

# 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

Dandan Liu<sup>a</sup>, Nengcheng Chen<sup>a,b,\*</sup>, Xiang Zhang<sup>a</sup>, Chao Wang<sup>a</sup>, Wenying Du<sup>a</sup>

<sup>a</sup> State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China

<sup>b</sup> Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, China



## ARTICLE INFO

### Keywords:

Urban land mapping  
Large-scale long time series  
Google Earth Engine  
OpenStreetMap  
Landsat  
Middle Yangtze River basin

## ABSTRACT

Long time series (e.g. 30 years) urban land observations from remote sensing images are important for urban growth modeling as well as for the goal of sustainable urban development. However, updates to regional and even global maps are infrequent due to the cost and difficulty of collecting representative training data and the requirement for high-performance computation in processing large amounts of images. In this study, a semi-automatic large-scale and long time series (LSTS) urban land mapping framework is demonstrated by integrating the crowdsourced OpenStreetMap (OSM) data with free Landsat images (~13,218 scenes) to generate annual urban land maps in the 317,000 km<sup>2</sup> middle Yangtze River basin (MYRB) from 1987 to 2017 facilitated by Google Earth Engine (GEE). Random training samples for latest year were generated based on the updated OSM land use data after a manual topological conflict processing and uploaded to GEE for automatic image classification. For each historical year, training samples were obtained with a proposed transferring schema by which only the unchanged were selected through a change detection analysis. The annual spectral indices and texture feature maps acquired from the surface reflectance dataset were also added to the original bands. Finally, the classified maps were downloaded from GEE and a spatial-temporal consistency checking was further performed. Based on independent samples, the overall accuracies and kappa coefficients of all years ranged from 98% to 99% and 0.65 to 0.85, respectively. Our product when compared with current 30 m land-cover products showed similar accuracies but more spatial details. The characteristics of pattern, traces, and hotspots of urban expansion were further explored. This study provides a more convenient procedure for LSTS urban land mapping especially for areas where large-scale field sample-collection is difficult and little historical crowdsourced datasets are available. The resultant dataset is expected to provide consistent details about the spatial distribution of urban land in MYRB. We highlight the potential use of this proposed framework to be applied and validated to other parts of the world to help better understand and quantify various aspects of urban-related problems.

## 1. Introduction

### 1.1. Scope

Spontaneous growth of urban land has become a common phenomenon all over the world, especially in developing countries (e.g., Asia and Africa) (Arnold and Gibbons, 1996; Seto et al., 2011; Wang et al., 2012), in the past two centuries. In the next few decades, the urban population is projected to expand by an additional 2.5 billion (Seto et al., 2011). The outward expansion of suburban and rural

development is also associated with important environmental and societal problems such as soil degradation (Yan et al., 2015), air pollution (Bereitschaft and Debbage, 2013; Zhu et al., 2017), surface runoff (Eshtawi et al., 2016), diminishing human health (Cao et al., 2018), spatial equity and increasingly expensive infrastructure facilities and services (Wei and Ewing, 2018). The urgency of understanding and addressing these problems for the sustainable development goal (SDG) (Boyd et al., 2018) necessitates timely and accurate monitoring of urban dynamics based on reliable data (Xie and Weng, 2017). However, at present, we are still unable to quantify the magnitude and pace of

\* Corresponding author at: State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China.

E-mail address: [cnc@whu.edu.cn](mailto:cnc@whu.edu.cn) (N. Chen).

urbanization in a consistent manner at higher spatial ( $\sim 30$  m) and temporal (e.g. annual) resolution at the regional or even global scale (Ban et al., 2015; Chen et al., 2015).

### 1.2. Large-scale and long time series (LSLTS) urban land mapping

Currently, for large-scale urban land mapping, accurate and consistent monitoring of urban land cover is frequently precluded by either coarse spatial (e.g., the 500 m Moderate Resolution Imaging Spectroradiometer (MODIS) land cover product (Friedl et al., 2010), the 1 km urban land product from Liu et al. (2012)) or temporal (e.g., the 10-year GlobeLand30 product from Chen et al. (2015), the 5-year global urban products from (Liu et al., 2018), and the Global Human Settlement Layer (GHSL) Built-up Grid created by the European Commission ([https://ghsl.jrc.ec.europa.eu/ghs\\_bu.php](https://ghsl.jrc.ec.europa.eu/ghs_bu.php))) resolution of such datasets. The former does not provide sufficient thematic information to capture human activities (Ban et al., 2015), and the latter lacks the temporal consistency needed to represent high-order complexity (Sexton et al., 2014) as well as perform trends analysis. Although there are also studies of urban impervious analysis at higher spatial and temporal resolution, e.g., (Li et al., 2015; Sexton et al., 2014; Song et al., 2016; Zhang et al., 2017), most of them have mainly focused on small areas that are covered by only one or several image scenes. Such a process over large areas, such as a continent, will always have limitations in terms of the availability of an extensive collection of reference data as well as large time and computation costs.

With the long-term archive and free availability of Landsat and similar image data, the development of global land cover (GLC) (Ban et al., 2015) data products at a 30 m resolution has become feasible. Such products are considered a superior option for the next generation of GLC maps, since most significant human activities and human impact on the land system can be captured by scrutinizing of shapes, sizes and patterns at this scale (Chen et al., 2015). Moreover, the freely available Landsat archive data from the past forty years (1972 to present) offer unprecedented opportunities to document historical land-cover change with a longer range than that of MODIS (Dong et al., 2015), which became available only after 2000. The increased availability of the Landsat archive has the potential to change not only the way regional land-cover maps are produced, but also how they are analyzed (Gomez et al., 2016).

### 1.3. Automatic land-use/land-cover (LULC) mapping

Nevertheless, there are two main problems that hinder the large-scale and long time series (LSLTS) mapping of urban land with supervised classification approaches: (1) the necessity of larger training datasets; and (2) the expensive computation cost for data processing in terms of both money and time (Yang et al., 2017). Automation of satellite-based human settlement mapping is urgently needed to utilize historical archives of satellite data for urgent issues of urban development at the global scale (Miyazaki et al., 2016).

For the first problem, traditional means of training samples collection (e.g., visual image interpretation or field surveys) can be difficult or expensive (Demir et al., 2012). Crowd-sourcing platforms may allow researchers to scale—at low cost—the labeling method and to construct larger and more comprehensive ground-truth datasets (Goldblatt et al., 2016). The possibility of using crowdsourced geographic data (CGD) to replace training data in the process of land use and land cover (LULC) classification has also been investigated by many scholars. Arsanjani et al. (2013a) compared the resultant land-use map with existing urban map and achieved a kappa index of 89%, proving the suitability of using OpenStreetMap (OSM) data for training site definition. See et al. (2015) used geographically weighted regression and crowdsourced validation data from Geo-Wiki to create two hybrid global land-cover maps. Based on OSM data and Landsat time series, Johnson and Iizuka (2016) explored the potential of automatically extracting crowd

sourced data for rapid LULC mapping. Fonte et al. (2017) combined OSM data with the GlobeLand30 product and finally improved the accuracy of the product. By combining OSM data and remote sensing data to fill the gaps, Schultz et al. (2017) produced a land cover prototype with complete coverage in Heidelberg, Germany with an overall accuracy of 87%, better than the already existed products. Nevertheless, none of these studies tried to apply the generated samples from CGD to LSLTS LULC mapping.

For the second problem, with the availability of cloud-based platforms such as Google Earth Engine (GEE) (<https://code.earthengine.google.com/>) (Gorelick et al., 2017), it is now feasible to monitor urbanization at multi-spatial and temporal resolutions and to understand urban dynamics globally (Goldblatt et al., 2016). GEE is a cloud-based platform for scientific analysis and visualization of petabyte-scale geospatial datasets. It stores several decades of historical images and scientific datasets and enables parallel computing and feasible big data processing in a large study area. Based on this platform, many large-scale or even LSLTS LULC mapping projects such as crops (Dong et al., 2016; Shelestov et al., 2017; Xiong et al., 2017), urbanization (Goldblatt et al., 2016; Liu et al., 2018; Patel et al., 2015), gain and lose of vegetation (Huang et al., 2017), and wetland landscape (Alonso et al., 2016) have been conducted successfully all over the world, indicating the usefulness of GEE for extensive earth observation.

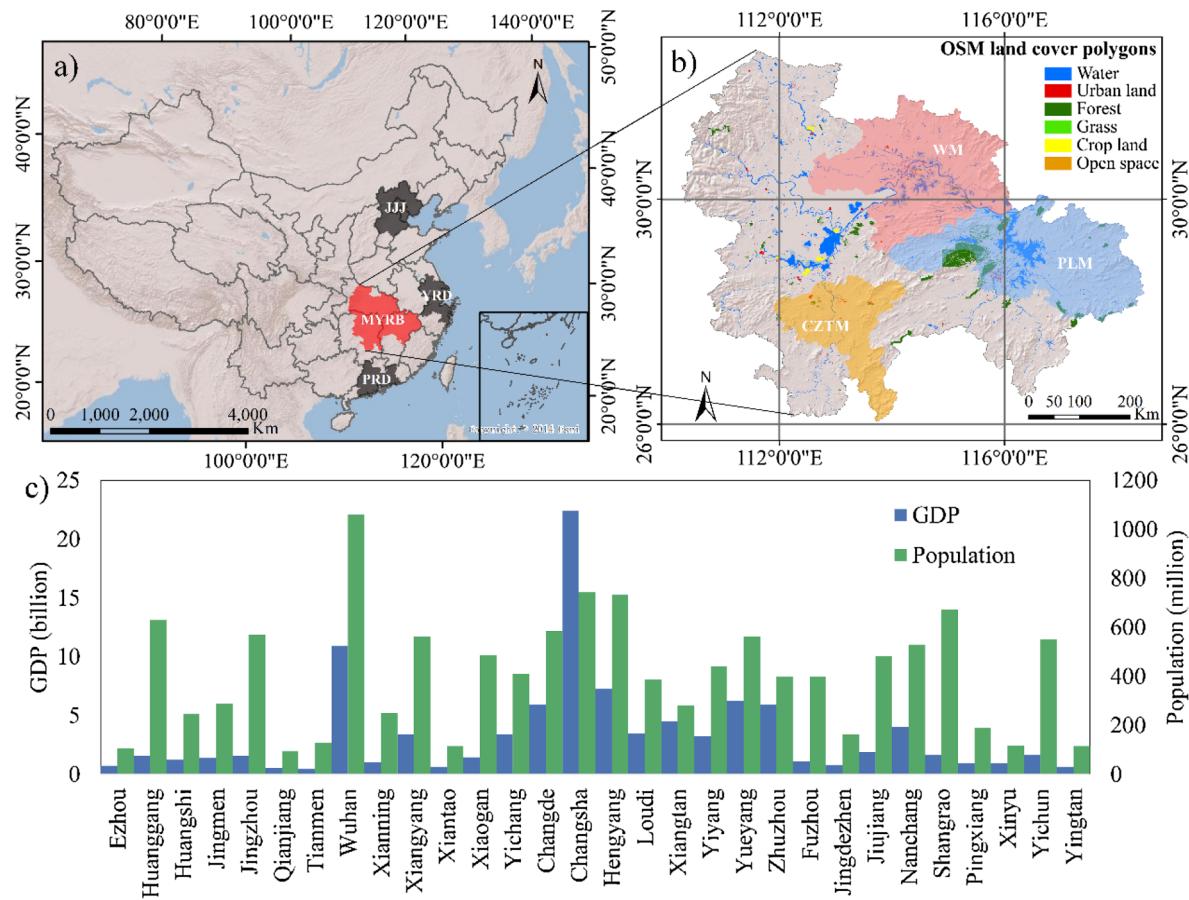
### 1.4. The goal of this study

The middle Yangtze River basin (MYRB) urban agglomeration has experienced unprecedented urbanization in recent decades (Liu and Chen, 2017). Moreover, with the issue of the “Development Plan for the MYRB urban agglomeration, 2014–2020” (the State Council of China), the MYRB urban agglomeration is going to be transformed into a new support hub for China’s economy, with an urbanization rate goal reaching 60% in 2020 (Yu et al., 2018). In this context, more of the rural population will move from rural areas to cities and this continuous migration will pose great threats to the urban environment in terms of resource consumption and urban expansion, especially for key national river basins (Poyang Lake and Dongting Lake) and the grain production base in the region.

Given the above discussions, the main objective of this study is to develop a semi-automatic framework facilitated by GEE to map long time series urban land on an annual and regional basis using Landsat imagery along with OSM data in the 317,000 km<sup>2</sup> MYRB urban agglomeration. Reference data sampling is often difficult and labor-intensive especially for regional and even global long time series mapping. The available crowdsourced data provides a new solution to this problem. However, for some areas such as MYRB, the usable historical data is limited. Thus, how to make full use of the up-to-date data for historical mapping is the key issue. The specific objectives of this study are as follows: (1) to demonstrate a methodology for 30 m LSLTS urban land mapping based on latest OSM data; (2) to evaluate how accurate it is to use OSM-derived training data in long time series image classification; and (3) to quantify and characterize the spatiotemporal dynamics of urban land during the last three decades in the MYRB urban agglomeration.

## 2. Study area

The MYRB urban agglomeration, which is located in the central part of China (Fig. 1), is the fourth largest urban agglomeration in China and the second largest in the Yangtze River Economic Belt (YREB) with an area of 317,000 km<sup>2</sup>, accounting for 3.30% of land in China. It consists of 31 cities of different sizes across Wuhan Metropolis (WM), Chang-Zhu-Tan Metropolis (CZTM), and Poyang Lake Metropolis (PLM) (Fig. 1). As the most important hub of Chinese transportation and the focal point of the economy, culture, education, and technology in central China, the study area has a population of 121 million, and a



**Fig. 1.** The location of the study area. (a) The four largest urban agglomerations in China (JJJ: Jing-Jin-Ji; YRD: Yangtze River Delta; PRD: Pearl River Delta); (b) OSM land cover data distribution in the study area; and (c) the GDP and population of all the cities within MYRB in 2014.

gross domestic product (GDP) of 980 billion US dollars in 2014, accounting for 8.8% and 8.8% of the whole country, respectively (National Development and Reform Commission).

### 3. Materials and methods

The proposed mapping framework is presented in Fig. 2. First, the annual available Landsat images and the related spectral indices were collected and calculated in GEE (Section 3.1). In addition, the OSM related data were collected and processed manually in ArcGIS to generate the training samples (Section 3.2). Then, the generated samples were uploaded to GEE, and the annual classification was performed by the proposed sample-transferring schema based on which the samples located on the unchanged areas were selected for previous years' classification (Section 3.3). Finally, the post processing (Section 3.4) and accuracy assessment (Section 3.5) were performed. Except for the generation of reference data for training and validation as well as post classification analysis, most of the data processing was performed automatically in GEE. In this study, urban areas on remote sensing images are defined as sites that are dominated by a built environment, including all non-vegetative, human-constructed elements and are defined as features with tags of all non-vegetative, human-constructed elements including road networks and buildings in OSM data.

#### 3.1. Remote sensing data and processing

Good quality observations for each month of every year were unavailable (Dong et al., 2015), and it was necessary to combine all images in one year for analysis. All available standard Level 1 Terrain-corrected (L1T) orthorectified surface reflectance images, with cloud

cover less than 70%, from Landsat TM/ETM + /OLI archived in the GEE platform were used for this study. Thirty-one path/rows cover the whole study area (Fig. 3). Poor observations (clouds and cloud shadows), were removed according to the Landsat surface-reflectance pixel-quality band (Roy et al., 2014).

Five spectral indices, including the normalized difference vegetation index (NDVI) (Tucker, 1979), normalized difference built-up index (NDBI) (Zha et al., 2003), modified normalized difference water index (MNDWI) (Xu, 2006), and the new built-up index (BuEI) and new soil index (SoEI) (Feyisa et al., 2016), were further calculated in GEE. The BuEI and SoEI can efficiently distinguish impervious surface and soil (Feyisa et al., 2016). The annual maximum vegetation composites (MVC) were generated based on NDVI to ameliorate seasonal and inter-annual fluctuations of NDVI and to mitigate the impact of agricultural lands in the suburban and peri-urban areas (Tsendlbazar et al., 2015). In this paper, we referenced the MVC, and the annual minimum built-up composites (MBC) based on NDBI were also generated to mitigate the impact of bare land. Moreover, three texture features, including the dissimilarity (DISS), entropy (ENT), and the angular second moment (ASM), were also added to the bands for image classification because they were identified as effective indicators for the texture description of different land-cover types (Puissant et al., 2005). The slope data were further computed from the 30 m digital elevation model (DEM) data in the Shuttle Radar Topographic Mission (SRTM) dataset (Wang et al., 2015) to mask out regions of high elevation and/or steep slopes where urban areas are not likely to occur.

All indices were computed using the GEE Code Editor and all images were clipped according to a common geographic extent and assembled into a time series dataset ready for use.

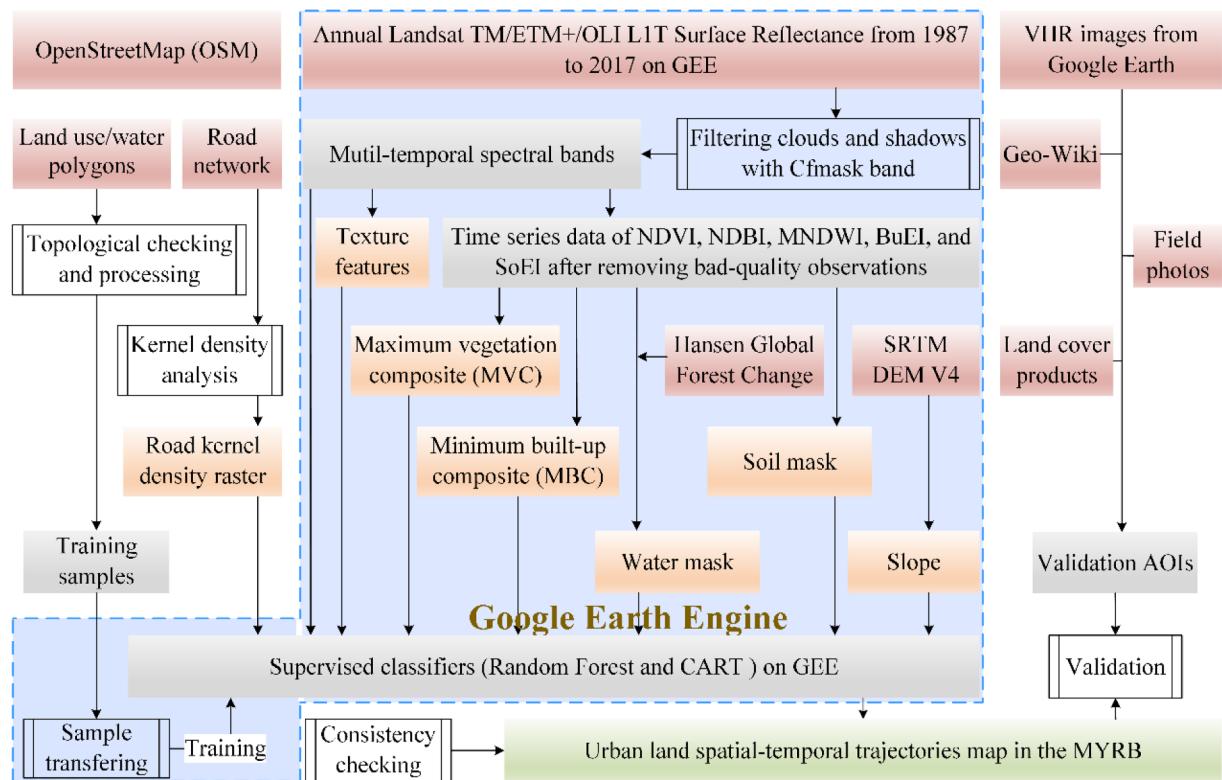


Fig. 2. The workflow for the semi-automatic annual urban mapping in this study. (Procedures in the blue frame are automatic.)

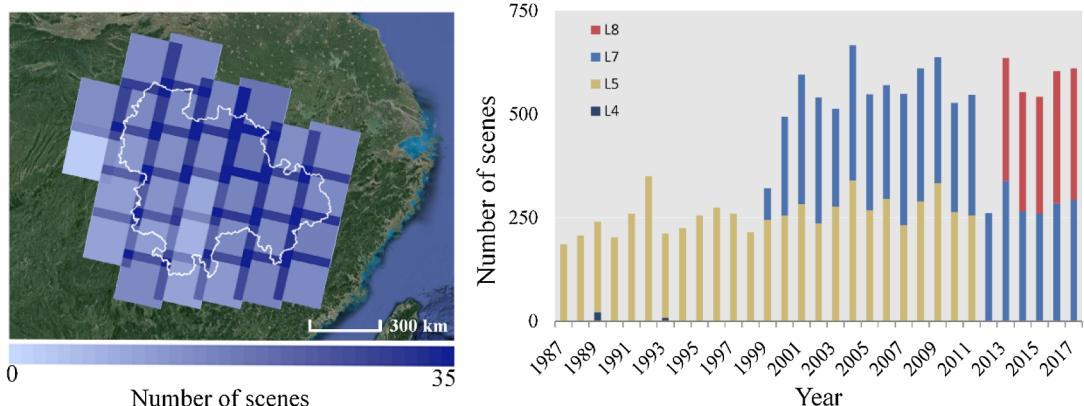


Fig. 3. Distribution of observation numbers of available images in 2017 and number of images by sensors (Landsat TM 4, Landsat TM5, and Landsat 7 ETM + and Landsat 8 OLI) used for each year.

### 3.2. Training samples generation

#### 3.2.1. OSM data

The OSM project was created by a user community dedicated to producing open-source, geocoded datasets of infrastructural resources (Weiss et al., 2018). The structure of the data is composed of nodes, ways (open and closed, as well as areas) and relations. The dataset can be available on many platforms. In this study, we downloaded the data in shapefile format from Geofabrik (<http://download.geofabrik.de/asيا/>, last accessed on 11/10/2017). We used the “landuse”, “natural” and “water” datasets from the OSM database as they represent a variety of manmade and natural LULC types and are usually sufficiently large to detect at this resolution (Johnson et al., 2017). However, the “natural” dataset only include 5 features in the study area. Therefore, only the “landuse” and “water” datasets were used in this paper (Fig. 1 b)). Moreover, major roads were also selected from the road polyline dataset based on the following constraints: [highway] = ‘construction’

or ‘cycleway’ or ‘primary’ or ‘living\_street’ or ‘motorway’ or ‘pedestrian’ or ‘residential’ or ‘secondary’ or ‘tertiary’ or ‘trunk’ or ‘subway’. Based on this data, a major road kernel density (Parzen 1962) layer was generated with the bandwidth set as 500 m in ArcGIS and added to the remote-sensing images assuming that the urban areas are always near roads (Song et al., 2016). All land-use classes that could not be clearly associated with a corresponding LULC class (further described in Section 3.2.2) were excluded, as were polygons with an area of less than 900 m<sup>2</sup> (since Landsat pixels are 30 m × 30 m in size).

#### 3.2.2. OSM data processing

Contributions have dissimilar geometrical accuracy and frequently overlap each other (i.e., some areas are given several attributes) (Arsanjani et al., 2013b). Therefore, the crowdsourced datasets were converted into LULC classes via a sequence of steps as follows:

- (1) Topological conflict processing: a challenge in using OSM data for

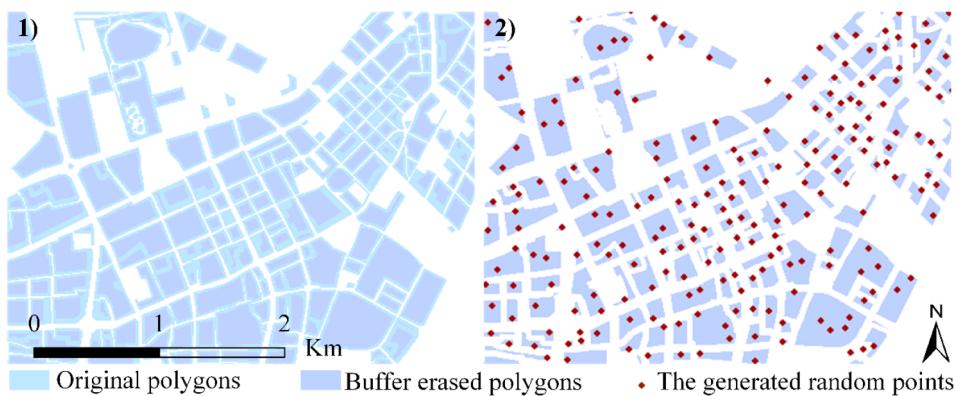


Fig. 4. Illustration of the random sample points generation process.

LULC mapping is the fact that OSM polygons can contain multiple LULC types (Johnson and Iizuka, 2016); therefore, in the first step, several rules were defined to remove the inconsistency among the polygons using a spatial topology rule in ArcGIS: for each single layer, a topological checking was conducted with any intersected polygons removed; for every two layers, the overlapping parts of two layers were calculated and extracted from both layers; finally, the left polygons were filtered with areas less than 900 m<sup>2</sup> (1 pixel) abandoned.

- (2) Sample points generation: in this step, random points were generated in the filtered polygons (Johnson and Iizuka, 2016). Considering that the random points may be located on the common boundary of two polygons with different classes, which may further result in mixed classes at one location when using these sample points training the images on cloud. Therefore, a buffer analysis with an interior buffer distance of 30 m was performed before generating the points (Fig. 4). Then, the left polygons with areas less than 8100 m<sup>2</sup> (9 pixels) were removed to further improve the class purity of the selected sample points. For the remaining polygons, random points were generated within each polygon (with a minimum point spacing of 90 m, or three Landsat pixels), to reduce the spatial autocorrelation in the training set.
- (3) Conversion of OSM features to LULC class: there is a large diversity of tags in OSM (Fonte et al., 2017), and some of the tags can be converted directly to LULC classes, such as “forest”, “meadow”, “orchard”, “grass”, “commercial”, and “residential”. However, there are also some classes, such as “park” and “military”, which include at least two types of land use and some of the land use only accounts for a small proportion of the whole area, e.g., wetland (only 23 polygons). Therefore, we removed these land-use polygons. We finally acquired 51,037 sample points (Fig. 5) and the details are listed in Table 1.

### 3.3. Classification on GEE based on the training samples

#### 3.3.1. Sample transferring

Based on the generated samples, further image classification can be conducted on GEE. However, for the mapping of previous years, the usable historical archives of OSM dataset is limited in the study area. Therefore, we proposed a training samples transferring schema (Fig. 6) to make use of the training samples of latest year for historical long time series urban land mapping in this study.

For each previous year, the tasseled cap transformation (Kauth and Thomas 1976) was performed on the corresponding year's image as well as the image of the latest year (2017 in this study). The components of the transformation were then used as the main input vectors for change vector analysis (CVA) (Chen et al., 2011) to detect the changes between two periods. There are three main types of methods for selecting the threshold of change/no-change for each year: the Percentile,

Otsu's method and Kapur's algorithm (Ilsever and Ünsalan, 2012). We chose the Otsu method in this study to maximize the variance between change objects and unchanged objects (Li et al., 2016). However, because the area of this research is relatively large, computation time-out errors always occur during processing. Therefore, we proposed an improved method that can reduce the data volume used for the threshold computation (Fig. 6d)). After CVA, edge detection was performed based on canny algorithm (Canny, 1986), and the two pixels adjacent to the central-edge pixel were selected as the edge buffer. Then the threshold was decided by Otsu. Finally, the unchanged region of the image was used to sample the training points to train the classifier for the corresponding year. This proposed approach is able to avoid some overestimations of change objects as shown in Fig. 6(g) and (h).

#### 3.3.2. Classification on GEE

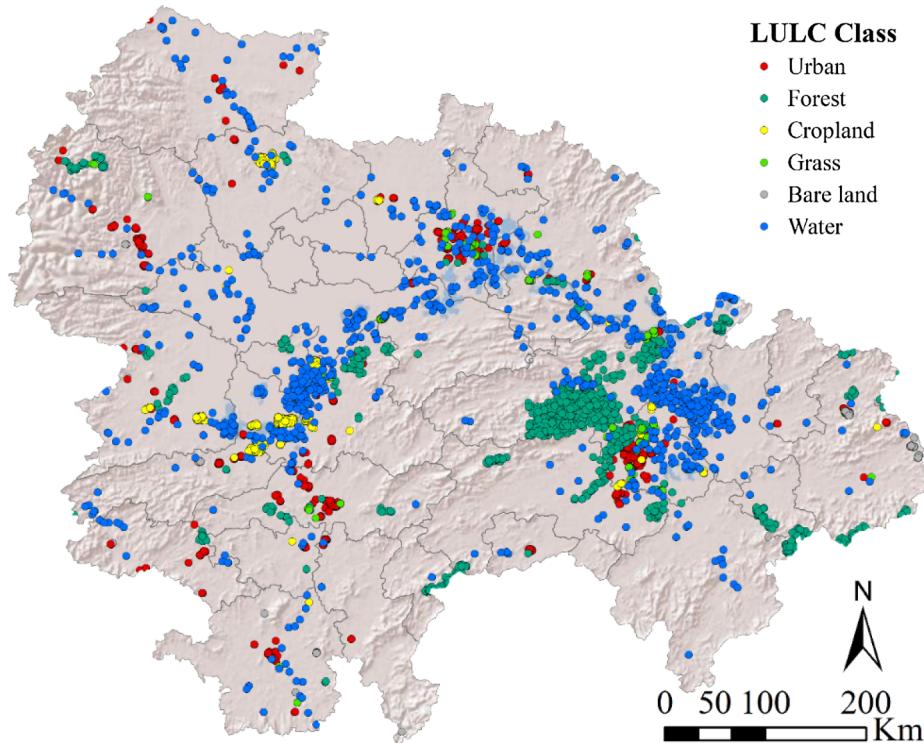
Considering that water in the greenest-pixel image appears to be white influenced by plants in it, the MNDWI band as well as the water band from Hansen Global Forest Change products (Hansen et al., 2013) archived on GEE were used to mask water out before the classification. In addition, as built-up surfaces and bare land show similar surface reflectance (Feyisa et al., 2016) and the volume of bare land samples is relatively small, a soil mask was also performed according to the SoEI. The threshold values of the two indices were 0 and 0.7 based on previous research (Xu, 2006) and trial and error through visual interpretation, respectively. Then, online classification was conducted through GEE to obtain the urban land map year by year.

GEE offers several classification algorithms, including Classification and Regression trees (CART), random forests (RF), support vector machine (SVM), and Naïve Bayes classifier (Shelestov et al., 2017). Preliminary analyses showed that CART provided superior land-cover extrapolations for phenologically invariant cover types but over-fit the training data in phenologically variable classes (Sexton et al., 2013); RF is superior to other classifiers, especially when applied to large geographic scales and when the data are noisy and high dimensional (Goldblatt et al., 2018). Based on the strength of each of these algorithms, a hybrid of these two approaches was used here to avoid overfitting of the training sample. A total of 500 trees were constructed for the RF, and a total of 100 trees were built for the CART to classify the images.

After obtaining the classified classes from both classifiers, the output images were further converted into two land categories (urban and non-urban) because the focus of this study is the spatiotemporal dynamics of urban land. Finally, the pixels that were both classified as urban land by the two methods were labeled as urban land and the others labeled as non-urban.

#### 3.4. Consistency correction

There often exists high variability in the urban-area estimates



**Fig 5.** Distribution of the generated training samples based on OSM data.

**Table 1**  
LULC classes derived from the OSM “landuse” and OSM “water” shapefiles.

OSM class	Number of polygons	Area (ha.)	LULC	Number of points
Commercial	386	1378	Urban	15,142
Residential	4070	54,568	Urban	
Industrial	721	15,420	Urban	
Recreation_ground	31	77	Urban	
Retail	176	647	Urban	
Farm	202	42,719	Cropland	1733
Forest	5374	455,225	Forest	22,534
Natural_reserve	11	122,729	Forest	
Scrub	30	404	Forest	
Orchard	8	73	Forest	
Grass	121	168	Grass	706
Meadow	130	1897	Grass	
Quarry	54	1200	Bare land	257
River	339	412,311	Water	10,665
Water	1230	746,934	Water	
Reservoir	37	6758	Water	

derived from the annual raw-classification results; for example, areas of higher impervious proportions may decrease in subsequent years at some locations (He et al., 2017; Li et al., 2015; Zomlot et al., 2017). In this study, we combined the methods described in Li et al. (2015) and Zomlot et al. (2017) for consistency correction of the resultant maps. We first performed a temporal filtering for the classified time-series raw urban products using a dynamic time-filtering window. Then, a change-trajectory analysis was performed according to the steps described in Zomlot et al. (2017). Because we only consider two classes, i.e., urban and non-urban, Rule III was modified, that is, “if the processed segment was urban dominated, we would modify it and keep it in the trajectories from non-urban to urban earlier or all as urban. Otherwise, we would maintain it as non-urban or from non-urban to urban more recently”.

The trajectory analysis began from 2017 to previous years because

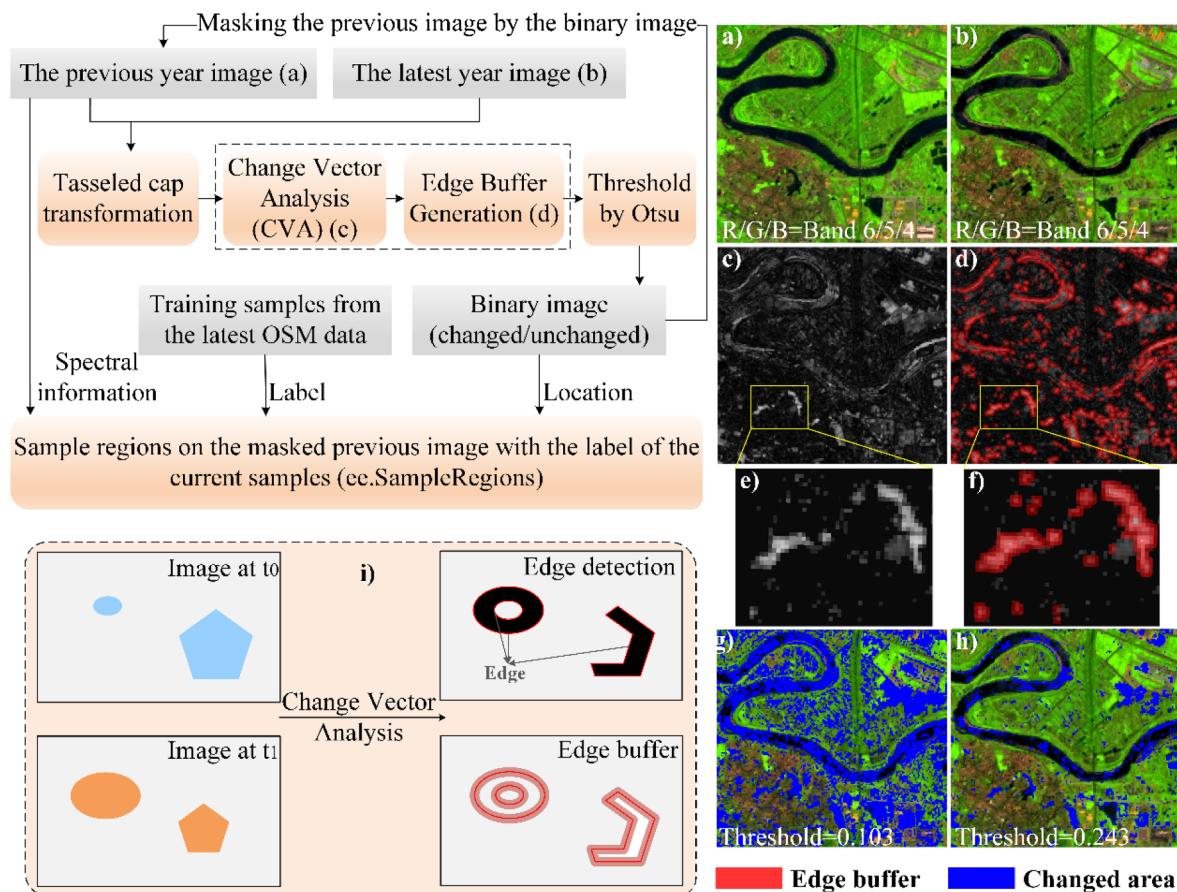
the samples were collected in 2017 and the classification result in this year has the highest certainty (as shown in Section 4.1.2); consequently, its temporal contextual information would be propagated to the other slices based on similar rules. Finally a spatial majority filtering was applied to the initial classification results to remove small urban and nonurban patches.

### 3.5. Validation and comparison

Validation of our resultant urban classifications was conducted in two aspects: (1) validation using ground truth samples; (2) comparison with the existing Landsat-based land-cover maps.

Due to the challenge of collecting reference data over the past thirty years and the availability of land-cover products, we collected reference data from multiple sources for seven epochs (2017, 2014, 2010, 2005, 2000, 1995, and 1990), including very high resolution (VHR) images from Google Earth (GE), human-impact land-cover points from the Geo-Wiki platform (Fritz et al., 2017), and the Global Geo-Referenced Field Photo Library (<http://www.eomf.ou.edu/photos/>). Since no fine-resolution imagery is available for the early periods, reference samples were also selected from Landsat image based urban land products using visual judgments.

We first used the stratified random sampling approach to acquire validation samples through visual interpretation on GE for the year 2017 (Fig. 7). The samples were within a stratification of five classes (urban, forest, cropland, bareland and water). Because urban land is the major concern of this study, the other four types of AOIs were combined as the non-urban category. For the Geo-Wiki and field photos data, we created areas of interest (AOIs) as square buffers of the points (100 m by 100 m) (Fig. 7c, d, e, and g)). Then, the labels of the AOIs were checked referring to the VHR images with the class occupying the area majority set as reference class, which provided a representation of reality (Schultz et al., 2017). Finally, a total of 2444 AOIs (283, 534 pixels) were generated for the validation of the resultant urban and non-urban map (Table 2). The AOIs were then expanded to other epochs as follows: for the years 2015, 1995, and 1990, GE, GHSL and



**Fig. 6.** The sample transfer procedures (upper-left). The data-processing procedures are shown for a small area as an example (right): (a) the false color composite ( $R/G/B = \text{Band } 6/5/4$ ) of the annual greenest surface reflectance pixel data in 2013 and (b) 2017; (c) the CVA result image, where the white pixels represent changed areas and the black ones represent unchanged areas; (d) overlapping the edge buffer on the CVA result image; (e) partial enlargement of figure (c); (f) partial enlargement of figure (d); (g) the detected changed area using global Otsu; (h) the detected changed area using the method shown in (i); (i) the change detection and edge buffer generation process.

the products from Liu et al. (2018) were used as references for samples checking; for the year 2010, GE, GlobeLand30 and the product from Liu et al. (2018) were used as references for samples checking; for the year 2005, GE and the product from Liu et al. (2018) were used as references for samples checking; for the year 2000, GE, GHSL, GlobeLand30, and the product from Liu et al. (2018) were used as references for samples checking. In each epoch, the produced temporally adjacent reference samples were loaded first, then the samples located on high quality GE images were visually checked to identify whether their types have been changed. For those changed samples, their labels will be reassigned or they will be slightly moved to ambient sites with more confidence (Grinand et al., 2013). For those samples located where no high quality GE images exist, only the samples located on both of the two or three products were labeled as “urban”, and the others were labeled as “non-urban”. Detailed information about the AOI numbers for different categories is shown in Table 2.

A confusion matrix including the overall accuracy (OA), user's accuracy (UA) and producer's accuracy (PA), kappa coefficient, commission error (CE) and omission error (OE) of the urban land map were calculated using the poststratification method described in Stehman and Foody (2019) to evaluate the accuracies of the results based on these validation samples.

In addition to validation, we compared our urban land classifications with the existing Landsat-based land-cover maps, including GlobeLand30 as well as the results from Liu et al. (2018). These high-resolution land-cover products still require further validation at the international level to determine if they are useful for different

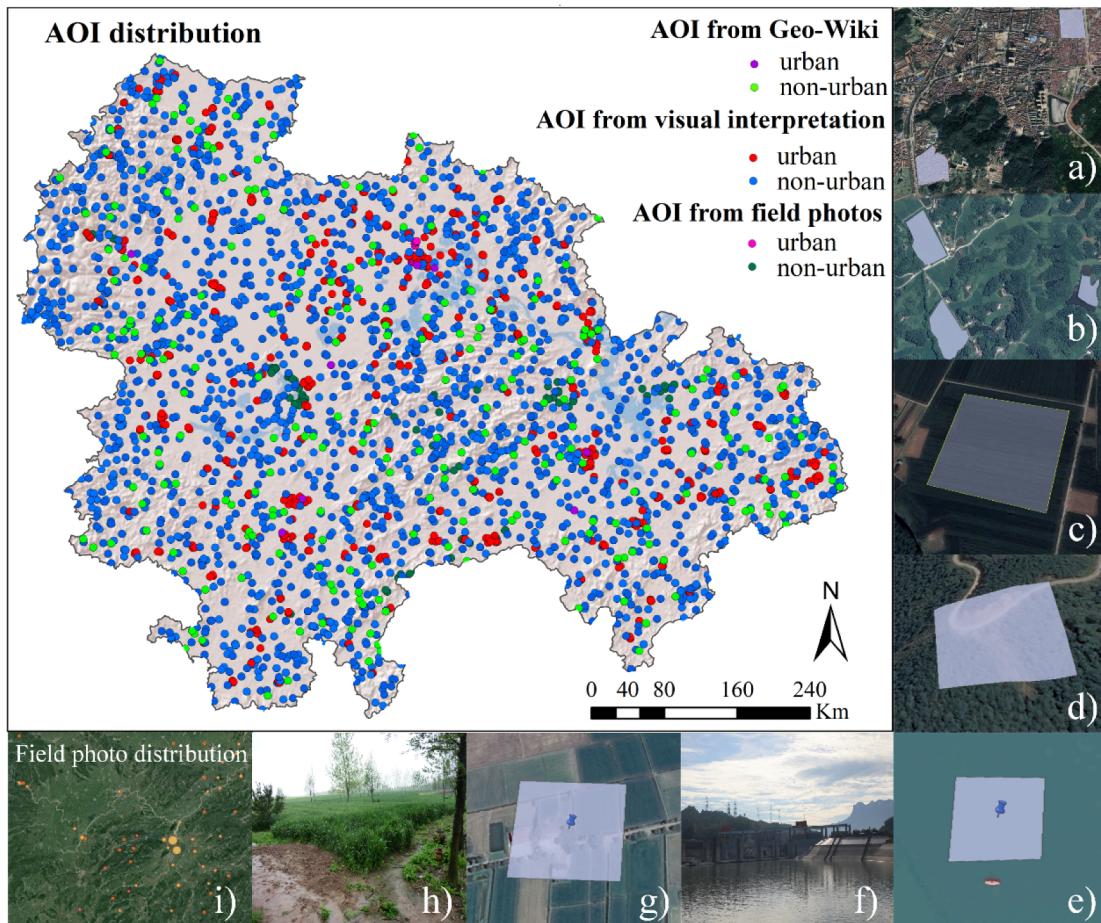
applications (Fonte et al., 2017).

## 4. Results

### 4.1. Reliability of the proposed urban land mapping framework

#### 4.1.1. Performance of the change detection scheme

We compared the results of the change detection scheme (i.e., generating the threshold using a canny edge buffer) (local Otsu) with that derived merely from the traditional Otsu threshold segmentation algorithm (global Otsu) on an example scene in the study area (path/row: 123/039) (Fig. 8). It is obvious that some unchanged regions were detected as changed regions, such as the Yangtze River, some lakes and even some of the green crop lands by using global Otsu. One reason for this error is that the spectral features of some ground objects may change over time. For example, the surface reflectance of water varies at different times as a result of sediment charge and aquatic organisms, etc. By contrast, by only considering the image features at the boundary region of the change-vector image, the modified Otsu threshold segmentation algorithm (local Otsu) can distinguish the true changes and false changes more effectively. The zoomed-in images enable a more intuitive comparison of the two approaches (the second, third, and fourth rows in Fig. 8). Region A, B, and C show complete change, partial change and no change, respectively, from 2013 to 2017. Both methods can detect the complete change in region A. However, for region B and C, there is an over-detection of change areas by the global Otsu. And the modified local Otsu performed better than the global Otsu.



**Fig. 7.** Spatial distribution of the validation AOIs in 2017. AOI examples from visual interpretation referring to VHR images: (a) urban land; (b) forest; AOI examples from Geo-Wiki sets referring to VHR images: (c) crop land; (d) forest; and AOI examples from field photos referring to VHR images: (e) water; (f) filed photo of (e); (g) cropland; and (h) field photo of (g); (i) the distribution of filed photos in MYRB.

**Table 2**

Number of AOIs collected for the validations of urban land maps during the seven epochs. The numbers inside the brackets are the validation pixels used.

Epochs	Urban	Non-urban	Total
1990	102(5506)	4166(564,265)	4268(569,771)
1995	149(7409)	4265(583,602)	4414(586,348)
2000	200(11,765)	4139(561,925)	4339(573,690)
2005	279(17,465)	4164(650,808)	4443(652,961)
2010	329(21,494)	4170(577,698)	4499(599,192)
2014	400(34,026)	4138(558,022)	4538(592,048)
2017	466(28,823)	4138(558,000)	4604(585,546)

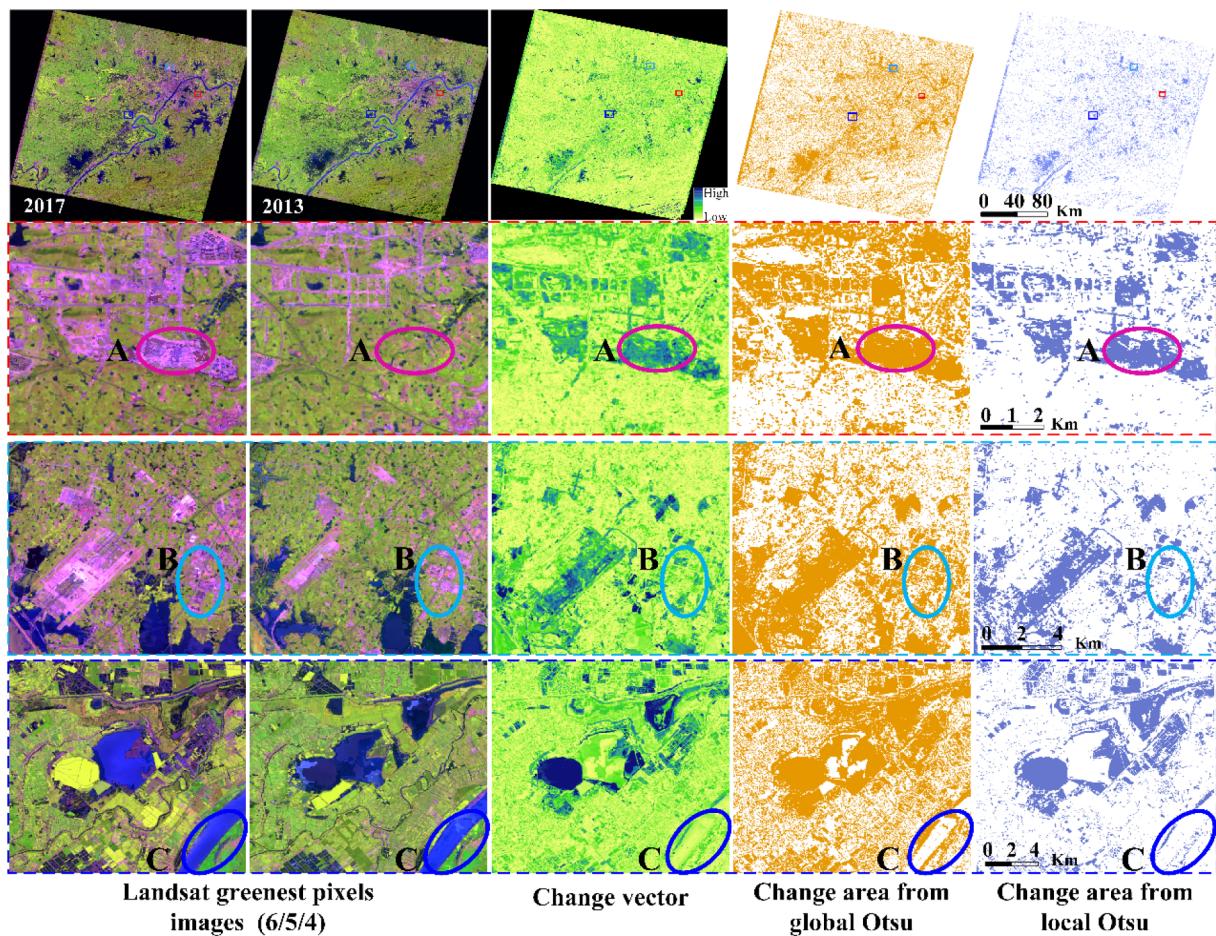
A further assessment of the accuracy of the two change detection methods was performed. The accuracy assessment of the change detection indicated good performance of our approach in both changed and non-changed areas for 1304 randomly collected samples during 2013–2017 based on GE images. The overall change-detection accuracies of the global Otsu and our approach are 76% and 88%, respectively. The quantitative results also show the existence of over-change detection of the non-changed area by the global Otsu method (200 non-changed points were detected as changed points compared with the local Otsu) (Table 3), indicating the efficiency of our method.

#### 4.1.2. Annual accuracies based on interpreted reference samples

Fig. 9 shows the annual accuracy assessments of the resultant urban land classifications both before and after the consistency correction in the last thirty years conducted by using the validation AOIs as described

in Section 3.5. The results indicated that the resultant urban land and non-urban land maps had high accuracies. As shown in Fig. 9, the original average OA and kappa coefficient of urban land can reach approximately 93% and 0.80, respectively. And even for the earlier years, the lowest OA and kappa coefficient (in 1993) can reach about 89% and 0.68, respectively. And they were improved up to 0.3% (96%) and 0.08 (0.88), respectively after consistency correction, and for the years 2010, 2016 and 2017, their accuracies are even better (above 98%). Although the original UA and PA for urban land are relatively low, they can averagely reach to 90% and 80%, respectively. Overall, the quality of the final obtained urban sequences is satisfactory. Therefore, we are confident about the use of this product for future practical applications in urban growth modeling or analysis.

We also find that the higher accuracies are mainly achieved in more recent years that are closer to the year in which the training samples were acquired. This phenomenon can be attributed to two reasons, on the one hand, the closer to the training samples acquiring year, the more certain the class of the samples are. On the other hand, there are more image scenes available in more recent years (Fig. 3) which ensures higher temporal frequency for observations (Roy et al., 2014) and further contributes to the classification accuracy. Nevertheless, it should be noted that we did not collect the AOIs at an annual frequency, and the accuracy assessments for five sequential years that two-year before and after the AOIs sampling year were done using the same validation samples. Moreover, there are missing data in the mapping results for the 1990s due to the shortage of high-resolution images for obtaining reference data. These problems impair the completeness and quality of the presented data product and cause a certain degree of



**Fig. 8.** Illustration of the proposed change detection schema. The first row shows the whole scene of the sample image for change detection and the detected results. The second, third and fourth rows are enlarged example areas in the red, blue and green frames.

**Table 3**

Accuracy assessment of change detection for the example image scene. (OA: overall accuracy; UA: user's accuracy; PA: producer's accuracy.)

		Reference		Sum	UA (%)
		Changed	Non-changed		
Global Otsu Map	Changed	378	280	658	57.45
	Non-changed	26	620	646	95.98
	Sum	404	900	1304	—
	PA	93.56	68.89	—	—
	OA	76.53	—	Kappa	0.53
Local Otsu Map	Changed	338	80	418	80.86
	Non-changed	66	820	886	92.55
	Sum	404	900	1304	—
	PA (%)	83.66	91.11	—	—
	OA (%)	88.80	—	Kappa	0.74

uncertainty.

#### 4.1.3. Comparison with other 30 m products

Statistics based on the spatial differences between the two existing products and our resultant maps in 2000 and 2010 are also calculated and compared (Fig. 10). The higher value of UU in G2000 and G2010 indicate that our resultant maps have more urban pixels consistent with GlobleLand30 than the results from Liu et al. (2018). As to the proportion of NU, GlobleLand30 shows less inconsistent with our product than that of Liu et al. (2018). However, the urban land difference between GlobleLand 30 and our product are larger than the differences with the Liu et al. (2018) product with higher proportion of UN.

Overall, GlobleLand30 demonstrated more inconsistencies with our products than that of Liu et al. (2018).

Fig. 11 shows several random selected sites from original Landsat images and the three 30 m urban land products in 2000 and 2010. A visual comparison with the referenced Landsat images shows that our mapping results and GlobleLand30 provide an accurate representation of urban land, but our results provide finer spatial details (e.g., row/column: 2/2 vs. 4/2 and 5/6 vs. 8/6). The products from Liu et al. (2018) also provide relatively fine spatial details of urban land, which can also be illustrated from Fig. 10, but always with an over estimation of the urban land (e.g., row/column: 6/2 vs. 7/2 and 6/6 vs. 7/6).

We also used the same validation samples to quantitatively assess the free acquired urban maps in the two epochs and the accuracies of Liu et al. (2018) were also lower than the results of this study and GlobleLand30 (Table 4). The result is also consistent with the above visual comparison results, which indicated that our results are acceptable and can be used for further analysis.

#### 4.2. Long time series urban land change mapping in MYRB

We finally mapped the expansion of urban land from 1987 to 2017 (Fig. 12). The map showed that the urban land continued to increase during the last thirty years and expanded from urban cores to suburban areas. Some “hotspots” were clearly delineated, including Wuhan, Changshang, and Nanchang, which are the capital cities of Hubei, Hunan and Jiangxi Provinces, respectively. The zoomed-in maps show that most cities are passed through by rivers, and thus balancing the relationship between urban development and water quality should also be paid attention to. Moreover, from the scatter point figure showing

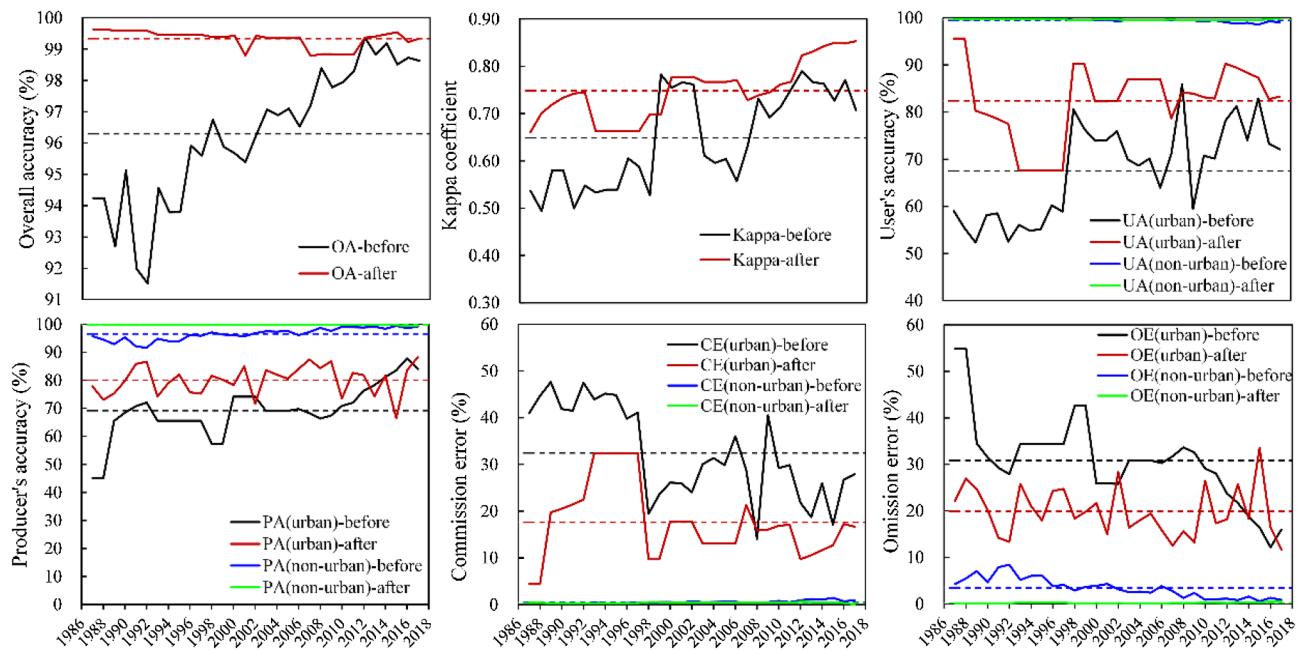


Fig. 9. Annual accuracy assessment of resultant urban land during 1987–2017. (-before: before consistency checking; -after: after consistency checking).

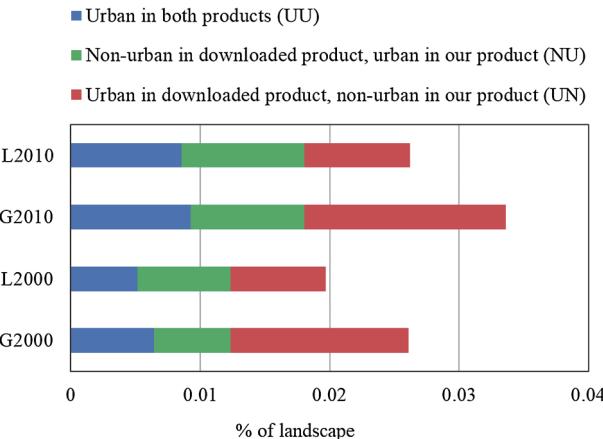


Fig. 10. Proportion of urban land difference maps between: (1) GlobleLand30 and our product in 2000 (G2000); (2) Liu et al. and our product in 2000 (L2000); (3) GlobleLand30 and our product in 2010 (G2010); and (4) Liu et al. and our product in 2010 (L2010).

urban area by year, we can draw some preliminary conclusions:

Urban land in MYRB showed a basic tendency of exponential growth. Specifically, in the period of 1987 to 2004, the MYRB did not experience a significant urban expansion. Urban land increased from 1445.9 km<sup>2</sup> in 1987 to 3219.2 km<sup>2</sup> in 2004, with an average annual growth rate of 4.82%. From 2005 to 2008, the MYRB experienced a rapid urban growth period with a sharp increase in urban area from 3462.2 km<sup>2</sup> in 2005 to 5357.8 km<sup>2</sup> in 2008, corresponding to an average annual growth rate of 15.67%. This was the period in which the “Rise of Central China Plan” was proposed by the national central government, which may have promoted the urban development of this region. For the period 2009–2012, the MYRB continued its fast urbanization but witnessed a slow-down compared with the former period, with an average annual growth rate of 5.29%. The period of 2013–2017 included another period of rapid urban growth and the urban cluster started to form. The urban area increased from 6747.8 km<sup>2</sup> to 8744.9 km<sup>2</sup> at an average annual growth rate of 6.69%.

## 5. Discussion

### 5.1. Feasibility of the proposed workflow and future implications

Higher-resolution (~30 m) land-cover characterization and monitoring enables the recognition of land change at the scale of most human activities and offers increased flexibility of environmental model parameterization needed for global change studies (Chen et al., 2015; Giri et al., 2013). However, a number of challenges must be overcome before these data sets can be produced, including the unavailability of consistent global coverage of satellite data, the sheer volume of data, the unavailability of timely and accurate training and validation data, difficulties in preparing image mosaics, and high-performance computation requirements (Giri et al., 2013). This study indicated the feasibility and reliability of yearly semi-automatic 30 m mapping of urban land at the regional scale in MYRB based on the acceptable accuracies of the resultant products (Section 4.1.2 and Section 4.1.3). The major contributions of this study are as follows:

Methodologically, this study provides a new solution for mapping long-term and large-scale urban land semi-automatically by integrating remote sensing images with crowdsourced in situ data, which frees human beings from duplicative work. It is a difficult task to acquire training samples for a large study area let alone for sequent years. The proposed framework investigates a method that incorporates the up-to-date OSM data for latest year’s image classification and transfers the unchanged land use information to historical years to realize automatic image classification without manual sampling year by year. On the other hand, benefiting from the large amounts of archived remote sensing resources, various designed image processing functions as well as high computation efficiency supported by millions of servers around the world (Huang et al., 2017) of GEE, this solution frees human beings from complex data-processing tasks such as image mosaicking, clipping, and compositing. Because the GEE platform and the OSM database are freely available to the public, the approach does not incur financial costs or any time spent exploring land through fieldwork or processing images one by one. Extendedly, owing to the openness of OSM data and GEE, the proposed framework has the potential to be applied and extended to other regions, even at the global scale for LSITS urban land mapping by only changing the geographic coordinates. It works especially for areas where little OSM data are available in the early years as

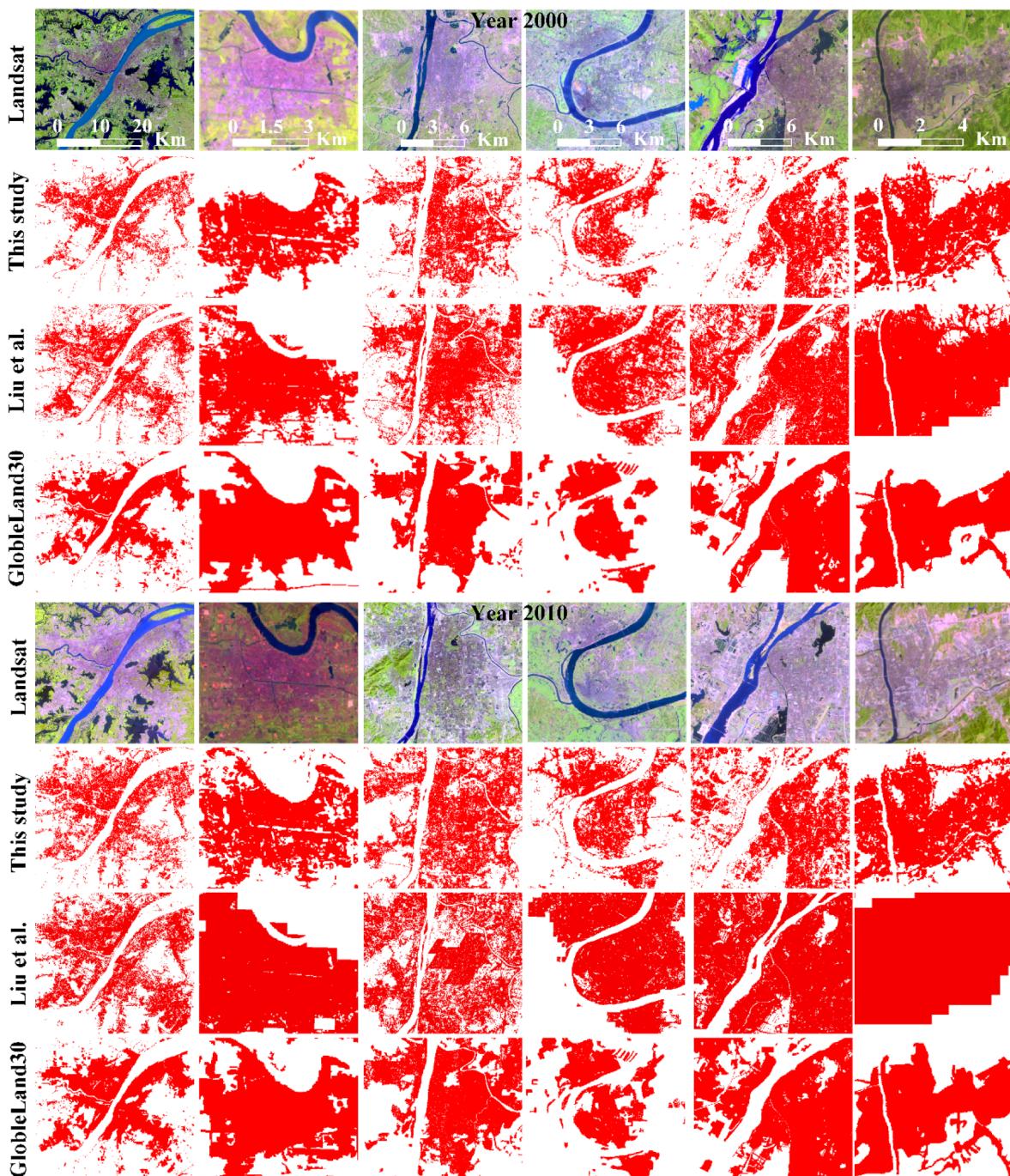


Fig. 11. Comparing the urban land classifications obtained by this study and other available 30 m products.

**Table 4**

Accuracy assessment on overall accuracy and Kappa coefficients of the products from this study, Liu et al. and GlobleLand30 based on the AOIs in the common spatial extent.

	Overall accuracy (%)		Kappa coefficient	
	2000	2010	2000	2010
This study	96.53	97.08	0.90	0.92
Liu et al. (2018)	95.63	95.70	0.87	0.88
GlobleLand30	98.46	98.66	0.95	0.96

our case. Moreover, we used the Otsu approach for change information extraction in the sample transferring schema owing to its widespread usage and easy implementation in GEE. Further studies can replace this method with other techniques for change detection, showing the flexibility of this framework. Resultantly, a continuous 30 m urban land cover product for the MYRB was produced from this study, providing supportive spatial information for trend analysis of urbanization as well as for predication of future urban development in the study area.

However, there were a few limitations need to be improved further. (1) The OSM data is inconsistent in spatial coverage due to the presence of most of human activities are in urban areas (Fonte et al., 2015). Thus, some LULC classes such as wetland, bare land, and farm land are not frequently uploaded, resulting in a number imbalance (Johnson and

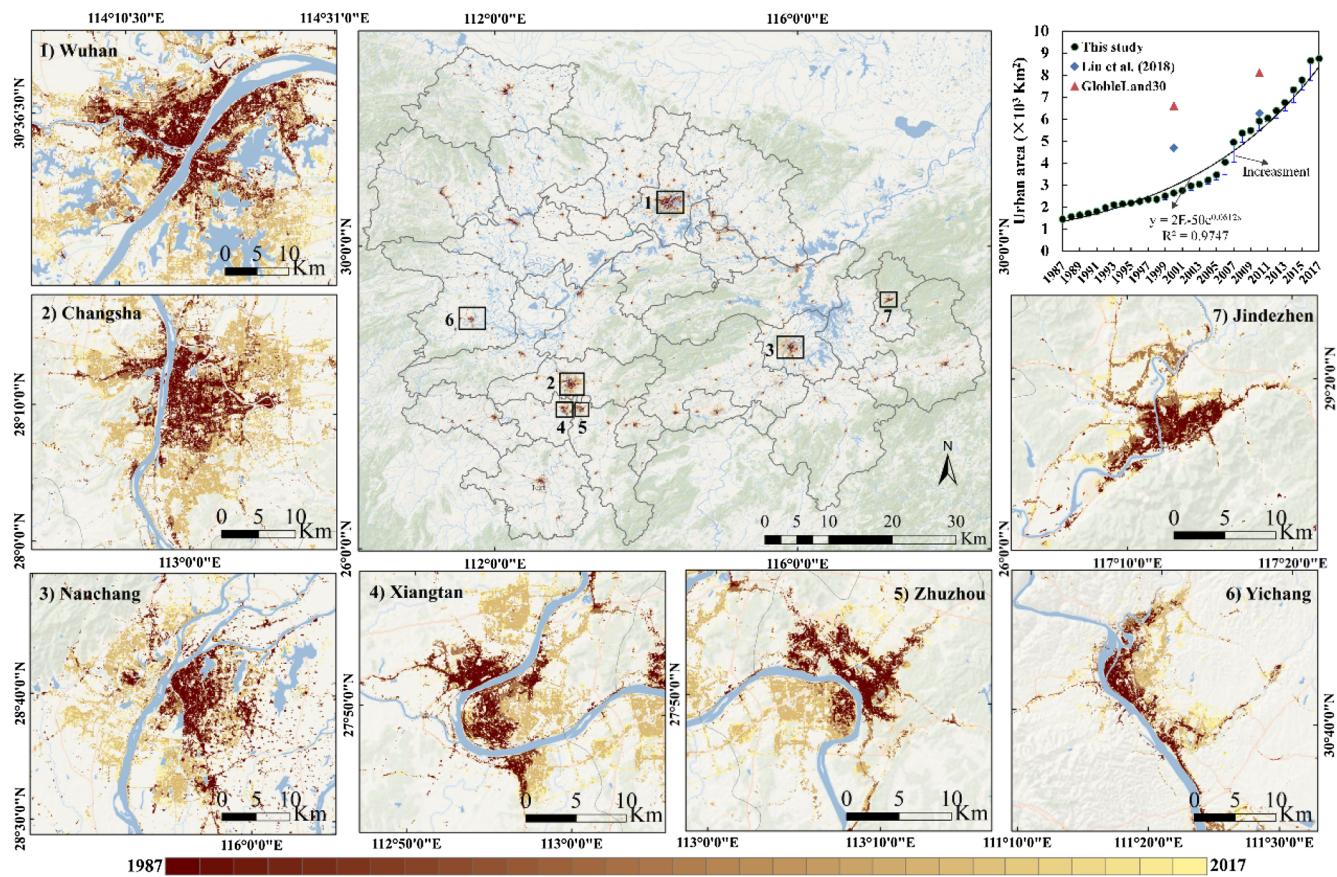


Fig. 12. The resultant temporal urban expansion map in the middle Yangtze River basin (1987–2017).

Iizuka, 2016) among different classes, making this approach inappropriate for mapping more than two classes or for the minority LULC mapping. In the future, integrating other sources of CGD such as Flickr photos and GeoNames (Li and Gong 2016) could be a practical task. We also intend to improve the mapping framework by employing an algorithm that handles the class imbalance, such as the synthetic minority over-sampling technique (SMOTE) (Chawla et al., 2002) based on GEE API. We believe better and more consistent increased urban areas can be identified using this solution. (2) The dataset archived on GEE are mainly Earth-observation imagery (Gorelick et al., 2017). The training samples generation is still conducted manually in this study. In the future, if some data catalog and GEE API related to CGD could be integrated on the platform, it will be possible for us to realize a more automatic framework for LSLTS mapping. Moreover, as the limitation of an interactive interface for post classification, the classified results often need to be downloaded from the cloud platform for further manual processing. We highly urge that the GE annual VHR images could be added to GEE so that users can set the background image by themselves and further annual validation samples could also be collected on the cloud. (3) Finally, we plan to create a web service based on the GEE platform to display and update our produced products as well as provide different types of urban expansion analysis results. Such a service would offer both the government and the public a more straightforward means of understanding their living space and how to make better use of that space.

## 5.2. Urban expansion in MYRB from 1987 to 2017

### 5.2.1. Pattern of urban growth in MYRB

Based on the obtained mapping results, further investigations were performed to evaluate the spatial evolution of urban land in MYRB. Three urban expansion types (i.e., infilling, edge-expansion and

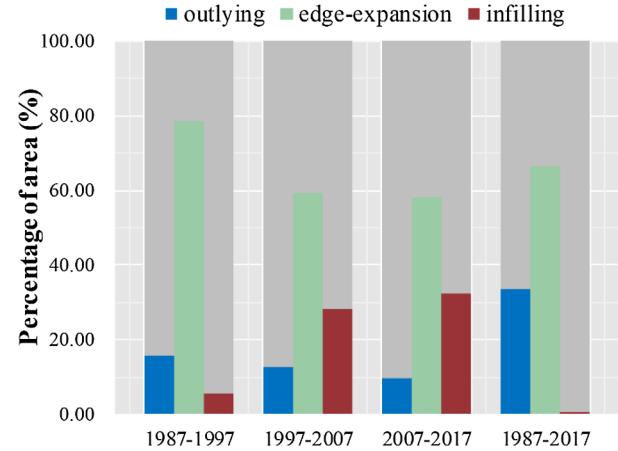
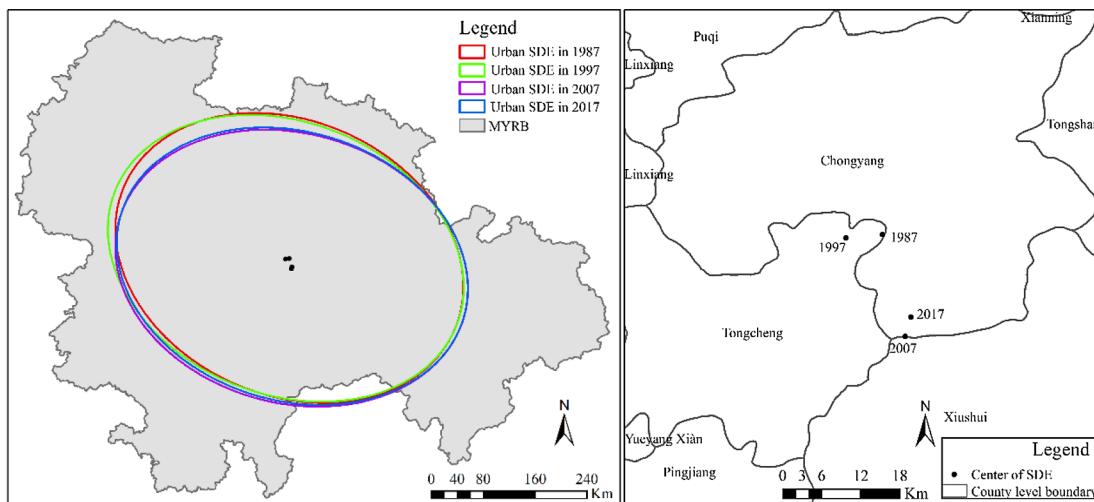


Fig. 13. Composition of urban growth types relative to the total expanded area in MYRB during 1987–2017.

outlying) (Liu et al., 2010) were identified for the newly developed areas (Fig. 13). The dominant urban growth type (edge-expansion), which is in line with our previous study (Liu and Chen, 2017), has experienced a decrease of approximately 21% from 79% during 1987–1997 to 58% during 2007–2017. The minor urban growth type has shifted from infilling to outlying during the last thirty years, indicating that urban development has become increasingly compact (Wu et al., 2015) over time. By contrast, when viewed with a 30-year time interval, edge-expansion and outlying growth nearly dominated the whole expansion, indicating the incredible urbanization process in MYRB in the last thirty years. The reason for such difference is that the calculation of urban growth type is based on the expanded urban land.



**Fig. 14.** Change in SDE and its gravity center for the urban land in MYRB.

When we analyzed the urban expansion types for a small time interval, an urban cluster may be identified as infilling growth. However, when analyzed from a larger time interval, it will be identified as outlying growth.

#### 5.2.2. Spatial evolution among cities in MYRB

Spatially, the distribution of urban land in MYRB generally shows a “northwest-southeast” pattern (Fig. 14). The change in the long axis of the standard deviation ellipse (SDE) (Lefever 1926) is much stronger than that of the short axis, indicating that the main force driving the spatial evolution of urban expansion in MYRB may come from the “northwest-southeast” direction. Specifically, the urbanization trajectory of MYRB is characterized by a movement to the northwest and then to the southeast. During 1997 and 2007, the SDE of urban land shrank significantly in the northwestern part, indicating that in this period the urban expansion mainly happened in the southeast part, such as Nanchang, Jiujiang, Shangrao. By contrast, cities in the northwest part, such as Yichang and Xiangyang, experienced a slowdown in their urbanization progress.

The spatial statistic methods, including Anselin local Moran's I (Anselin, 1995) and Getis-Ord G\* (Getis and Ord, 1992) were further used to reveal the spatial distribution pattern of urban expansion intensity (UEI) among the 31 cities within MYRB based on the resulting urban land map through ArcGIS. These four indices measure the degree of global and local autocorrelation and concentration of the attribute of interest, respectively.

Fig. 15 displays the significant locations color coded by type of spatial autocorrelation as well as the “hot-spot” and “cold-spot” areas of UEI among the 31 cities from 1987 to 2017. The result of local Moran's I is relatively consist with that of local Getis-Ord as a whole. The locations with low UEI were unchanged and mainly located in the western part of MYRB urban agglomeration in all the three periods in both results. The reason could be attributed, to some degree, to the fact that there are less highly developed cities in the west as well as due to the constraint of terrain. However, in the central and southeastern part of MYRB, some outliers, which are shown as hot spots by the local Getis-Ord method, can be detected by local Moran's I. Only a few cities such as Wuhan and its one or two neighbors, Changsha, and Nanchang are true hot spots with higher UEI. Whereas most of their neighbors, in fact, witnessed low UEI. This results demonstrate that the link between cities in MYRB is still very weak and the core cities currently have not effected too much in driving their neighbors' urban development.

From the above analysis, two primary conclusions can be drawn: 1) during the past thirty years, spatial autocorrelation and a weak high-value spatial clustering existed in the intensity of urban expansion in

MYRB in general; 2) connections among cities were relatively weak and great spatial disparity still existed in the process of urban development in MYRB. A key tenet for sustainable development and smart growth is promoting equitable urbanization and land development and mitigating land use conflicts (Wei, 2016). Strengthening social and economic cooperation between cities will help strengthen the interactions between cities, form regional development synergies, and accelerate the process of urbanization.

## 6. Conclusion

Currently, Landsat is the only remote sensing data source to track continuous regional land use change back to the 1980s at a 30 m spatial resolution; however, the use of time-series Landsat imagery in large-scale LULC studies faces a series of challenges, including large computing cost and labor-intensive images and samples collection and processing. By taking advantages of current open-sourced satellite and in situ data, the high-performance GEE computing platform as well as the large volume images archived on it, we proposed a semi-automatic framework to map urban land and applied it to the 317,000 km<sup>2</sup> middle Yangtze River basin from 1987 to 2017. First, random training samples were elaborated from collected latest OSM land use data after a manual topological conflict processing to exclude ambiguous polygons with mixed tags. Second, the samples of latest year were uploaded to the cloud platform for further image classification. Training samples for each historical year were acquired based on a schema that detected the unchanged area which was further used as a mask for filtering the latest year training samples. Third, the samples and the annual greenest pixel composites were input to both a RF and a CART classifiers to acquire urban land map series. For each previous year, urban land map will be automatically produced without further manual operations. To our knowledge, this study is the first to use the crowdsourced OSM data to map large-scale 30 m urban land annually. The accomplishment of it was largely attributed to the utility of cloud-computing technology, thus showcasing the potentials and prospects of emerging parallel/cloud-computing platforms. In addition, this approach can considerably reduce the cost of time and labor compared with traditional manual sample collection method.

Based on independent quantitative accuracy assessment as well as qualitative comparison with current 30 m land-cover products, the OA of the classification results can reach to 93% and 96% before and after a consistency correction, respectively. And the mapping results of this research show more spatial details compared with the other two existed products (i.e., GlobleLand30 and the products from Liu et al. (2018)), and remained similar even higher accuracy than the other two

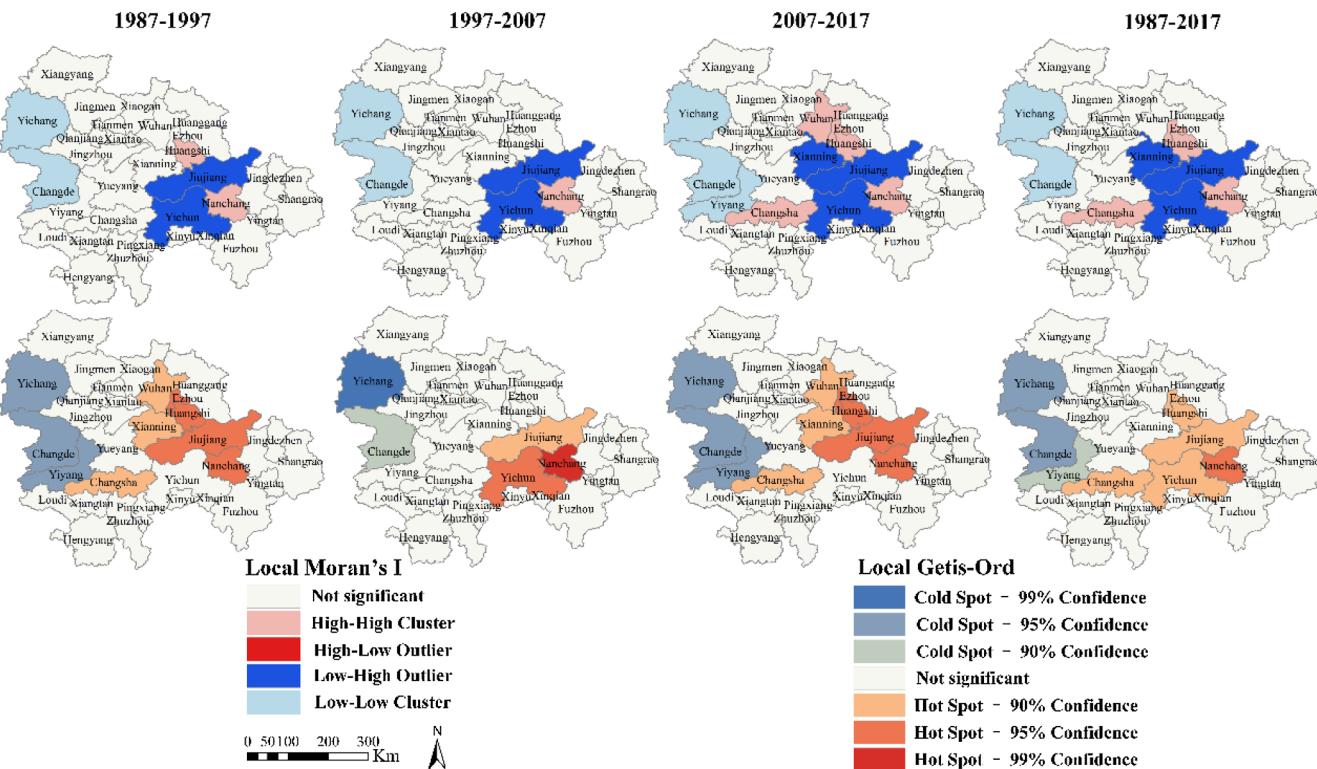


Fig. 15. Spatial autocorrelation and hot spots mapping of UEI in MYRB.

products. Thus, the research question is now answered and the OSM can be a potential data source for mapping urban landscape. The produced mapping result inventories spatially explicit urban extents in the 31 target cities across the MYRB at a one-year temporal frequency during the study period. The results clearly demonstrated that the urban agglomeration in MYRB experienced accelerated growth in the past thirty years with urban land increased from 1445.9 km<sup>2</sup> in 1987 to 8744.9 km<sup>2</sup> in 2017. However, most of its urban land increase still concentrated on several core cities, indicating that future policies related to urban planning need to focus more on balancing urbanization in MYRB.

#### 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

The authors would like to thank all the open data organizations including the OSM community, Geo-Wiki community and the Earth Observation and Modeling Facility for providing such diverse data for us; the GEE developers and the working group of GEE who helped us a lot during this study; Dr. Brian A. Johnson who provided constructive suggestions during the design of this research. We would also like to express our respects and gratitude to the anonymous reviewers and editors for their professional comments and suggestions.

This research was supported by the National Key R&D Program of China (no. 2018YFB2100500), National Natural Science Foundation of China program (no. 41890822) and Creative Research Groups of Natural Science Foundation of Hubei Province of China (no. 2016CFA003).

#### References

- Alonso, A., Munoz-Carpena, R., Kennedy, R.E., Murcia, C., 2016. Wetland landscape spatio-temporal degradation dynamics using the new google earth engine cloud-based platform: opportunities for non-specialists in remote sensing. *T Asabe* 59, 1333–1344.
- Anselin, L., 1995. Local indicators of spatial association—LISA. *Geogr. Anal.* 27 (2), 93–115.
- Arnold, C.L., Gibbons, C.J., 1996. Impervious surface coverage – the emergence of a key environmental indicator. *J. Am. Plann. Assoc.* 62, 243–258.
- Arsanjani, J.J., Helbich, M., Bakillah, M., 2013a. Exploring volunteered geographic information to ease land use mapping of an urban landscape. *ISPRS – Int. Arch. Photogramm., Rem. Sens. Spatial Inf. Sci.* XL-4/W1 51–55.
- Arsanjani, J.J., Helbich, M., Bakillah, M., Hagenauer, J., Zipf, A., 2013b. Toward mapping land-use patterns from volunteered geographic information. *Int. J. Geogr. Inf. Sci.* 27, 2264–2278.
- Ban, Y.F., Gong, P., Gini, C., 2015. Global land cover mapping using Earth observation satellite data: recent progresses and challenges. *Isprs J. Photogramm.* 103, 1–6.
- Bereitschaft, B., Debbage, K., 2013. Urban form, air pollution, and CO<sub>2</sub> emissions in large U.S. Metropolitan areas. *Prof. Geogr.* 65, 612–635.
- Boyd, D.S., Jackson, B., Wardlaw, J., Foody, G.M., Marsh, S., Bales, K., 2018. Slavery from Space: demonstrating the role for satellite remote sensing to inform evidence-based action related to UN SDG number 8. *Isprs J. Photogramm.* 142, 380–388.
- Canny, J., 1986. A computational approach to edge-detection. *IEEE T Pattern Anal.* 8, 679–698.
- Cao, Q., Yu, D.Y., Georgescu, M., Wu, J.G., Wang, W., 2018. Impacts of future urban expansion on summer climate and heat-related human health in eastern China. *Environ. Int.* 112, 134–146.
- Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* 16, 321–357.
- Chen, J., Chen, J., Liao, A.P., Cao, X., Chen, L.J., Chen, X.H., He, C.Y., Han, G., Peng, S., Lu, M., Zhang, W.W., Tong, X.H., Mills, J., 2015. Global land cover mapping at 30 m resolution: a POK-based operational approach. *Isprs J. Photogramm.* 103, 7–27.
- Chen, J., Chen, X.H., Cui, X.H., Chen, J., 2011. Change vector analysis in posterior probability space: a new method for land cover change detection. *IEEE Geosci. Remote S* 8, 317–321.
- Demir, B., Bovolo, F., Bruzzone, L., 2012. Detection of land-cover transitions in multi-temporal remote sensing images with active-learning-based compound classification. *IEEE T Geosci. Remote* 50, 1930–1941.
- Dong, J.W., Xiao, X.M., Kou, W.L., Qin, Y.W., Zhang, G.L., Li, L., Jin, C., Zhou, Y.T., Wang, J., Biradar, C., Liu, J.Y., Moore, B., 2015. Tracking the dynamics of paddy rice planting area in 1986–2010 through time series Landsat images and phenology-based algorithms. *Remote Sens. Environ.* 160, 99–113.
- Dong, J.W., Xiao, X.M., Menarguez, M.A., Zhang, G.L., Qin, Y.W., Thau, D., Biradar, C., Moore, B., 2016. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. *Remote Sens.*

- Environ. 185, 142–154.
- Eshtawi, T., Evers, M., Tischbein, B., 2016. Quantifying the impact of urban area expansion on groundwater recharge and surface runoff. *Hydrolog. Sci. J.* 61, 826–843.
- Feyisa, G.L., Meilby, H., Jenerette, G.D., Pauillet, S., 2016. Locally optimized separability enhancement indices for urban land cover mapping: Exploring thermal environmental consequences of rapid urbanization in Addis Ababa, Ethiopia. *Remote Sens. Environ.* 175, 14–31.
- Fonte, C.C., Bastin, L., See, L., Foody, G., Lupia, F., 2015. Usability of VGI for validation of land cover maps. *Int. J. Geogr. Inf. Sci.* 29, 1269–1291.
- Fonte, C.C., Minghini, M., Patriarca, J., Antoniou, V., See, L., Skopeliti, A., 2017. Generating Up-to-Date and Detailed Land Use and Land Cover Maps Using OpenStreetMap and GlobeLand30. *Isprs Int Geo-Inf. 6*.
- Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X.M., 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* 114, 168–182.
- Fritz, S., See, L., Perger, C., McCallum, I., Schill, C., Schepaschenko, D., Duerauer, M., Karner, M., Dresel, C., Laso-Bayas, J.C., Lesiv, M., Moorthy, I., Salk, C.F., Danylo, O., Sturm, T., Albrecht, F., You, L., Kraxner, F., Obersteiner, M., 2017. A global dataset of crowdsourced land cover and land use reference data. *Sci. Data 4*.
- Getis, A., Ord, J.K., 1992. The analysis of spatial association by use of distance statistics. *Geogr. Anal.* 24, 189–206.
- Giri, C., Pengra, B., Long, J., Loveland, T.R., 2013. Next generation of global land cover characterization, mapping, and monitoring. *Int. J. Appl. Earth Obs.* 25, 30–37.
- Grinand, C., Rakotomalala, F., Gond, V., Vaudry, R., Bernoux, M., Vieilledent, G., 2013. Estimating deforestation in tropical humid and dry forests in Madagascar from 2000 to 2010 using multi-date Landsat satellite images and the random forests classifier. *Remote Sens. Environ.* 139, 68–80.
- Goldblatt, R., Stuhlmacher, M.F., Tellman, B., Clinton, N., Hanson, G., Georgescu, M., Wang, C.Y., Serrano-Candela, F., Khandelwal, A.K., Cheng, W.H., Balling, R.C., 2018. Using Landsat and nighttime lights for supervised pixel-based image classification of urban land cover. *Remote Sens. Environ.* 205, 253–275.
- Goldblatt, R., You, W., Hanson, G., Khandelwal, A.K., 2016. Detecting the boundaries of urban areas in India: a dataset for pixel-based image classification in Google Earth engine. *Remote Sens.* 8.
- Gomez, C., White, J.C., Wulder, M.A., 2016. Optical remotely sensed time series data for land cover classification: a review. *Isprs J. Photogramm.* 116, 55–72.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27.
- Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. *Science* 342, 850–853.
- He, Y.Q., Lee, E., Warner, T.A., 2017. A time series of annual land use and land cover maps of China from 1982 to 2013 generated using AVHRR GIMMS NDVI3g data. *Remote Sens. Environ.* 199, 201–217.
- Huang, H.B., Chen, Y.L., Clinton, N., Wang, J., Wang, X.Y., Liu, C.X., Gong, P., Yang, J., Bai, Y.Q., Zheng, Y.M., Zhu, Z.L., 2017. Mapping major land cover dynamics in Beijing using all Landsat images in Google Earth Engine. *Remote Sens. Environ.* 202, 166–176.
- 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.
- Johnson, B.A., Iizuka, K., Bragaas, 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. *Comput. Environ. Urban.* 64, 184–193.
- LeFever, D.W., 1926. Measuring geographic concentration by means of the standard deviational ellipse. *Am. J. Sociol.* 32, 88–94.
- Li, L., Li, X., Zhang, Y., Wang, L., Ying, G.W., 2016. Change detection for high-resolution remote sensing imagery using object-oriented change vector analysis method. *IEEE International Geoscience And Remote Sensing Symposium (Igars)* 2016, 2873–2876.
- Li, X.C., Gong, P., 2016. An “exclusion-inclusion” framework for extracting human settlements in rapidly developing regions of China from Landsat images. *Remote Sens. Environ.* 186, 286–296.
- Li, X.C., Gong, P., Liang, L., 2015. A 30-year (1984–2013) record of annual urban dynamics of Beijing City derived from Landsat data. *Remote Sens. Environ.* 166, 78–90.
- Liu, D.D., Chen, N.C., 2017. Satellite Monitoring of Urban Land Change in the Middle Yangtze River Basin Urban Agglomeration, China between 2000 and 2016. *Remote Sens. 9*.
- Liu, X.P., Hu, G.H., Chen, Y.M., Li, X., Xu, X.C., Li, S.Y., Pei, F.S., Wang, S.J., 2018. High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform. *Remote Sens. Environ.* 209, 227–239.
- Liu, X.P., Li, X., Chen, Y.M., Tan, Z.Z., Li, S.Y., Ai, B., 2010. A new landscape index for quantifying urban expansion using multi-temporal remotely sensed data. *Landscape Ecol.* 25, 671–682.
- Liu, Z.F., He, C.Y., Zhang, Q.F., Huang, Q.X., Yang, Y., 2012. Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008. *Landsc. Urban Plan.* 106, 62–72.
- Miyazaki, H., Shibasaki, R., Nagai, M., 2016. An automated method for time-series human settlement mapping using landsat data and existing land cover maps. *IEEE International Geoscience And Remote Sensing Symposium (Igars)* 2016, 1784–1787.
- Patel, N.N., Angiuli, E., Gamba, P., Gaughan, A., Lisini, G., Stevens, F.R., Tatem, A.J., Trianni, G., 2015. Multitemporal settlement and population mapping from Landsat using Google Earth Engine. *Int. J. Appl. Earth Obs.* 35, 199–208.
- Puissant, A., Hirsch, J., Weber, C., 2005. The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *Int. J. Remote Sens.* 26, 733–745.
- Roy, D.P., Wulder, M.A., Loveland, T.R., Woodcock, C.E., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., Scambos, T., Schaaf, C.B., Schott, J.R., Sheng, Y., Vermote, E.F., Belward, A.S., Bindschadler, R., Cohen, W.B., Gao, F., Hippel, J.D., Hostert, P., Huntington, J., Justice, C.O., Kilic, A., Kovalsky, V., Lee, Z.P., Lymbumer, L., Masek, J.G., McCorkel, J., Shuai, Y., Trezza, R., Vogelmann, J., Wynne, R.H., Zhu, Z., 2014. Landsat-8: science and product vision for terrestrial global change research. *Remote Sens. Environ.* 145, 154–172.
- Schultz, M., Voss, J., Auer, M., Carter, Sarah, Zipf, A., 2017. Open land cover from OpenStreetMap and remote sensing. *Int. J. Appl. Earth Obs.* 63, 206–213.
- See, L., Schepaschenko, D., Lesiv, M., McCallum, I., Fritz, S., Comber, A., Perger, C., Schill, C., Zhao, Y., Mau, V., Siraj, M.A., Albrecht, F., Cipriani, A., Vakolyuk, M., Garcia, A., Rabia, A.H., Singha, K., Marcarini, A.A., Kattenborn, T., Hazarika, R., Schepaschenko, M., van der Velde, M., Kraxner, F., Obersteiner, M., 2015. Building a hybrid land cover map with crowdsourcing and geographically weighted regression. *Isprs J. Photogramm.* 103, 48–56.
- Seto, K.C., Fragkias, M., Guneralp, B., Reilly, M.K., 2011. A meta-analysis of global urban land expansion. *Plos One 6*.
- Sexton, J.O., Song, X.P., Huang, C.Q., Channan, S., Baker, M.E., Townshend, J.R., 2014. Urban growth of the Washington, D.C.-Baltimore, MD metropolitan region from 1984 to 2010 by annual, landsat-based estimates of impervious cover. *Remote Sens. Environ.* 155, 379.
- Sexton, J.O., Urban, D.L., Donohue, M.J., Song, C.H., 2013. Long-term land cover dynamics by multi-temporal classification across the Landsat-5 record. *Remote Sens. Environ.* 128, 246–258.
- Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., Skakun, S., 2017. Exploring Google Earth Engine platform for big data processing: classification of multi-temporal satellite imagery for crop mapping. *Front. Earth Sci.* 5, 1–10.
- 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 Sens. Environ.* 175, 1–13.
- Stehman, S.V., Foody, G.M., 2019. Key issues in rigorous accuracy assessment of land cover products. *Remote Sens. Environ.* 231, 111199.
- Tsendbazarr, N.E., de Bruin, S., Herold, M., 2015. Assessing global land cover reference datasets for different user communities. *Isprs J. Photogramm.* 103, 93–114.
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150.
- Wang, J., Zhao, Y.Y., Li, C.C., Yu, L., Liu, D.S., Gong, P., 2015. Mapping global land cover in 2001 and 2010 with spatial-temporal consistency at 250 m resolution. *Isprs J. Photogramm.* 103, 38–47.
- Wang, L., Li, C.C., Ying, Q., Cheng, X., Wang, X.Y., Li, X.Y., Hu, L.Y., Liang, L., Yu, L., Huang, H.B., Gong, P., 2012. China's urban expansion from 1990 to 2010 determined with satellite remote sensing. *Chinese Sci. Bull.* 57, 2802–2812.
- Wei, Y.D., 2016. Towards Equitable and Sustainable Urban Space: Introduction to Special Issue on “Urban Land and Sustainable Development”. *Sustainability*-Basel 8.
- Wei, Y.D., Ewing, R., 2018. Urban expansion, sprawl and inequality. *Landsc. Urban Plan.* 177, 259–265.
- 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., 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.
- Wu, W.J., Zhao, S.Q., Zhu, C., Jiang, J.L., 2015. A comparative study of urban expansion in Beijing, Tianjin and Shijiazhuang over the past three decades. *Landsc. Urban Plan.* 134, 93–106.
- Xie, Y.H., Weng, Q.H., 2017. Spatiotemporally enhancing time-series DMSP/OLS nighttime light imagery for assessing large-scale urban dynamics. *Isprs J. Photogramm.* 128, 1–15.
- Xiong, J., Thenkabail, P.S., Gumma, M.K., Teluguntla, P., Poehnelt, J., Congalton, R.G., Yadav, K., Thau, D., 2017. Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. *Isprs J. Photogramm.* 126, 225–244.
- Xu, H.Q., 2006. Modification of normalized difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int. J. Remote Sens.* 27, 3025–3033.
- Yan, Y., Kuang, W.H., Zhang, C., Chen, C.B., 2015. Impacts of impervious surface expansion on soil organic carbon - a spatially explicit study. *Sci. Rep. – UK* 5.
- Yang, D., Fu, C.S., Smith, A.C., Yu, Q., 2017. Open land-use map: a regional land-use mapping strategy for incorporating OpenStreetMap with earth observations. *Geo-Spat. Inf. Sci.* 20, 269–281.
- Yu, Y., Tong, Y., Tang, W.W., Yuan, Y.B., Chen, Y., 2018. Identifying spatiotemporal interactions between urbanization and eco-environment in the urban agglomeration in the middle reaches of the Yangtze River, China. *Sustainability*-Basel 10.
- Zha, Y., Gao, J., Ni, S., 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens.* 24, 583–594.
- Zhang, L., Weng, Q.H., Shao, Z.F., 2017. An evaluation of monthly impervious surface dynamics by fusing Landsat and MODIS time series in the Pearl River Delta, China, from 2000 to 2015. *Remote Sens. Environ.* 201, 99–114.
- Zhu, K.G., Xie, M., Wang, T.J., Cai, J.X., Li, S.B., Feng, W., 2017. A modeling study on the effect of urban land surface forcing to regional meteorology and air quality over South China. *Atmos. Environ.* 152, 389–404.
- Zomlot, Z., Verbeiren, B., Huysmans, M., Batelaan, O., 2017. Trajectory analysis of land use and land cover maps to improve spatial-temporal patterns, and impact assessment on groundwater recharge. *J. Hydrol.* 554, 558–569.