SPATIOTEMPORAL ANALYSIS OF URBANIZATION CAUSED LAND COVER CHANGE: CASE STUDY IN THE SUBURBAN AREA OF THE YANGTZE RIVER DELTA, CHINA

Xuan Xie(1), Chaomin Chen(1), Yun Du(1), Hasi Bagan(1)

[1] School of Environmental and Geographical Sciences, Shanghai Normal University, Shanghai,200234, China

Email: 1000480061@smail.shnu.edu.cn; 1000480060@smail.shnu.edu.cn; 1000459348@smail.shnu.edu.cn; hasibagan@staff.shnu.edu.cn

ABSTRACT: China formally elevated the academic concept of "Yangtze River Delta integration" to a national strategic decision in 2008. Over the past 12 years, the integration of the Yangtze River Delta has made great progress and development. As urbanization accelerates, in order to promote the integration of the Yangtze River Delta, it is important to explore the land cover changes caused by urbanization. The study area is a demonstration zone of green and integrated ecological development of the Yangtze River Delta, located at the junction of Jiangsu Province, Zhejiang Province, and Shanghai City. Based on multi-source remote sensing data, this study obtained land cover results for 2008 and 2020 in the study area using a random forest classifier. First, we combined the Landsat-5 TM optical image and dual-polarized PALSAR data to classify the land cover in 2008. Then we classified the land cover in 2020 by combining the Sentinel-2 multispectral optical images and dual-polarized PALSAR-2 data. The land cover changes and their driving forces over the 12 years of the study area were analyzed based on the classified maps. The results show that from 2008 to 2020, the area of the built-up in the demonstration zone increased significantly from 863.83 km^2 to 1231.02 km^2, while the area of cropland decreased significantly from 535.92 km^2 to 197.26 km^2, and woodland and shrubland were also reduced. In the past ten years, the permanent population of the demonstration zone has increased from 2.4368 million to 3.1150 million. The proportion of primary and secondary industries is decreasing, while the proportion of tertiary industries is increasing. The main drivers of land cover change are urban expansion, reconstruction of the old city, population growth, optimization of industrial structure, and the rapid advancement of the Yangtze River Delta integration strategy. This paper can provide a reference for the study of the land cover change in suburban areas during urbanization, and supply a decision-making basis for promoting the construction of the demonstration zone of green and integrated ecological development of the Yangtze River Delta and promoting the integration of the Yangtze River Delta.

KEY WORDS: Land cover change; Yangtze River Delta; Random Forest; Multi-source remote sensing data

1. INTRODUCTION

Since the 1990s, Land Use/Cover Change (LUCC) has been the focus and hotspot of global change research (Li, 1996; Chen et al., 2001). With the acceleration of urbanization, land use/cover change is becoming more and more significant. Therefore, land use/cover change during urbanization is attractive in land use research (Liu, et al., 2011; Bagan, et al., 2014; Nolè, et al., 2015). The Yangtze River Delta Urban Agglomeration is one of the most urbanized and rapidly developing areas in China. Therefore, land use/cover change in the Yangtze River Delta has become one of the hotspots for LUCC in China.

In 2008, the State Council formally elevated the academic concept of "Yangtze River Delta integration" to a strategic decision at the national level. Qingpu District, Wujiang District, and Jiashan County are located at the junction of Jiangsu Province, Zhejiang Province, and Shanghai City, and play an important role in the integration process of the Yangtze River Delta. In 2019, the State Council approved the establishment of a demonstration zone of green and integrated ecological development of the Yangtze River Delta in these three counties. Therefore, it is important to investigate the land cover changes caused by urbanization in the demonstration area to promote the integration of the Yangtze River Delta.

Land use/cover classification system is a key issue for LUCC research, and the determination of the classification system is related to the purpose and significance of the study. Therefore, a great deal of work has been done by domestic and foreign scholars (Anderson, et al., 1976; Liu, et al., 2002; Hu, et al., 2017). Synthetic Aperture Radar (SAR) sensors can operate regardless of day or night, with less interference from clouds and rains (Jia, et al., 1998; Hu, et al., 2009; Zhu, et al., 2009). SAR data are capable of capturing the scattering information on the intensity of features, thus compensating for the deficiencies of optical images(Bagan, et al, 2012; La, et al, 2020). This paper combines Landsat-5 TM spectral images and dual-polarized PALSAR data in 2008, Sentinel-2 MSI multispectral data and dual-polarized PALSAR-2 data in 2020, obtained land cover results for 2008 and 2020 in the study area using a Random Forest classifier, and analysis of land cover change and its drivers in the demonstration zone from 2008 to 2020.

2. MATERIALS AND METHODS

2.1 Study area

The demonstration zone of green and integrated ecological development of the Yangtze River Delta spans Jiangsu Province, Zhejiang Province, and Shanghai City (Figure 1). The demonstration zone is adjacent to Dianshan Lake and located in Qingpu District, Wujiang District, and Jiashan County. It covers an area of nearly 2,300 km^2, of which 676 km^2 is in Qingpu District, 1,092 km^2 in Wujiang District, and 506 km^2 in Jiashan County. The demonstration zone is a first step in implementing the national strategy of the Yangtze River Delta integration, which is of great significance in promoting the construction of the Yangtze River Delta integration and exploring new mechanisms for cross-regional governance.

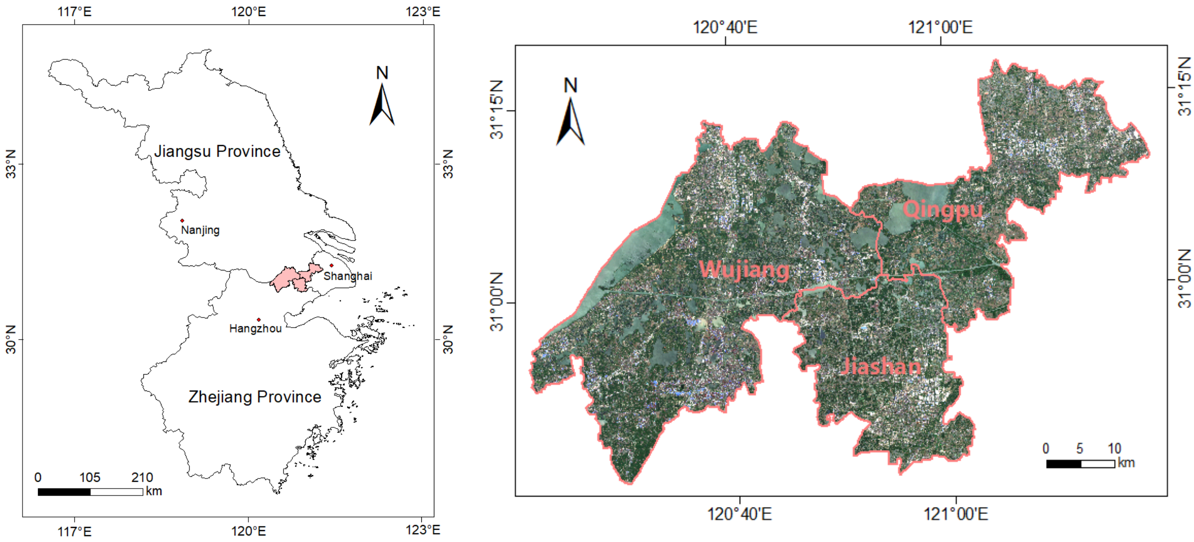

Figure 1 Left: Location of the study area; Right: Sentinel-2 image acquired on April 28, 2020 (RGB = 4, 3, 2).

2.2 Remote sensing data

We choose the Landsat-5 TM, Sentinel-2A MSI, and PALSAR, PALSAR-2 data for supervised classification (Table.1). The data were obtained from the United States Geological Survey (USGS), the European Space Agency (ESA), and the Japan Aerospace Exploration Agency (JAXA), respectively. All remotely sensed data were unified to the same coordinate system (Universal Transverse Mercator (UTM) Zone 50 North map projection, World Geodetic System 84 (WGS-84) datum).

Table 1 Summary of remote sensing data used in this study

Sensors	Product	Time	Spatial resolution (m)
Landsat-5 TM	Collection1 Level 2	2008/04/25	30
		2008/05/02	
Sentinel-2A MSI	Level-2A	2020/04/28	10, 20, 60
ALOS PALSAR-1	RTC	2008/04/28	
		2008/05/15	20
		2008/06/01	
ALOS-2 PALSAR-2	L3.1	2019/05/18	
		2020/05/02	10
		2020/05/30	

2.3 Land-cover class

The reference sites were divided into training and testing sets to ensure spatial disjointing and to reduce the potential for correlation between the sample data and the test data (Table 2). In this study, we defined six landcover classes: built-up, woodland, shrubland, cropland, bare soil, and water (Table.2).

Table 2 The number of training and test sample pixels in 2008 and 2020

Land-cover class	2008		2020	
	Training	Test	Training	Test
Built-up	2565	1235	3259	1046
Woodland	2525	1091	3100	1073
Shrubland	2752	1316	3080	1389
Cropland	2559	1082	3103	1144
Bare soil	3273	1111	2939	975
Water	3046	1239	2968	1263

2.4 Random Forest (RF) Classification

Based on the *Scikit-learn* (Pedregosa, et al., 2011) in Python, random forest classifiers were implemented to classify land cover. The random forest is an integrated machine learning algorithm consisting of a set of multiple decision trees (Breiman, et al., 2001), which exhibits powerful performance in many realistic tasks. In this paper, a grid search based on the scores of out-of-bag (OOB) samples was used to perform the selection of two important hyperparameters (i.e., the number of trees (t) and the number of randomly selected features at each node (f)) in the random forest model, while all other hyperparameters were kept in their default settings. For the classification in 2008, t was set from 100 to 1000 in intervals of 100; f was set from 1 to 8 in intervals of 1. For the classification in 2020, t was set from 100 to 1000 in intervals of 100; f was set from 1 to 12 in intervals of 1. By comparing the OOB scores of random forest models trained by two different hyperparameter combinations, two optimal hyperparameter combinations (i.e., t = 1000, f = 3 for the classification in 2008; t = 1000, f = 8 for the classification in 2020) were finally selected.

3. RESULTS

3.1 Accuracy assessment

The land-cover results for the demonstration zone in 2008 and 2020 are shown in Figure 2. The classification maps were evaluated by test samples, and the results of the accuracy evaluation were represented in the form of a confusion matrix (Table 3), with the evaluation indicators being overall accuracy (OA), producer's accuracy (PA), and user's accuracy (UA). The overall accuracy of the classification results in 2020 (94.46%) was higher than that in 2008 (89.27%), which indicates that images from different sensors have a significant impact on classification accuracy. Water was classified with the highest accuracy (close to 100%). The classification accuracy was relatively high for woodland, followed by bare soil and shrubland. The built-up and cropland had relatively low classification accuracy.

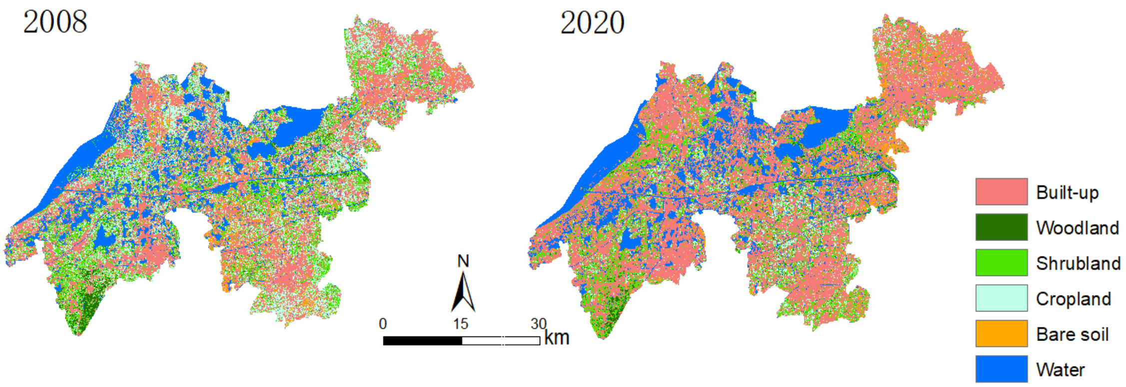

Figure 2 Land-cover classification maps for 2008 and 2020

Table 3 Classification accuracy for the land-cover maps for 2008 and 2020

Class	2008		2020	
	PA (%)	UA (%)	PA (%)	UA (%)
Built-up	78.38	91.49	96.75	90.84
Woodland	90.74	94.56	96.09	99.23
Shrubland	91.41	84.66	93.95	90.31
Cropland	83.18	78.60	84.27	90.18
Bare soil	92.17	87.90	95.69	100.00
Water	99.27	99.35	100.00	97.91
OA (%)	89.27		94.46	

3.2 Area change in land-cover class

As shown in Table 4 and Figure 3, the land cover in the demonstration zone changed significantly between 2008 and 2020. The highest percentage of the area was in the built-up class, followed by shrubland and water. It is observed that there was a large increase in the area of the built-up class and a sharp decrease in the area of the cropland class.

Table 4 Area of each land-cover class in 2008 and 2020 (unit: km^2)

Class	2008	2020
Built-up	863.83	1231.02
Woodland	64.62	44.17
Shrubland	392.84	354.29
Cropland	535.92	197.26
Bare soil	122.71	139.88
Water	439.82	453.10

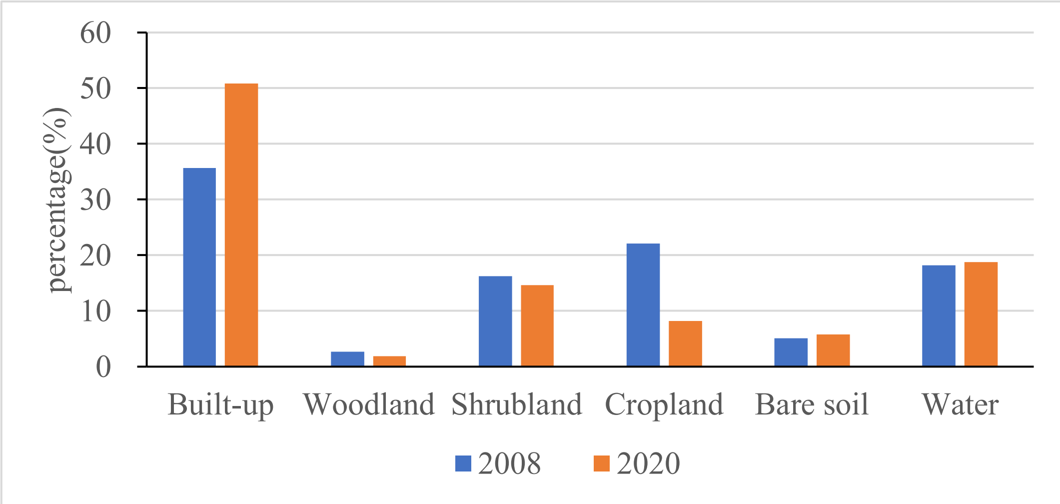

Figure 3 Percentage of each land-cover class in 2008 and 2020

The change in area by land-cover class from 2008 to 2020 is shown in Table 5. Land cover change in the demonstration zone between 2008 and 2020 was mainly in the form of the conversion between shrubland, cropland, and built-up.

Table 5 Land-cover conversion matrix from 2008 to 2020 (unit: km^2)

	Built-up	Woodland	Shrubland	Cropland	Bare soil	Water	Class Total
Built-up	**645.95**	16.60	***165.10***	***273.71***	***65.13***	***64.53***	1231.02
Woodland	3.50	**13.78**	15.67	8.07	2.43	0.73	44.17
Shrubland	***71.36***	***25.58***	**104.79**	120.58	18.67	13.31	354.29
Cropland	47.77	4.80	55.14	**65.03**	16.94	7.60	197.26
Bare soil	45.73	2.34	28.98	40.20	**11.39**	11.24	139.88
Water	49.52	1.52	23.16	28.34	8.16	**342.41**	453.10
Class Total	863.83	64.62	392.84	535.92	122.71	439.82	2419.73
Class Changes	217.87	50.84	288.05	470.89	111.32	97.41	1236.39

4. DISCUSSION

Over the past 12 years, the Yangtze River Delta integration has made great progress and development. The land cover of the demonstration zone has undergone significant changes. There was a significant increase in the built-up and a decrease in cropland. This reflects the acceleration of urbanization in the demonstration zone over the past decade, accompanied by a sharp decline in green space coverage.

The increase in the built-up land area is closely related to population growth. Population growth has led to rapid urbanization and increased demand for construction land, prompting the conversion of large amounts of land into construction land. However, urban development cannot be promoted by unlimited population expansion. The excessive population will put great pressure on land resources and affect the healthy development of urbanization. The integration process of the Yangtze River Delta should be accompanied by effective population control. Over the past ten years, the permanent population of the demonstration zone has grown from

2.4368 million to 3.1150 million. Compared to Wujiang District and Jiashan County, Qingpu District shows the most significant population growth (from 0.7898 million to 1.1290 million). Qingpu District has the fastest population growth due to its location in Shanghai, which is more attractive to the population. Whereas Jiashan County has a small population base and the population of Wujiang District is growing slowly. Therefore, the labor-intensive industries of Qingpu District can be relocated to the surrounding counties. In order to promote the coordinated development of population and land in the demonstration zone, it is suggested that relevant policies can be combined to increase the attractiveness of the surrounding counties and divert the population.

The economic development and urbanization of the demonstration zone will inevitably require large amounts of land resources. The occupation of arable land is the most direct way, which can be reflected in the conversion of cropland to built-up areas. In the past ten years, the proportion of industrial structure has changed from 2.17% to 2.06% (for the primary industries), 61.96% to 49.58% (for the secondary industries), and 35.87% to 48.36% (for the tertiary industries), respectively. With the rapid growth of gross domestic product and the continuous optimization and upgrading of the industrial structure, the structure of land use is constantly being transformed. Since the development of tertiary industries requires considerable construction land as a foundation, a large amount of cropland and shrubland is converted into built-up areas.

It is noteworthy that a large proportion of cropland was converted to shrubland. On the one hand, the policy of returning farmland to forests in the demonstration zone has achieved satisfactory results. On the other hand, there has been a change in cropping patterns. In the pursuit of economic development, cash crops have replaced food crops. For instance, rice paddies have been replaced by orchards. Over the past decade, the level of intensive land use in the demonstration zone has been rising, and agricultural output has been rising. However, while developing the economy, attention should also be paid to the protection of basic farmland.

5. CONCLUSION

Based on multi-source remote sensing data, this paper acquired the land cover results of the demonstration zone of green and integrated ecological development of the Yangtze River Delta in 2008 and 2020 using a random forest classifier. The spatiotemporal characteristics and driving forces of land cover change in the demonstration zone from 2008 to 2020 were further analyzed. The results show a significant increase in the amount of construction land in the demonstration zone, mainly due to urban expansion, renovation of old cities, and population growth.

Additionally, the area of cropland has decreased, and a part of it has been converted to shrubland. The main manifestation of land cover change in the demonstration zone is the transformation of vegetation types into building types. The main reasons are the optimization and upgrading of the industrial structure, the policy of returning farmland to forests, and changes in planting methods.

This study can serve as a reference for the land cover change in the process of urbanization, support the study of urban heat island and climate change, and provide a decision-making basis for the construction of the demonstration zone of green and integrated ecological development of the Yangtze River Delta and the promotion of Yangtze River Delta integration.

6. ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Grant no. 41771372), and the Shanghai Science and Technology Committee (Grant no.18511102300).

References

Anderson, J. R., Hardy, E. E., Roach, J. T., *et al.*,1976. A land use and land cover classification system for use with remote sensor data. Geological Survey Professional Paper 964.
Bagan, H., Kinoshita, T., Yamagata, Y., 2012. Combination of AVNIR-2, PALSAR, and Polarimetric Parameters for Land Cover Classification, IEEE Transactions on Geoscience and Remote Sensing, 50(4): 1318–1328.
Bagan, H., Yamagata, Y., 2014. Land-cover change analysis in 50 global cities by using a combination of Landsat data and analysis of grid cells. Environmental Research Letters, 9(6): 064015.
Breiman, L.,2001. Random forests. Machine Learning, 45(1): 5-32.
Chen, Y., Yang, P., 2001. Recent progresses of international study on land use and land cover change. Economic Geography, 2001(01): 95-100.
Hu, Z., Li, H., Du, P., 2009. Study on the extraction of texture features and its application in classifying SAR images. Journal of China University of Mining &Technology, 38(03): 422-427.
Hu, C., Guo, X., Lian ,G., *et al.*,2017. Effects of land use change on ecosystem service value in rapid urbanization areas in Yangtze River Delta——a case study of Jiaxing City. Resources and Environment in the Yangtze Basin, 26(03): 333-340.
Jia, Y., 1999. Fusion of Landsat-TM and SAR images based on principal component analysis. Remote Sensing Technology and Application, 1998(01): 3-5.
Li, X., 1996. Review of the international research on land use/land cover change. Acta Geographica Sinica., 1996(06): 553-558.
Liu, H., Zhang, Y., Zheng, W., 2011. Evaluation on spatiotemporal development and interaction of intensive urban land use and urbanization: Case studies of the cities in the Bohai Rim Region. Geographical Research, 30(10): 1805-1817.
Liu, J. Y., Zhuang, D. F., Luo, D., *et al.*, 2002. Land-cover classification of China: integrated analysis of AVHRR imagery and geophysical data. International Journal of Remote Sensing, 2002: 1-16.
Nolè, G., Murgante, B., Calamita, G., *et al*., 2015. Evaluation of Urban Sprawl from space using open source technologies. Ecological Informatics, 26: 151-161.
Pedregosa, F., Varoquaux, G., Gramfort, A., *et al*, 2011. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12: 2825-2830.
La, Y., Bagan, H., Yamagata, Y.,2020. Urban land cover mapping under the Local Climate Zone scheme using Sentinel-2 and PALSAR-2 data, Urban Climate, 33.
Zhu, L., Guo, W., Yu, W., 2009. Analysis of SAR satellite development history and tendency. Modern Radar, 31(04): 5-10.