zhu, Z.; Wang, S.; Woodcock, C.E. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sens. Environ.* **2015**, *159*, 269–277. [[CrossRef](#)]
42. Adiri, Z.; Lhissou, R.; El Harti, A.; Jellouli, A.; Chakouri, M. Recent advances in the use of public domain satellite imagery for mineral exploration: A review of Landsat-8 and Sentinel-2 applications. *Ore Geol. Rev.* **2020**, *117*, 103332. [[CrossRef](#)]
43. Weiss, D.J.; Nelson, A.; Gibson, H.S.; Temperley, W.; Peedell, S.; Lieber, A.; Hancher, M.; Poyart, E.; Belchior, S.; Fullman, N.; et al. A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **2018**, *553*, 333–336. [[CrossRef](#)] [[PubMed](#)]
44. Helbich, M.; Amelunxen, C.; Neis, P.; Zipf, A. Comparative spatial analysis of positional accuracy of OpenStreetMap and proprietary Geodata. In Proceedings of the GI_Forum 2012: Geovisualization, Society and Learning, Salzburg, Austria, 3–6 July 2012; pp. 4–6.
45. Liu, D.; Chen, N.; Zhang, X.; Wang, C.; Du, W. Annual large-scale urban land mapping based on Landsat time series in Google Earth Engine and OpenStreetMap data: A case study in the middle Yangtze River basin. *ISPRS J. Photogramm. Remote Sens.* **2020**, *159*, 337–351. [[CrossRef](#)]
46. Wan, T.; Lu, H.; Lu, Q.; Luo, N. Classification of high-resolution remote-sensing image using OpenStreetMap information. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 2305–2309. [[CrossRef](#)]
47. Liu, X.; Long, Y. Automated identification and characterization of parcels with OpenStreetMap and points of interest. *Environ. Plann. B* **2015**, *43*, 498–510. [[CrossRef](#)]
48. Rozenfeld, H.D.; Rybski, D.; Gabaix, X.; Makse, H.A. The area and population of cities: New insights from a different perspective on cities. *Am. Econ. Rev.* **2011**, *101*, 2205–2225. [[CrossRef](#)]
49. Ministry of Housing and Urban-Rural Development (MoHURD). *Code for Design of Urban Road Engineering CJJ37*; China Architecture & Building Press: Beijing, China, 2012.
50. Kang, J.; Körner, M.; Wang, Y.; Taubenböck, H.; Zhu, X.X. Building instance classification using street view images. *ISPRS J. Photogramm. Remote Sens.* **2018**, *145*, 44–59. [[CrossRef](#)]
51. Hoffmann, E.J.; Wang, Y.; Werner, M.; Kang, J.; Zhu, X.X. Model fusion for building type classification from aerial and street view images. *Remote Sens.* **2019**, *11*, 1259. [[CrossRef](#)]
52. Gong, P.; Chen, B.; Li, X.; Liu, H.; Wang, J.; Bai, Y.; Chen, J.; Chen, X.; Fang, L.; Feng, S.; et al. Mapping essential urban land use categories in China (EULUC-China): Preliminary results for 2018. *Sci. Bull.* **2020**, *65*, 182–187. [[CrossRef](#)]
53. Stehman, S.V.; Foody, G.M. Key issues in rigorous accuracy assessment of land cover products. *Remote Sens. Environ.* **2019**, *231*, 111199. [[CrossRef](#)]
54. Weng, Q. Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sens. Environ.* **2012**, *117*, 34–49. [[CrossRef](#)]
55. Zhao, W.; Li, Q.; Li, B. Extracting hierarchical landmarks from urban POI data. *J. Remote Sens.* **2011**, *15*, 973–988.
56. Puissant, A.; Hirsch, J.; Weber, C. The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *Int. J. Remote Sens.* **2006**, *26*, 733–745. [[CrossRef](#)]
57. Lu, D.; Weng, Q. Use of impervious surface in urban land-use classification. *Remote Sens. Environ.* **2006**, *102*, 146–160. [[CrossRef](#)]
58. Carlson, T.N.; Ripley, D.A. On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sens. Environ.* **1997**, *62*, 241–252. [[CrossRef](#)]
59. Heiden, U.; Segl, K.; Roessner, S.; Kaufmann, H. Determination of robust spectral features for identification of urban surface materials in hyperspectral remote sensing data. *Remote Sens. Environ.* **2007**, *111*, 537–552. [[CrossRef](#)]
60. Deng, J.; Huang, Y.; Chen, B.; Tong, C.; Liu, P.; Wang, H.; Hong, Y. A methodology to monitor urban expansion and green space change using a time series of multi-sensor SPOT and sentinel-2A images. *Remote Sens.* **2019**, *11*, 1230. [[CrossRef](#)]
61. Foody, G.M. Status of land cover classification accuracy assessment. *Remote Sens. Environ.* **2002**, *80*, 185–201. [[CrossRef](#)]
62. Card, D.H. Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogramm. Eng. Remote Sens.* **1982**, *48*, 431–439.

63. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [[CrossRef](#)]
64. Yu, B.; Lian, T.; Huang, Y.; Yao, S.; Ye, X.; Chen, Z.; Yang, C.; Wu, J. Integration of nighttime light remote sensing images and taxi GPS tracking data for population surface enhancement. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 687–706. [[CrossRef](#)]
65. Su, M.; Guo, R.; Chen, B.; Hong, W.; Wang, J.; Feng, Y.; Xu, B. Sampling strategy for detailed urban land use classification: A systematic analysis in Shenzhen. *Remote Sens.* **2020**, *12*, 1497. [[CrossRef](#)]
66. Johnson, B.A.; Iizuka, K. Integrating OpenStreetMap crowdsourced data and Landsat time-series imagery for rapid land use/land cover (LULC) mapping: Case study of the Laguna de Bay area of the Philippines. *Appl. Geogr.* **2016**, *67*, 140–149. [[CrossRef](#)]
67. Hu, T.; Yang, J.; Li, X.; Gong, P. Mapping urban land use by using Landsat images and open social data. *Remote Sens.* **2016**, *8*, 151. [[CrossRef](#)]
68. Pan, H.; Tong, X.; Xu, X.; Luo, X.; Jin, Y.; Xie, H.; Li, B. Updating of land cover maps and change analysis using GlobeLand30 product: A case study in Shanghai metropolitan area, China. *Remote Sens.* **2020**, *12*, 3147. [[CrossRef](#)]