Remote Sensing of Environment 196 (2017) 56–75

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Multi-level monitoring of subtle urban changes for the megacities of China using high-resolution multi-view satellite imagery

Xin Huang [a,b,*], Dawei Wen [b,*], Jiayi Li [a], Rongjun Qin [c]

[a] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, PR China
[b] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, PR China
[c] Department of Civil, Environmental and Geodetic Engineering (CEGE), Ohio State University, Columbus, OH 43210, USA

ARTICLE INFO

Article history:
Received 24 April 2016
Received in revised form 28 April 2017
Accepted 3 May 2017
Available online 6 May 2017

Keywords:
Urban
Change detection
Land-cover classification
Landscape
Landsat
ZY-3

ABSTRACT

Megacities have evolved at an unprecedented rate under the pressure of urban population growth and economic development, particularly in the developing countries. For instance, many cities in China are experiencing a large number of rapid but subtle changes resulting from urban infrastructure construction. Information concerning such small changes is imperative for understanding the local environment and human activity, and can also provide key insights for urban planners. Undoubtedly, high-resolution remotely sensed data should play an essential role for the monitoring of such subtle changes, due to the improved observation capacity of the spatial details. However, few studies have investigated high-resolution data for change detection at a large geographic scale, due to the multi-temporal heterogeneity of the data, e.g., spatial mis-registration, parallax distortion for high architectures, and the different viewing angles. In this study, we attempted to fill this gap and resolve these problems by the use of multi-view ZY-3 satellite data, which are used to generate multi-temporal orthographic images through photogrammetric derivation. In the meantime, we present a general framework for precise urban change analysis in a multi-level (pixel, grid, and city block) approach. Two typical Chinese megacities—Beijing and Wuhan—are chosen in the experiments. The results confirm the accuracy of the proposed multi-level method for monitoring subtle urban changes, achieving Kappa coefficients of ~0.8 at the pixel level and a correctness of 93–95% at the grid level. The landscape analysis further indicates that the rapid urban construction led to greater fragmentation and spatial heterogeneity of buildings and decreased minimum distance between building patches (by ~1.0 m between 2012 and 2013). Moreover, the performances of ZY-3 and Landsat for the monitoring of subtle urban changes are compared, revealing that the high-resolution sensor—ZY-3 (2.5 m)—is essential for precisely detecting subtle urban changes, whereas the Landsat data (30 m) are not sensitive to most of the subtle changes that occur in the urban areas.

1. Introduction

Urban areas, which cover only 0.5% of the Earth's land surface (Schneider et al., 2009), host more than half of the world's population, and the urban population is projected to continually increase, particularly in developing countries (UN, 2014). As human-dominated habitants, the megacities have experienced a rapid process of urban development, resulting in a series of negative effects, e.g., resource depletion, land degradation, traffic congestion, environmental and ecological problems, and urban heat island effects (Lasanta and Vicente-Serrano, 2012; Ma et al., 2012; Susaki et al., 2014; Zhou et al., 2014). In-depth studies of urban changes are essential to promote sustainable urbanization (Dewan and Yamaguchi, 2009; He et al., 2006).

China, as the largest developing country, has been undergoing rapid and frequent changes (Fig. 1) due to the "reform and opening-up" policy (Lu et al., 2007). In addition, the pace of urbanization in a lot of Chinese cities has been further accelerated due to the great reliance on the revenue from land transactions and real estate construction (Bai et al., 2011). Specifically, land-related incomes can account for 30–70% of a city's financial revenue (Bai et al., 2014; Lin, 2007). China's macro-control policy, i.e., the central government on behalf of the Chinese public controls and empowers local governments to make land-use decisions, was implemented in order to standardize the land market (Long, 2014). However, authorities uncovered and investigated a total of 53,000 cases of illegal land use across the country in 2010 alone (ChinaDaily, 2011). Moreover, large areas of land (up to 113 km^2), which could provide housing for 1 million households, were found idle nationwide after land developers purchased land-use rights (People'sDaily, 2010).

* Corresponding authors at: School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, PR China.
E-mail addresses: xhuang@whu.edu.cn (X. Huang), daweiwen@whu.edu.cn (D. Wen).

http://dx.doi.org/10.1016/j.rse.2017.05.001
0034-4257/

Fig. 1. An example from Wuhan, China, showing the subtle changes of urban land cover and landscapes. (Source of images: Google Maps.)

Generally speaking, a large number of land parcels have either been illegally occupied or remain idle due to the lack of public participation and the unclear demarcation for land development projects, leading to ineffective use of land resources and challenges to land-use management (Koroso et al., 2013; Li, 2016).

Recently, remote sensing techniques have shown their notable advantages in monitoring urban change, as a result of the large geographic coverage, rich spatial details, and high temporal frequency (Baus et al., 2014). A number of studies concerning urban development have been carried out using remote sensing data, such as the Advanced Very High Resolution Radiometer (AVHRR) data of the National Oceanic Atmospheric Administration (NOAA) (Kressler and Steinnocher, 1999; Stow and Chen, 2002), the Moderate Resolution Imaging Spectroradiometer (MODIS) data (Lunetta et al., 2006; Mertes et al., 2015), Landsat data (Schneider, 2012; Shalaby and Tateishi, 2007), and the Defense Meteorological Satellite Program's Operational Linescan System (DMSP/OLS) nighttime light data (Liu et al., 2012; Ma et al., 2012). It should be noted that these studies have largely focused on the use of coarse- or moderate-resolution data, which are effective for monitoring the macro changes in a large-scale area, but can miss subtle urban changes (Yu et al., 2016).

Therefore, precise urban monitoring, which can detect subtle changes at a very fine scale, is of vital importance for regulating urban land cover/use and landscapes. The availability of high spatial resolution remote sensing data can support such detailed urban monitoring, e.g., the demolition and construction of urban infrastructure, and hence allow in-depth change analysis for landscapes. However, high-resolution change detection has rarely been studied in the existing literature, due to the heterogeneity existing in the multi-temporal high-resolution imagery (e.g., spatial mis-registration, parallax distortion for high architectures, illumination conditions, and different viewing angles) (Stumpf et al., 2014). These problems make it difficult to accurately detect micro changes from high-resolution imagery, leading to a large number of false alarms, especially for pixel-based processing. In this regard, the ZiYuan-3 (ZY-3) satellite, which is China's first civilian high-resolution three-line array stereo satellite (launched in January 2012), has the ability to provide multi-temporal orthographic images, so as to minimize the spatial inconsistency as much as possible using the multi-view mode. In this way, it becomes possible to perform change detection from the high-resolution and multi-temporal orthographic images, courtesy of the merits of the multi-view images.

With respect to the change detection methods, although much progress has been achieved in recent years (Du et al., 2013; Tewkesbury et al., 2015), it is a big challenge for accurate multi-temporal classification and change detection using high-resolution images, especially in the complex urban scenes. Most of the current change detection algorithms employ pixel-based methods (Song et al., 2016). However, it is widely agreed that pixel-based approaches do not exploit spatial or contextual information, and can be subject to the "salt and pepper" effect (Hussain et al., 2013; Yu et al., 2016). Therefore, in this study, a multi-level framework (i.e., pixel, grid, and block) for change detection is presented, in order to make full use of spatial details in high-resolution images and analyze urban changes at various scales.

In the meantime, how to perform accurate multi-temporal classification in an effective way, is also an essential issue, which provides underlying basis for the subsequent change analysis. For the high-resolution image classification, it is not sufficient to only consider spectral bands. In order to improve separability among spectrally similar objects (e.g., soil, roofs, and roads), researchers have proposed a series of spatial and structural features (e.g., pixel shape index (Zhang et al., 2006), morphological profiles (Mura et al., 2010), and textural metrics (Pacifici et al., 2009)). However, in most cases, it seems impossible to select one optimal feature set for different objects and scenes (Huang and Zhang, 2013). In addition, different spatial features can compensate each other in classifying different land covers by characterizing image properties from different perspectives. Consequently, multi-feature image classification is attempted in this paper, in order to improve interpretation efficacy of high-resolution imagery. On the other hand, conducting classification in each time series separately is the most ad hoc way in current literature (El-Kawy et al., 2011; Tarantino et al., 2016). In this way, sample collection is required for each separated classification process, which is time-consuming and labor-intensive, especially at a large geographic scale. Furthermore, it hampers the possibility of classifying time series of images accurately and consistently in a more automated fashion. Considering this, we aim to propose a sample migration strategy to reduce work in sample collection.

In summary, the objective of this study is to address the key research questions for change detection using high-resolution multi-view and multi-temporal satellite imagery, concerning how to

(1) minimize the spatial heterogeneity existing in the multi-temporal high-resolution imagery;
(2) conduct a multi-temporal classification in an efficient way; and
(3) analyze multi-level urban changes.

The proposed methodology is conducted in two representative Chinese cities—Beijing and Wuhan—over a short time period (2012–2013),

in order to test the performance of the proposed method in detecting subtle urban changes.

2. Study sites and data sets

In our study, two typical megacities in China—Beijing and Wuhan—were chosen for the monitoring of urban changes. Beijing, which had a population of over 20 million in 2012 (Beijing Municipal Statistics Bureau, 2013), is the capital of China, and the center for Chinese political and cultural activities and international exchanges. It is an international metropolis located in north China, characterized by the typical chessboard shape of the city structure pattern. In order to protect the historic and cultural conservation areas in the Old City of Beijing, large-scale demolition and reconstruction is forbidden in the Old City (Shin, 2010). Therefore, the city center of Beijing is appropriate for our research into subtle urban change detection.

The other study area, Wuhan, the capital of Hubei province, is one of the biggest metropolises in central China. Wuhan is located in the middle-lower Yangtze Plain, and had a population of over 10 million in 2012 (Wuhan Municipal Statistics Bureau, 2013). The unique locational characteristics of Wuhan, which is referred to as the "thoroughfare to nine provinces", have resulted in it becoming China's largest inland rail and road transportation hub (Huang and Wei, 2014), providing favorable conditions for the rapid development of the city.

ZY-3 remote sensing data were used for monitoring the urban changes. The parameters of the ZY-3 satellite are shown in Fig. 2(b). A notable advantage of ZY-3 is that it carries three panchromatic cameras pointing in forward (3.5 m), backward (3.5 m), and nadir (2.1 m) directions, as well as a multispectral (5.8 m) camera (Wang et al., 2014) (Fig. 2(a)). These merits make ZY-3 an excellent platform for providing multi-temporal high-resolution orthographic imagery with the in-track multi-view mode. Furthermore, the ZY-3 satellite can provide wall-to-wall coverage of the area between 84 degrees north and 84 degrees south of the equator, which covers almost all the area of human habitation. Hence, we can continuously and stably monitor detailed urban dynamics at a large geographic scale with the ZY-3 multi-view images.

ZY-3 images with orthorectification and pansharpening pre-processing (see Section 3.2) are shown in Fig. 3(a) and (b), respectively, for Wuhan and Beijing. The digital surface models (Fig. 3(c) and (d)) derived from the ZY-3 multi-view images are also displayed. The images covering the city centers of Beijing and Wuhan were acquired in 2012 and 2013, corresponding to 8000 × 7500 pixels (375 km^2), and 5884 × 6715 pixels (247 km^2), respectively. Since the ZY-3 satellite was launched in 2012, long time series data are not yet available. However, the existing data can be used to routinely monitor annual urban changes.

3. Methodology

3.1. Overview

In this study, we first conduct the change detection at the pixel level, as this is able to accurately indicate the urban change trajectories. The accuracy of the high-resolution pixel-level change detection can be guaranteed by the multi-view mode of the ZY-3 images and the sophisticated multi-feature image interpretation framework. In spite of this, the pixel-based approach can result in a number of commission errors and the "salt-and-pepper" effect (Yu et al., 2016). Therefore, we further aggregate the pixel-based result into the grid-based one, which is called "hot-spot" change detection (Pacifici and Del Frate, 2010; Wen et al., 2016), in order to mitigate the salt-and-pepper noise occurring at the pixel level. Subsequently, above the grid level, change detection at the landscape level is explored by considering the changes of the composition and configuration in each city block, which is a basic management unit for Chinese cities. Identifying the subtle change information at this level is meaningful for urban planning. Although a lot of papers have addressed landscape analysis over urban areas (Li et al., 2011; Li et al., 2016), few studies have investigated the change of the landscape as cities grow, especially based on high-resolution images (Liu and Yang, 2015). This research attempts to fill this gap.

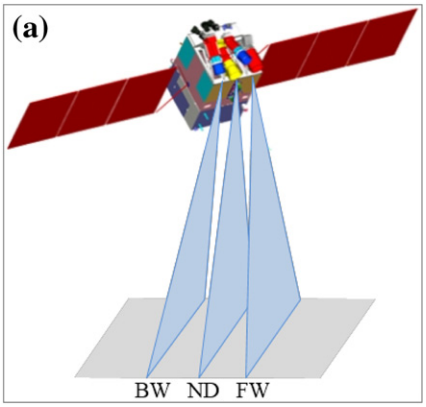
In summary, this study presents a multi-level framework (pixel, grid, and city block) for change detection from multi-temporal and multi-view high-resolution imagery. The key steps of the proposed scheme are: 1) generation of multi-temporal orthographic images from ZY-3 multi-view data; 2) accurate interpretation of the high-resolution imagery based on multi-feature interpretation; 3) multi-temporal land-cover mapping with a sample migration strategy; and 4) multi-level urban dynamics monitoring (Fig. 4).

3.2. Generation of multi-temporal orthographic images

As China's first civilian high-resolution satellite, ZY-3, which carries multiple optical cameras (i.e., three cameras in forward, nadir, and backward modes), can acquire multi-view images over the same area without body tilting (Fig. 2(a)). This notable advantage enables us to obtain multi-temporal orthographic images, as well as photogrammetrically derived digital surface models (DSMs). Courtesy of the multi-view mode of the ZY-3 satellite, it has become possible to conduct pixel-based change detection using high-resolution images. The procedure for generating the DSMs and orthographic images is as follows.

3.2.1. DSM generation

The DSMs used in this study were processed by the hierarchical semi-global matching (SGM) method (Hirschmuller, 2008). Firstly, a



(b) **The parameters of the ZY-3 satellite**

Sensor	Band	Wavelength	Spatial resolution
PAN	PAN-FW	500~800nm	3.5m
	PAN-BW	500~800nm	3.5m
	PAN-ND	500~800nm	2.1m
MS	Blue	450~520nm	5.8m
	Green	520~590nm	5.8m
	Red	630~690nm	5.8m
	Infrared	770~890nm	5.8m

Fig. 2. (a) Schematic diagram showing the imaging parameters of the ZY-3 satellite. The three panchromatic cameras of the ZY-3 satellite pointing in forward, backward, and nadir directions can simultaneously collect multi-view panchromatic images, allowing for stereo mapping on the same orbital pass. (b) The parameters of the ZY-3 satellite (PAN = panchromatic bands, MS = multispectral bands, FW = forward, BW = backward, and ND = nadir).

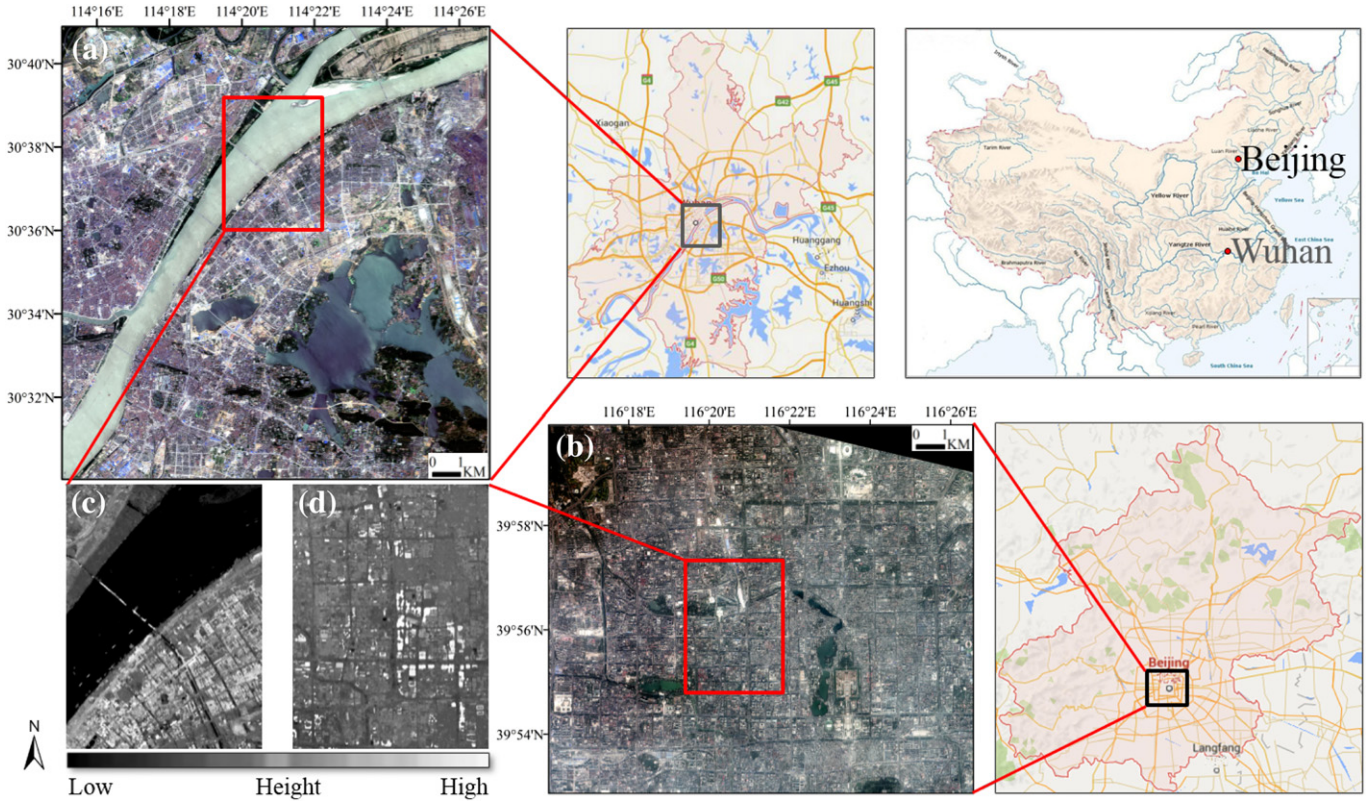


Fig. 3. Overview of the study sites and data sets: (a) and (b) are the orthographic images for the central areas of the two megacities—Wuhan and Beijing—respectively; (c) and (d) are the corresponding digital surface models.

quasi-epipolar stereo pair was generated by performing first-order bias correction (Fraser and Hanley, 2003) on the stereo pairs. The disparity map of the epipolar images can be obtained by a hierarchical SGM, where a coarse-to-fine strategy is adopted by performing the matching in each pyramid image, in order to increase the efficiency, in terms of time and memory, of the processing (Rothermel et al., 2012). The 3D point clouds can be subsequently generated through forward intersection followed by raster grid resampling. Finally, inverse distance interpolation was used to fill the missing pixels on the building edges caused by matching failure (Qin, 2014).

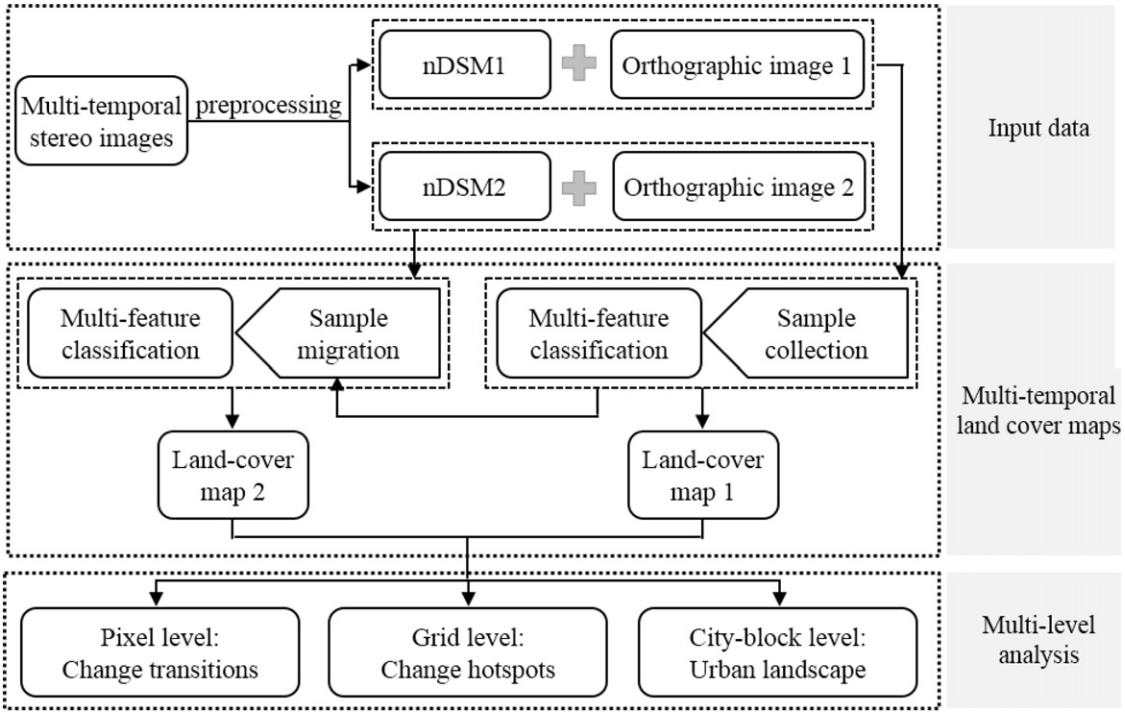


Fig. 4. Workflow showing the procedure of the proposed method for urban landscape monitoring (nDSM: normalized digital surface model).

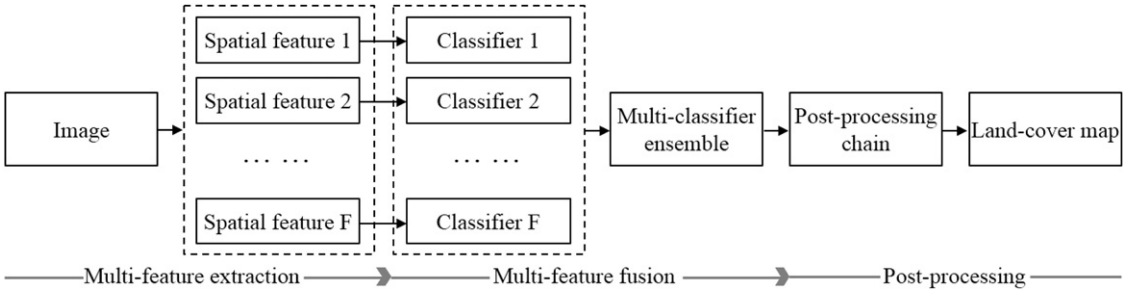


Fig. 5. The multi-feature classification framework for urban land-cover mapping.

DSMs describe vertical attributes of the Earth's surface, including the terrain and elevated land-cover objects. Normalized DSMs (nDSMs), delineating the height of objects above the Earth's surface by removing topographic effects, can be calculated using the top-hat morphological filter (Qin and Fang, 2014; Tian and Reinartz, 2013). In this study, in order to evaluate the accuracy of the nDSMs, 100 ground points and 100 off-ground points (i.e., buildings) were randomly selected. The corresponding reference height values of the off-ground points were obtained from the local urban planning department. The root-mean-square error (RMSE) was used to evaluate the overall error of the nDSM derivation.

3.2.2. Orthorectification

Orthorectification is defined as the spatial manipulation that projects each pixel from satellite images onto a ground reference system. In this study, the perspective distortion of the satellite images was corrected to a parallel projection using the generated DSM. To avoid occlusions, all the pixels from the stereo images contributed to the resulting orthographic image. The depth-buffer algorithm (Merrell et al., 2007) was used to accommodate the occluded area.

3.3. Accurate land-cover mapping

Our goal was to achieve a land-cover classification that was as accurate as possible, in order to guarantee the effectiveness of the urban dynamics monitoring. To achieve this, a multi-classifier ensemble method was adopted since it can exploit the strengths of the individual classifiers and obtain an enhanced performance (Mangai et al., 2010; Woźniak et al., 2014b). More specifically, multiple classifiers, containing different spatial features and spectral bands, were integrated, through

Table 1
The multiple features considered in this study.

Features	Principle	Formula
Texture		
Gray level co-occurrence matrix (GLCM) (Haralick et al., 1973)	The GLCM is defined as a matrix, with each entry estimating the probability of the pair of gray levels that occur in a certain distance and direction. A set of textural measures such as homogeneity (HOM), entropy (ENT), and contrast (CON) can then be extracted based on this co-occurrence matrix to delineate the textural statistics of the images.	$\mathrm{HOM} = \sum_i \sum_j \frac{f(i,j)}{1+\lvert i-j\rvert}$ $\mathrm{ENT} = \sum_i \sum_j f(i,j)\ \log(f(i,j))$ $\mathrm{CON} = \sum_i \sum_j f(i,j)(i-j)^2$ where $f(i,j)$ is the probability of the pairs of gray levels i and j that occur in a specified distance and direction.
Morphology		
Differential morphological profiles (DMPs) (Aytekin and Ulusoy, 2011; Benediktsson et al., 2003)	DMPs can record image structural information by measuring the slope of the opening-closing profiles for every step of the scale parameters. They can be viewed as the shape spectrum of an image.	$\mathrm{DMP} = \begin{cases} DMP_O = \lvert\gamma^{\lambda}(I) - \gamma^{\lambda-1}(I)\rvert \\ DMP_C = \lvert\phi^{\lambda}(I) - \phi^{\lambda-1}(I)\rvert \end{cases}$ where $\gamma^{\lambda}(I)$ and $\varphi^{\lambda}(I)$ represent morphological opening and closing by reconstruction for an image I, with λ being the scale parameter of the structural element.
Multiple indices		
Morphological building index (MBI) (Huang and Zhang, 2011)	The MBI is an effective index to highlight building structures from high-resolution imagery by representing the spectral-spatial properties of buildings (e.g., brightness, contrast, size, and directionality) with a series of morphological operators.	$\mathrm{MBI} = \frac{\sum_{d,s} \mathrm{DMP}_{\mathrm{W-TH}}(d,s)}{D \times S}$ where $\mathrm{DMP}_{\mathrm{W\text{-}TH}}$ denotes the differential morphological profiles of the white top-hat, and d and s indicate the scale and direction of the structural element, respectively.
Morphological shadow index (MSI) (Huang and Zhang, 2012a)	The derivation of the MSI is based on the fact that shadows show low reflectance but high local contrast; hence, the MSI is defined as the dual function of the MBI, i.e., the black top-hat morphological profiles, to highlight the shadow structures.	$\mathrm{MSI} = \frac{\sum_{d,s} \mathrm{DMP}_{\mathrm{B-TH}}(d,s)}{D \times S}$ where $\mathrm{DMP}_{\mathrm{B\text{-}TH}}$ denotes the differential morphological profiles of the black top-hat in contrast to the white top-hat in the MBI.
Urban complexity index (UCI) (Huang and Zhang, 2012b)	The UCI, constructed on the basis of the energy function of the 3-D wavelet transformation, is able to distinguish between urban classes (e.g., buildings and roads) and natural classes (e.g., water, soil, trees, and grass) based on the fact that natural classes have a relatively smaller spatial variation than spectral variation, but urban classes show more variation in the spatial domain.	$\mathrm{UCI} = \frac{E(HLL)+E(LHL)+E(HHL)}{E(LLH)+E(LHH)+E(HLH)}$ $E_f = \sum_i \sum_j \sum_m (f(i,j,m))^2$ where E_f is the energy function of sub-band f of the 3-D wavelet decomposition, and i, j, m denote the coordinates in the spatial-spectral domain; E(**H) and E(**L) indicate the spectral variation, spatial variation with high- and low-pass filters in the spectral domain, respectively.
Normalized difference vegetation index (NDVI) (Huete et al., 2002)	The NDVI is constructed by exploiting the great difference in the spectral reflectance of vegetation in the near-infrared and red bands.	$\mathrm{NDVI} = \frac{\mathrm{NIR}-\mathrm{R}}{\mathrm{NIR}+\mathrm{R}}$ where NIR, R is the reflectance in the near-infrared and red bands, respectively.
Normalized difference water index (NDWI) (McFeeters, 1996)	The NDWI is used to enhance the water feature based on the principle that water has strong reflectance in the green band but strong absorption in the near-infrared band.	$\mathrm{NDWI} = \frac{\mathrm{G}-\mathrm{NIR}}{\mathrm{G}+\mathrm{NIR}}$ where G, NIR is the reflectance in the green and near-infrared bands, respectively.

Table 2
Rules for the classification post-processing.

Mixed classes Class 1 & Class 2[a]	Attributes	Principles	Rules Class 1 → Class 2[b]
Buildings & road Buildings & soil	Relative border (RB); distance to shadow (DS); height (H)	Buildings and shadows are spatially adjacent, and buildings have a relatively large height.	✓ Class (O) = buildings ✓ RB to nearest road/soil object > T_1 ✓ DS > 0 ✓ H = 0 Buildings → road/soil ✓ Class (O) = road/soil ✓ RB to nearest building object > T_1 ✓ DS = 0 ✓ H > 0 Road/soil → buildings
Shadow & water	Relative border (RB); distance to buildings (DB)	Buildings and shadows are spatially adjacent.	✓ Class (O) = shadow ✓ RB to nearest water object > T_2 ✓ DB > 0 Shadow → water ✓ Class (O) = water ✓ RB to nearest shadow object > T_2 ✓ DB = 0 Water → shadow
Trees & grass	Relative border (RB); height (H)	Trees are always higher than grass.	✓ Class (O) = trees ✓ RB to nearest grass object > T_3 ✓ H = 0 Trees → grass ✓ Class (O) = grass ✓ RB to nearest trees object > T_3 ✓ H > 0 Grass → trees

[a] The possible misclassification between *Class 1* and *Class 2* is labeled as "*Class 1 & Class 2*" in the table.
[b] "*Class 1 → Class 2*" indicates that *Class 1* is reclassified as *Class 2* when the listed conditions are satisfied.

which the multiple features were effectively fused. The whole processing chain is shown in Fig. 5, and a detailed description of each step is introduced in the following subsections.

3.3.1. Multi-feature extraction

In remote sensing image classification, different features (e.g., spectral, textural, and morphological features) can describe image characteristics from different perspectives (Zhang et al., 2012). Spectra reflect the physical attributes of land-cover classes and play a key role in intuitively describing the land covers (Zhao et al., 2012). Textural features are able to represent the spatial distribution of image primitives, such as smoothness, coarseness, regularity, contrast, or other patterns (Peña-Barragán et al., 2011; Trias-Sanz et al., 2008). Recently, mathematical morphology has been the subject of much attention for image

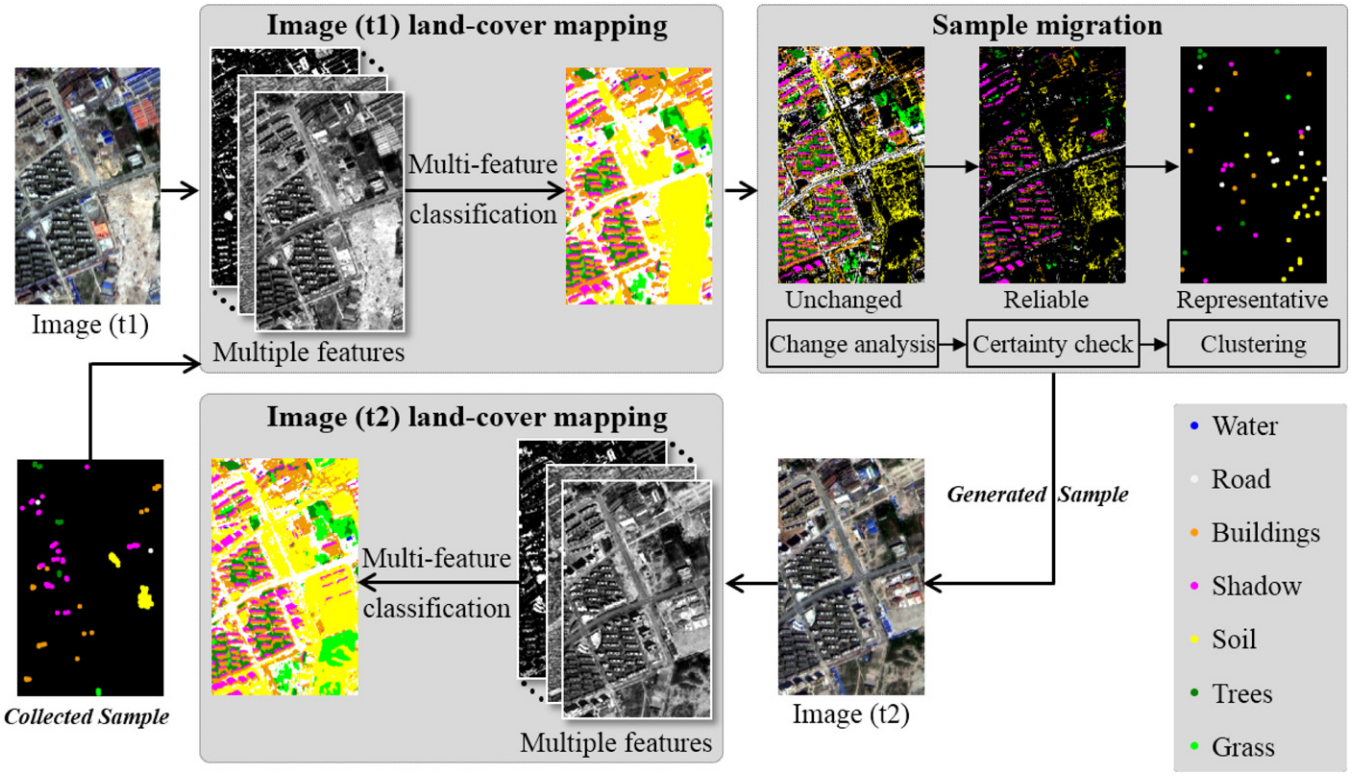


Fig. 6. The framework for multi-temporal land-cover mapping with a sample migration strategy.

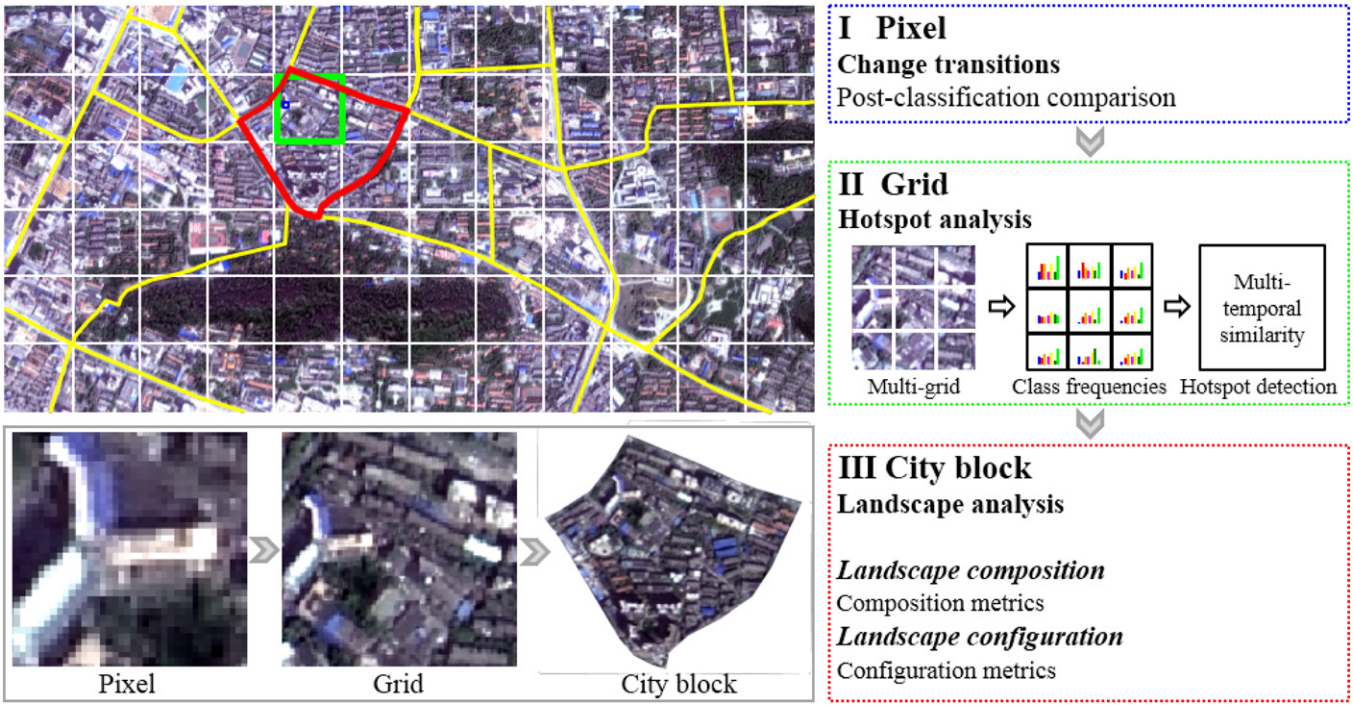


Fig. 7. The multi-level analysis of urban dynamics.

classification and urban mapping from high-resolution data (Plaza et al., 2009; Rogge et al., 2007; Serna and Marcotegui, 2014; Tuia et al., 2011). This approach involves extracting the structural information through analyzing the spatial relationship between groups of pixels with a series of structural elements (Benediktsson et al., 2003). In addition, a multi-index feature set was considered in this study, where a set of low-dimensional information indices (e.g., buildings, vegetation, water, and shadow indices) were used to represent the complex urban scenes (Huang et al., 2014a). A detailed description of the features considered in this research is provided in Table 1.

3.3.2. Multi-feature fusion

A key issue for multi-feature image classification is how to effectively integrate and interpret the multiple feature sources, e.g., spectral, textural, morphological, and multi-index features. The objective of the multi-feature fusion is to adequately exploit the characteristics derived from the various feature sources, and thus enhance the separability among different land-cover classes, especially for the similar classes (e.g., roads and buildings, shadow and water) (Dell'Acqua et al., 2004; Huang and Zhang, 2010; Plaza et al., 2009). Typically, there are two approaches to multi-feature classification: the first is feature stacking, i.e., concatenating the multiple features into a single vector and then classifying the stacked vector via a classifier (Bruzzone and Carlin, 2006; Pagot and Pesaresi, 2008); the other approach is the multiple classifier ensemble, i.e., feeding one feature source to a classifier, and then assembling the multiple classifiers in a decision-fusion manner (Huang and Zhang, 2010, 2013; Pelletier et al., 2016; Woźniak et al., 2014a; Zhang, 2015; Zhong and Wang, 2007). The latter approach, i.e., the multi-classifier ensemble, can overcome the over-fitting problem caused by the high-dimensional feature space in the vector-stacking method (Du et al., 2012; Liu et al., 2011). In addition, the multi-classifier method can reduce the uncertainty by the information fusion at the decision level.

In this research, a recently developed multi-feature land-cover mapping method for high-resolution imagery, i.e., the multi-feature classifier ensemble, was adopted to generate the land-cover maps. Initially, each classifier, referring to the textural, morphological, and multi-index features, respectively, was performed to generate crisp (class label) and soft (class probability) land-cover maps. The pixel-wise probabilistic outputs were then spatially aggregated into objects generated

Table 3
The metrics for landscape configuration considered in this study.

Category	Metric	Description
Area-edge	Largest patch index (LPI)	The percentage of the total landscape area that is comprised by the largest building patch.
	Edge density (ED)	The length of all the building patches per unit area.
	Mean patch area (MPA)	The average area of the building patches within a city block.
	Standard deviation of patch area (SDPA)	The standard deviation of the area of the building patches within a city block.
Shape	Mean shape index (MSI)	The average shape complexity for the building patches.
	Standard deviation of shape index (SDSI)	The standard deviation of the shape complexity for the building patches.
Aggregation	Mean nearest neighbor distance (MNN)	The average distance of a building patch to its nearest neighbor building patch.
	SD of the mean nearest neighbor distance (SDMNN)	The standard deviation of the distance of a building patch to its nearest neighbor building patch.
	Patch density (PD)	The number of building patches per unit area.
	Cohesion index (CI)	The indication of the connectedness of the building patches.

Table 4
Sample number and accuracy (correctness) of the land-cover mapping.

	Accuracy % (sample #)			
	Beijing 2012	Beijing 2013	Wuhan 2012	Wuhan 2013
Water	**87.9** (41219)	**95.1** (45663)	**91.1** (110114)	**99.1** (103567)
Road	**92.8** (68821)	**86.1** (73038)	**90.8** (106331)	**92.5** (105846)
Buildings	**83.4** (75672)	**81.9** (80203)	**87.2** (109487)	**85.0** (104491)
Shadow	**80.3** (35374)	**79.2** (39317)	**96.8** (55039)	**95.8** (52383)
Soil	**82.7** (17277)	**81.4** (15581)	**95.0** (55238)	**90.1** (52037)
Trees	**93.7** (14576)	**92.5** (16730)	**94.2** (54643)	**83.0** (25520)
Grass	**87.1** (13448)	**82.9** (12584)	**91.3** (56187)	**94.9** (26182)
Overall	**86.8** (266387)	**85.4** (283116)	**91.5** (547039)	**92.0** (470026)

Table 5
Error matrix and accuracy assessment for the pixel-level change transitions in Beijing (UA = user's accuracy, PA = producer's accuracy, and OA = overall accuracy).

Classified data	Reference data						Total	UA (%)
	Soil → buil	Buil → soil	Soil → grass	Grass → soil	Water → grass	No-change		
Soil → buil	23	0	0	0	0	0	23	100.0
Buil → soil	0	25	0	0	0	5	30	83.3
Soil → grass	0	0	28	0	0	0	28	100.0
Grass → soil	0	0	0	25	0	0	25	100.0
Water → grass	0	0	0	0	23	0	23	100.0
No-change	7	5	2	5	7	25	51	49.0
Total	30	30	30	30	30	30	180	
PA (%)	76.7	83.3	93.3	83.3	76.7	83.3		
OA (%)	82.8							
Kappa coefficient	0.79							

by adaptive mean-shift segmentation (Huang and Zhang, 2013). The object-based outputs were calculated through a weighted combination of the different soft land-cover maps. Specifically, the object-based probabilistic outputs were calculated by:

$$P^k(O) = \frac{\sum_{x \in obj} \sum_{f=1}^{N} P_f^k(x) \cdot w_f(x)}{N \cdot F} \quad (1)$$

$$w(x) = \sum_{k=1}^{K-1} \left[\hat{P}_k(x) - \hat{P}_{k+1}(x) \right] \cdot \frac{1}{k} \quad (2)$$

where $P^k(O)$ is the object-based class probability of object O for class k. $P_f^k(x)$ refers to the pixel-based class probability of pixel x for classifier f. The number of pixels in the object and the number of base classifiers are denoted by N and F, respectively. The classification certainty, which is measured according to the descending multi-class probabilistic outputs, i.e., $\hat{p}_1(x), \ldots, \hat{p}_k(x), \ldots, \hat{p}_K(x)$, is employed as the weight value for the multiple classifier fusion (denoted as $w(x)$ in Eq. (2)). A larger value indicates a more reliable classification.

3.3.3. Classification post-processing

In order to guarantee the accuracy of the land-cover mapping, a set of post-processing rules were applied to further refine the initial classification result, according to implicit rules related to the spatial arrangement or surrounding contexts (Hüllermeier and Brinker, 2008; Huang et al., 2014b). Based on the object-based probabilistic output (Eq. (1)), all the image objects can be divided into reliable and unreliable ones (Eqs. (3) and (4)), and the post-processing rules (Table 2) were applied to the unreliable objects only:

$$O_{un} : P_{\max}(O) < T \quad (3)$$

$$P_{\max}(O) = \max\left\{ P^k(O), k = 1, \ldots, K \right\} \quad (4)$$

where $P_{\max}(O)$ is the object-based probability for the winning label (i.e., the maximum probabilistic output) for object O. The threshold value T is used to control the proportion of unreliable objects.

3.4. Multi-temporal land-cover mapping

In traditional multi-temporal land-cover mapping (Shalaby and Tateishi, 2007; Yuan et al., 2005), each classification process in the multiple time series is conducted separately. However, sample collection is required for each classification process, which is time-consuming and labor-intensive, especially at a large geographic scale. To tackle this problem, we propose a multi-temporal land-cover mapping method with a sample migration strategy (Fig. 6). The idea is to apply the unchanged class labels from the classification at time 1 to classify the image at time 2. The unchanged pixels (their positions and class labels) that satisfy the three conditions (i.e., unchanged, reliable, and representative) are migrated to the image at time 2 and considered as training samples. In this way, the cost of collecting the training samples can be significantly reduced. As emphasized in Fig. 6, the sample migration from time 1 to time 2 should refer to the unchanged, reliable, and representative pixels:

- Unchanged: Naturally, the class labels of pixels whose land cover type is not changed can be reused in the subsequent time phase. In this research, change vector analysis (CVA) (Johnson and Kasischke, 1998) was employed to make sure that the migrated samples were unchanged. In this situation, therefore, we could simply set a strict threshold for CVA (0.005 in this paper).
- Reliable: The degree of certainty for a pixel at time 1 can be measured by its multi-class probabilistic output (Yager, 1992) (see Eq. (2)). The certainty metric measures the confidence of the pixel label by calculating the differences between the probability values of the ordered classes. The larger the difference, the higher the certainty. Only the positions with high-certainty samples are considered.

Table 6
Error matrix and accuracy assessment for the pixel-level change transitions in Wuhan (UA = user's accuracy, PA = producer's accuracy, and OA = overall accuracy).

Classified data	Reference data						Total	UA (%)
	Soil → buil	Buil → soil	Soil → grass	Grass → soil	Water → grass	No-change		
Soil → buil	23	0	0	0	0	0	23	100.0
Buil → soil	1	23	0	0	0	0	24	95.8
Soil → grass	0	0	26	0	0	0	26	100.0
Grass → soil	0	0	0	20	0	0	20	100.0
Water → grass	0	0	2	0	30	0	32	93.8
No-change	6	7	2	10	0	30	55	54.6
Total	30	30	30	30	30	30	180	
PA (%)	76.7	76.7	86.7	66.7	100.0	100.0		
OA (%)	84.4							
Kappa coefficient	0.81							

- Representative: The spectral-spatial k-nearest neighbor (KNN) method (Rajadell et al., 2014) was adopted to further purify the reliable and unchanged samples.

3.5. Multi-level urban dynamics monitoring

In order to conduct a comprehensive and systematic urban dynamics analysis, three levels (Fig. 7) were considered in our study, i.e., pixel, grid, and city block. Courtesy of the multi-temporal orthographic images provided by the ZY-3 sensor, we could achieve pixel-level change analysis using high-resolution images, concerning the detailed change trajectories. Post-classification comparison (PCC) (Aguirre-Gutiérrez et al., 2012; Sesnie et al., 2008; Shalaby and Tateishi, 2007) was employed at this level. Secondly, the grid-based analysis was aimed at highlighting the changed areas (i.e., hotspots (Pacifici and Del Frate, 2010)), and, in the meantime, reducing the salt-and-pepper effect from the pixel-level processing. Finally, the landscape analysis using the city block as the unit of calculation focused on the change of the composition (e.g., building coverage ratio and vegetation fraction) and configuration (landscape configuration metrics).

3.5.1. Pixel-level change transitions

PCC (Aguirre-Gutiérrez et al., 2012; Sesnie et al., 2008; Shalaby and Tateishi, 2007) was performed to investigate the land-cover change transitions, based on the multi-temporal land-cover maps. The PCC method is considered effective in minimizing the impact related to the atmospheric and environmental difference between multi-temporal images (El-Kawy et al., 2011). Additionally, it can not only indicate change but also provide "from-to" trajectories. It is well known that the key issue for PCC is the error accumulation (Tewkesbury et al., 2015). Therefore, in this research, a sophisticated land cover classification chain was used, involving the multi-feature classifier ensemble, post-processing, and the multi-temporal sample migration, in order to guarantee the accuracy of the multi-temporal land cover maps.

3.5.2. Grid-level hot-spot change analysis

Grid-based hot-spot analysis is aimed at discovering changed areas rather than analyzing individual pixels (Pacifici and Del Frate, 2010; Wen et al., 2016), based on an assumption that changed areas in the high-resolution images refer to patches instead of pixels. In this research, the hot spots were further divided into two categories: i.e., changes of artificial objects and natural objects, according to the land cover maps.

The hot-spot analysis is demonstrated in Fig. 7. Firstly, an image is divided into a series of grids. Every grid is then further partitioned into a set of non-overlapping cells, where the frequencies of the land cover types in each cell are used to characterize the spatial distribution and arrangement of the classes in the grid. The change intensity for each grid is then calculated by comparing the multi-temporal frequencies of the land cover types in the grid:

$$Dist = \sum_i \sum_j |H_1(i,j) - H_2(i,j)| \quad (5)$$

where *Dist* is the distance measure indicating the similarity between the multi-temporal image grids. $H_1(i,j)$ and $H_2(i,j)$ stand for the frequency of the *i*th land-cover class within the *j*th cell, for times 1 and 2, respectively. A large Dist value indicates a high degree of change at the grid level.

3.5.3. City block level landscape analysis

The city block, segmented by street networks, is the smallest administrative unit for Chinese cities. The city block landscape parameters serve as an important data base for urban management and planning (Leitão et al., 2012). In this study, the city block networks were derived from OpenStreetMap (OSM), where both primary- and secondary-level roads were considered. The landscape metrics include both composition and configuration. Landscape composition is expressed as a series of quantitative indices assessing the proportion of the land-cover classes, whereas landscape configuration measures the arrangement, distribution, and spatial characteristic of patches (McGarigal and Marks, 1995). It should be noted that the landscape metrics were calculated for each city block in this study.

In this research, two major landscape composition parameters were of interest: the vegetation fraction (VF) (Berger et al., 2013; Faryadi and Taheri, 2009; Hur et al., 2010; Leslie et al., 2010) and the building coverage ratio (BCR) (Heiden et al., 2012; Salomons and Pont, 2012; Yu et al., 2010). The VF (BCR) for each block is the ratio between the area covered by vegetation (buildings) and the total area. Please note that the indicator of urban land cover in moderate-resolution images always refers to the impervious layer (Sexton et al., 2013; Song et al., 2016), due to the limitation of the spatial resolution, whereas high-resolution data can be used to accurately calculate the BCR, which is more directly correlated to the urban heat island effect (Peng et al., 2011) and population distribution (Tatem et al., 2007). To date, few studies have addressed building density change. Therefore, in this study, we aimed to investigate the change of BCR at the landscape level.

Furthermore, in order to measure the spatial distribution of the land-cover classes, a set of configuration metrics (Table 3) were considered, including: 1) the patch area and edges: largest patch index (LPI), edge density (ED), and mean and standard deviation of the patch area (MPA and SDPA); 2) shape: mean and standard deviation of the shape index (MSI and SDSI); and 3) aggregation: mean and standard deviation of the nearest neighbor distance (MNN and SDMNN), patch density (PD), and cohesion index (CI) (McGarigal et al., 2002; McGarigal and Marks, 1995; Taubenböck et al., 2014). In this study, the considered landscape configuration metrics were only calculated for the buildings in each city block.

4. Results and analysis

4.1. Multi-temporal land-cover mapping

The reference data were collected by visual inspection (courtesy of the high spatial resolution of the ZY-3 data) and additional information from field survey. The training samples for the 2012 images included 500 randomly selected pixels for each class (Beijing or Wuhan), and the remaining pixels were used to test the classification accuracy. No training samples were required for the 2013 images since they could be automatically generated by the sample migration method (Section 3.4). The test samples for the 2013 images were still required to validate the accuracy of the land cover mapping. The classification accuracies are presented in Table 4. In general, the proposed framework is effective for classifying the multi-temporal high-resolution images. The overall

Table 7
The percentages of the changed types for the pixel-level change detection results.

Change trajectory	Beijing	Wuhan
Soil → buildings	28.9% (84.7 ha)	6.1% (82.1 ha)
Buildings → soil	55.6% (163.2 ha)	34.3% (458.1 ha)
Soil → grass	4.5% (13.1 ha)	40.5% (541.4 ha)
Grass → soil	10.7% (31.3 ha)	11.8% (157.3 ha)
Water → grass	0.3% (1.0 ha)	7.3% (98.2 ha)

Table 8
Accuracy of the hot-spot analysis.

	Beijing		Wuhan	
	Accuracy (#)	Accuracy (%)	Accuracy (#)	Accuracy (%)
Artificial objects	74/81	91.4	259/287	90.2
Natural objects	294/332	88.6	1400/1697	82.5
All	384/413	93.0	1885/1984	95.0

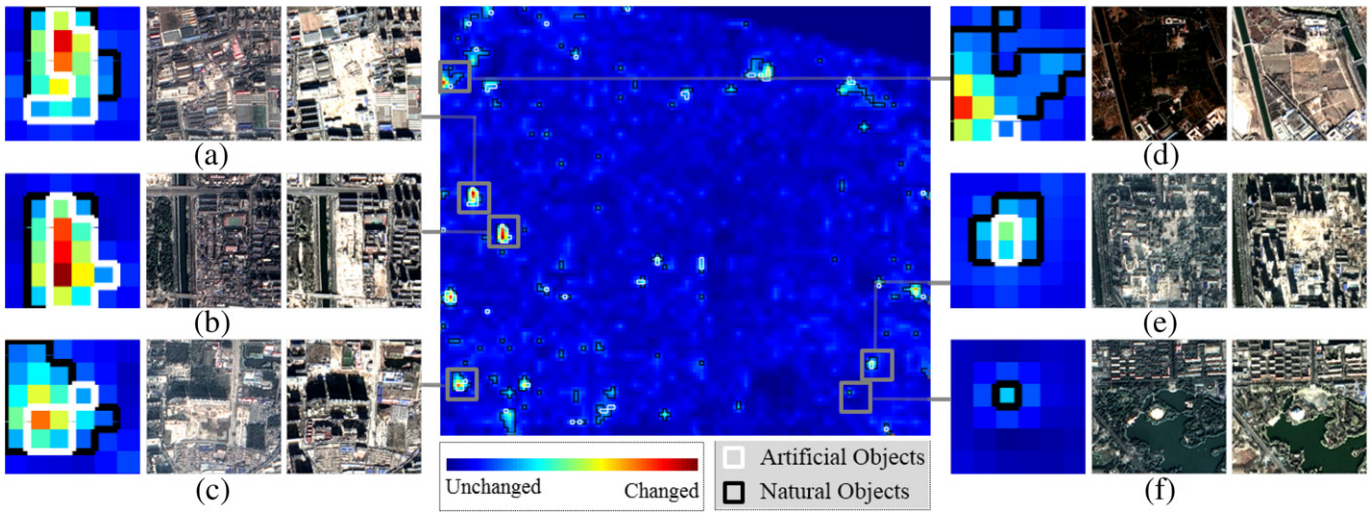


Fig. 8. The hot-spot change detection results for Beijing.

accuracy for the land-cover mapping was 85–87% for Beijing and 91–92% for Wuhan. The correctness for the specific land-cover categories ranged from 79% to 100%, with nearly all of the accuracies being larger than 80%. To sum up, the proposed multi-temporal land-cover mapping method was able to guarantee the accuracy and reliability of the change analysis.

4.2. Pixel-level change transitions

We randomly generated 30 check points for each change transition. Theoretically, seven land-cover classes involve 42 (7×6) types of change transitions. Therefore, when including the "no-change" category, 43 categories could be defined. However, in fact, there are only six types of change occurred in the study areas (Tables 5 and 6). Please note that some change categories rarely happen, e.g., buildings to water, trees to road, water to trees, and road to buildings. In addition, some change categories were not found in either study area, e.g., soil to trees, trees to soil, and road to soil, due to the short time span. The overall accuracy, the Kappa coefficient, and the producer's and user's accuracy for each change category were generated from the error matrices (Tables 5 and 6). The results show that reliable accuracies for the pixel-level change detection were achieved, with the overall accuracy being 82.8% and 84.4% and the Kappa being 0.79 and 0.81, for Beijing and Wuhan, respectively.

The proportion of each change trajectory was further analyzed (Table 7). In Beijing, 28.9% of the changes (84.7 ha) were from soil to buildings, and 55.6% of the changes (163.2 ha) were from buildings to soil, indicating that the area of changes related to building demolition was approximately twice that of the area of changes related to building construction. On the other hand, in the case of Wuhan, a more significant discrepancy between building construction (6.1%, 82.1 ha) and demolition (34.3%, 458.1 ha) was observed. Comparing the conversion between buildings and soil in the two cities, it can be found that more demolition activities took place in Wuhan (458.1 ha) than in Beijing

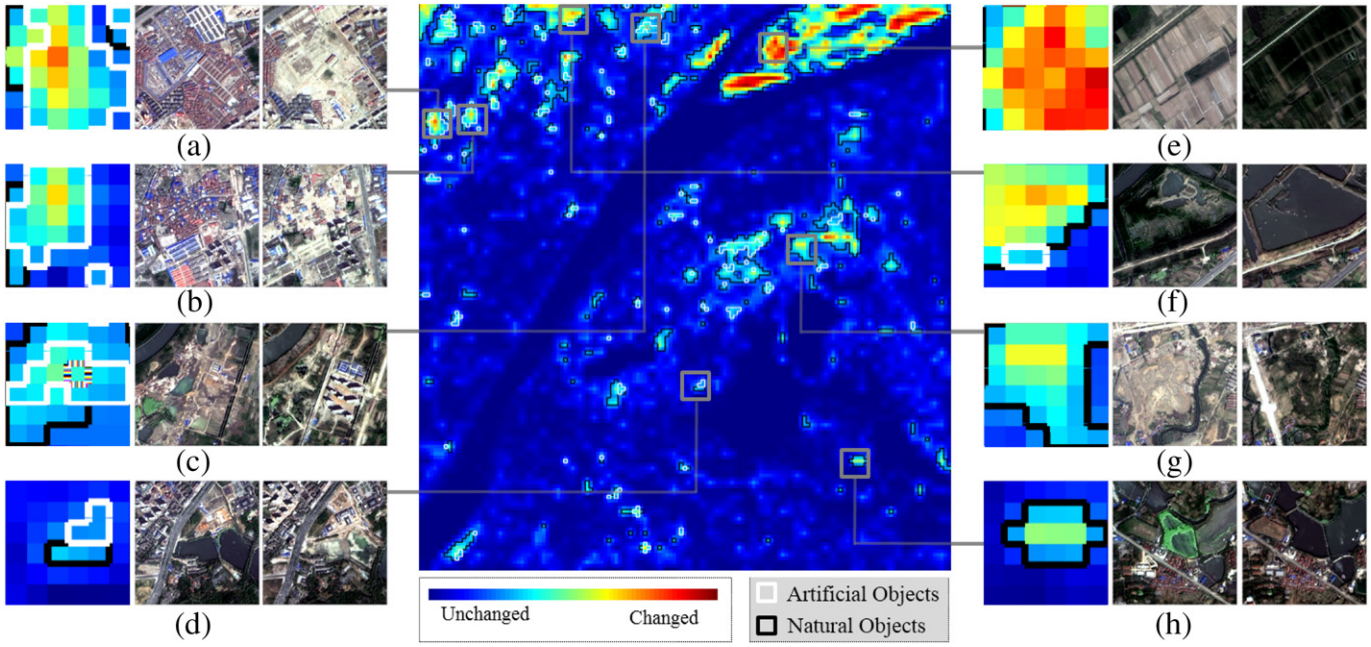


Fig. 9. The hot-spot change detection results for Wuhan.

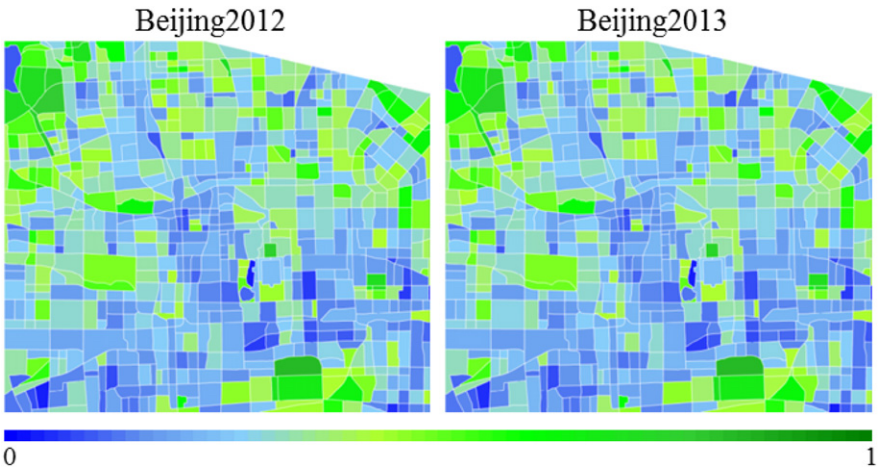


Fig. 10. Vegetation fraction for Beijing.

(163.2 ha). This phenomenon can be attributed to the urban planning policy of Wuhan, i.e., remolding the Old City for infrastructure construction (Cheng and Zhou, 2015; Liu et al., 2015). On the other hand, a strict policy of land development was implemented in Beijing, for the purpose of protecting Beijing Old City (Shin, 2010).

In Table 7, it can be observed that the natural changes that occurred in Wuhan were more active than in Beijing, since the seasonal effect was more significant for the former, i.e., the acquisition time of the Wuhan imagery was April 22, 2012, and August 12, 2013, but the acquisition time of the Beijing imagery was October 11, 2012, and November 28, 2013. For instance, the conversion from soil to grass accounted for 40.5% of the total changes in Wuhan (in contrast to 4.5% in Beijing), as the bare land observed in the non-growing season was covered by vegetation in the growing season.

4.3. Grid-level hot-spot analysis

The accuracy of the grid-level hot-spot change detection is assessed in Table 8. In general, for the Beijing data set, 384 out of 413 blocks were correctly detected (93.0%), and for Wuhan, the hot-spot detection generated 1885 correct blocks among 1984 grids, giving a detection accuracy of 95.0%. Moreover, 74 of 81 (91.4%) and 259 of 287 (90.2%) grids were correctly identified as changes related to artificial objects for Beijing and Wuhan, respectively. Concerning natural objects, the accuracies were 294/332 (88.6%) and 1400/1697 (82.5%) for Beijing and Wuhan, respectively.

The result of the hot-spot change detection is depicted in Fig. 8 (Beijing) and Fig. 9 (Wuhan), where the changed areas are highlighted by polygons. The black and white polygons are separately used to indicate the hot spots related to artificial and natural changes, respectively.

The hot-spot maps can be applied to analyze and monitor the urban development, as well as the change patterns. In the case of Beijing, most of the changed grids are distributed at the urban fringes, with a small quantity of changed grids located in the urban core, due to the policy of strict control over the land development in the downtown for the preservation of the historical style in the pattern and shape of the Old City (Shin, 2010). On the other hand, more changed grids occur in Wuhan, both in the core and fringe of the city. This phenomena can be attributed to the fact that Wuhan, a city in central China, is still at the acceleration stage of urban development, involving a large amount of infrastructure construction, e.g., tunnels and subways (Cheng and Masser, 2003; Cheng and Zhou, 2015). In the meantime, a large number of old or informal settlements (e.g., urban villages (Huang et al., 2015)) have been demolished to make room for the new urban facilities.

Representative examples of the hot-spot change detection results are provided. In the result for Beijing (Fig. 8), cases (a), (b), and (c) correspond to changes of artificial objects, and (d), (e), and (f) are related to natural objects. In order to provide land resources for real estate

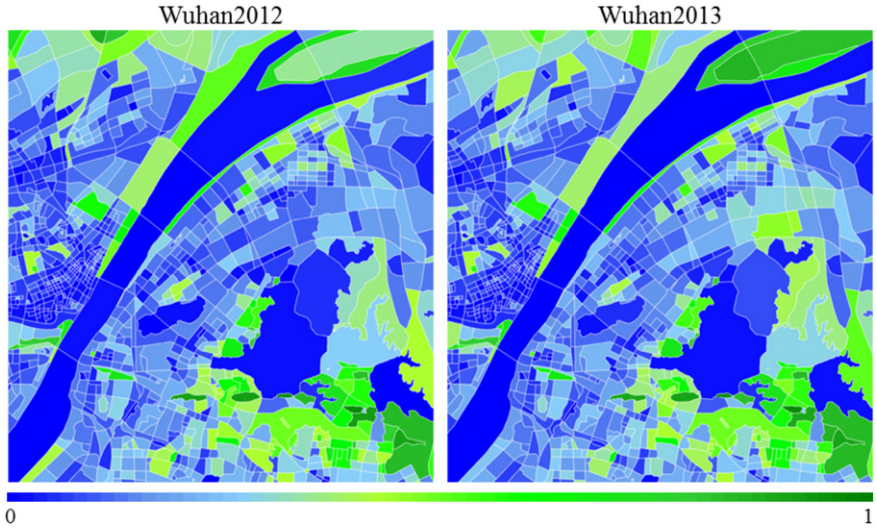


Fig. 11. Vegetation fraction for Wuhan.

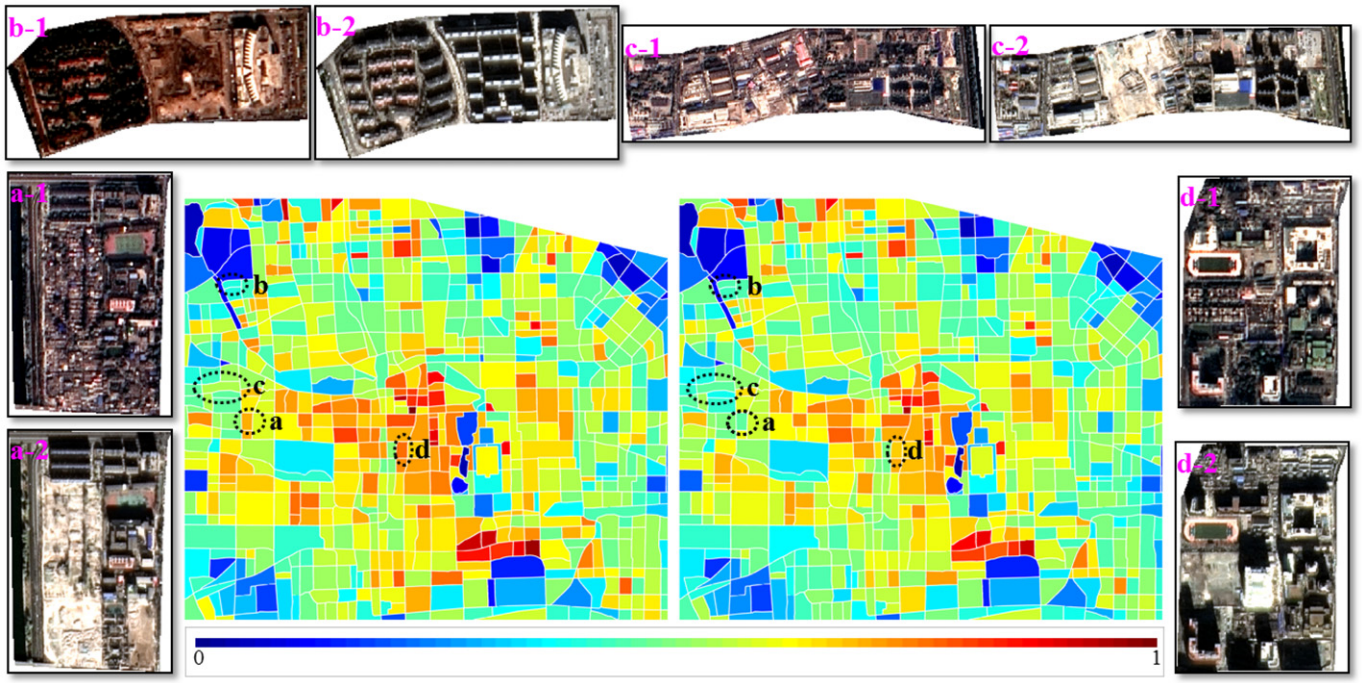


Fig. 12. Changes of the building coverage ratio for Beijing.

development (more housing space), old buildings were removed (Fig. 8(a)). In Fig. 8(b), the alley was demolished for redevelopment since the illegal buildings may result in security problems. Fig. 8 case (c) refers to the construction of high-rise buildings to address the increasing pressure of urban population (approximately 1 million new immigrants in Beijing during 2012–2013 (Beijing Municipal Statistics Bureau)). As a result of the seasonal effects, degraded vegetation and a shrunken lake are presented in Fig. 8(d) and (f), respectively. A case of vegetation removal for the future use of a construction site is shown in Fig. 8(e).

Likewise, in the result for Wuhan (Fig. 9), cases (a)–(d) and cases (e)–(h) correspond to changes of artificial and natural objects, respectively. Specifically, an example of building demolition is shown in Fig.

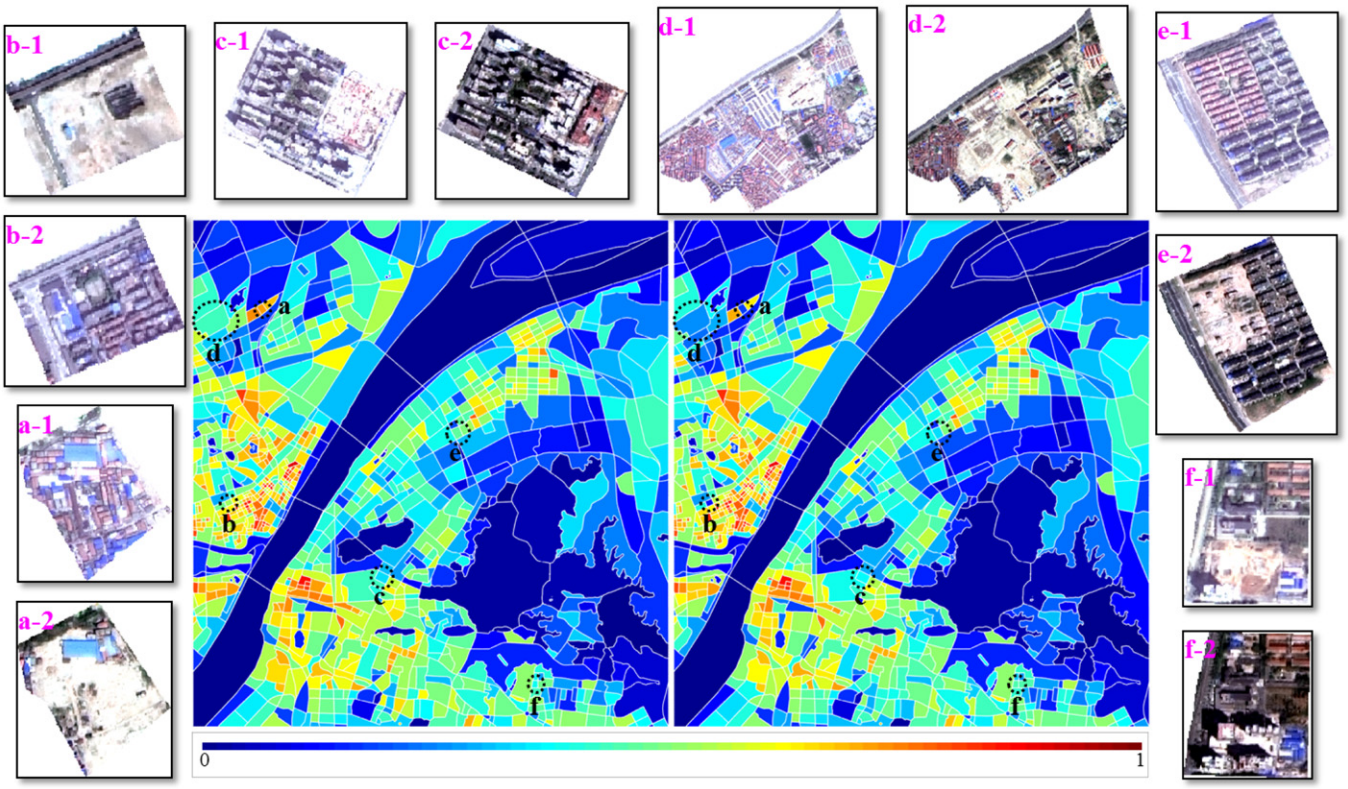


Fig. 13. Changes of the building coverage ratio for Wuhan.

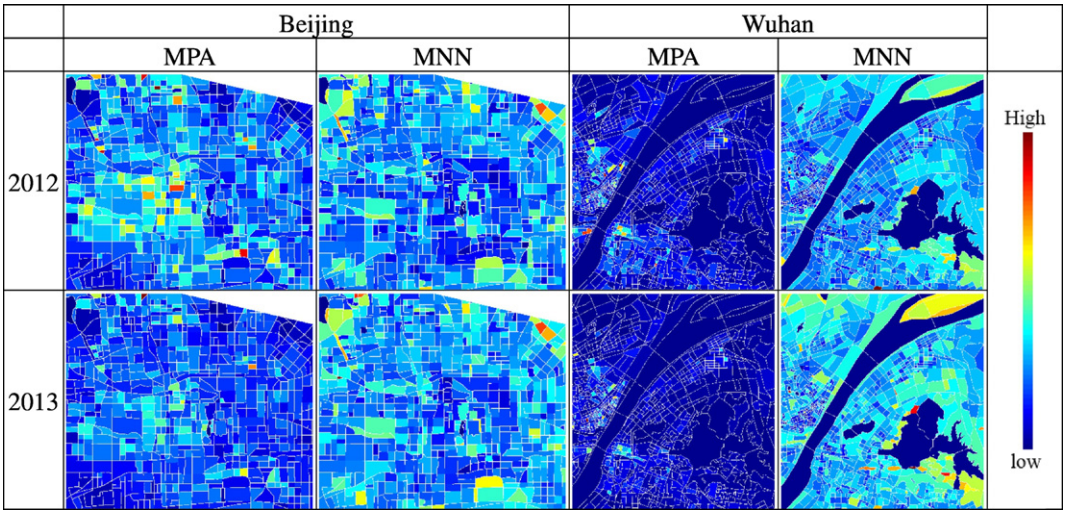


Fig. 14. Spatial maps of mean patch area (MPA) and mean nearest neighbor distance (MNN).

9(a), where an industrial zone was removed to make room for the construction of a residential area. In case (b), an urban village was demolished for the further construction of high-rise buildings, to provide more housing space. Cases (c) and (d) are also related to the construction of buildings. Examples for crop growth and water-level rise caused by the seasonal effect are given in (e) and (f), respectively. Case (g) shows the seasonal change for a patch of bare land (from soil to weeds), which has the potential for further land development and urban expansion. The change of water quality can also be monitored, as shown in Fig. 9(h), where water pollution caused by the spread of quantic plants can be observed.

4.4. Landscape change analysis

Urban green space plays an important role in urban ecosystems. VF at the landscape level was used to reveal the spatial distribution of green space. In Beijing, the city blocks where parks (e.g., Haidian Park, Yuyuantan Park, and Ditan Park) exist correspond to high VF values, and these areas are scattered across the city (Fig. 10). The city blocks with high VF in Wuhan include parks (e.g., Shouyi Park and Liberation Park), the East Lake scenic area, and marshland along the Yangtze River. Unlike Beijing, most of the blocks with higher VF in Wuhan (Fig. 11) are distributed in the scenic areas (e.g., the East Lake, parks, and Yangtze River), and are rarely in the city center and residential areas. This type of landscape configuration may lead to a high urban heat island intensity (UHII) (Li et al., 2011).

By focusing on the BCR at the landscape level, Beijing (Fig. 12) shows a typical radial distribution, whereas, in Wuhan (Fig. 13), most of the blocks with high BCR are located in the Old City. Representative examples of the changes of BCR are provided (Figs. 12 and 13). Again, the changes of BCR are associated with reclamation of the old residential areas (Fig. 13(b) and (e)), renovation of urban villages (Figs. 12(c), 13(a), (d)), alley removal (Fig. 12(a), (d)), and building construction for the real estate (Figs. 12(b), 13(c), (f)). Compared to the grid-based change detection, the analysis at the landscape level is more beneficial and convenient to the urban planning, as city block is the basic urban management unit of China.

In addition to the landscape composition (i.e., VF and BCR), we further analyzed the changes of the landscape configuration. A series of landscape metrics focusing on buildings were calculated. The spatial maps for certain metrics (i.e., MPA and MNN) are presented in Fig. 14. The mean values for the metrics measuring the configuration of buildings are presented in Table 9 to show the general trend in the change of building configuration for 2012–2013. The results reveal the following points: The decrease of the average LPI and increase of the average PD indicate greater fragmentation and more spatial heterogeneity of building patches, resulting from the demolition of old residential areas (e.g., Fig. 13(a) and (e)) and the construction of detached buildings (e.g., Fig. 13(f)). The increase of PD resulted in larger ED. Meanwhile, a slight decrease of the landscape shape index (MSI) indicates that the newly constructed building patches show smaller complexity in shape compared to the old ones. The patch area measures, e.g., MPA and SDPA, show a decreased trend, reflecting the increment of smaller-than-average building patches (e.g., detached and high apartments), as shown in Fig. 13(f), or reduction of bigger-than-average building patches (e.g., dense and compact residential areas), as shown in Fig. 13(a). MNN measures the average nearest neighbor distance between building patches. It was found that the minimum distance between building patches decreased by ~1.0 m during 2012–2013. SDMNN decreased by 1.36 and 1.10 for Beijing and Wuhan, respectively, showing that the variation of MNN values among different building patches became smaller. CI, measuring physical connectedness of building patches, showed a tiny reduction.

5. Discussions

5.1. Merits and limitations of the ZY-3 multi-view images

ZY-3 is China's first civilian high-resolution stereo mapping satellite, carrying triple linear cameras (forward, nadir, and backward modes). The capacity of the in-track stereo imaging ensures consistent image

Table 9
Mean values of the landscape metrics measuring the configuration of buildings.

	BJ 12	BJ 13	WH 12	WH 13
LPI	13.0	11.0	34.1	32.5
ED	510.3	591.7	644.1	699.0
MPA	0.14	0.08	0.28	0.21
SDPA	0.49	0.35	0.99	0.79
MSI	1.49	1.43	1.69	1.59
SDSI	0.85	0.82	1.28	1.17
MNN	11.3	9.8	10.8	9.8
SDMNN	8.1	6.8	7.7	6.6
PD	253.8	380.9	342.0	448.9
CI	97.4	96.9	97.7	97.3

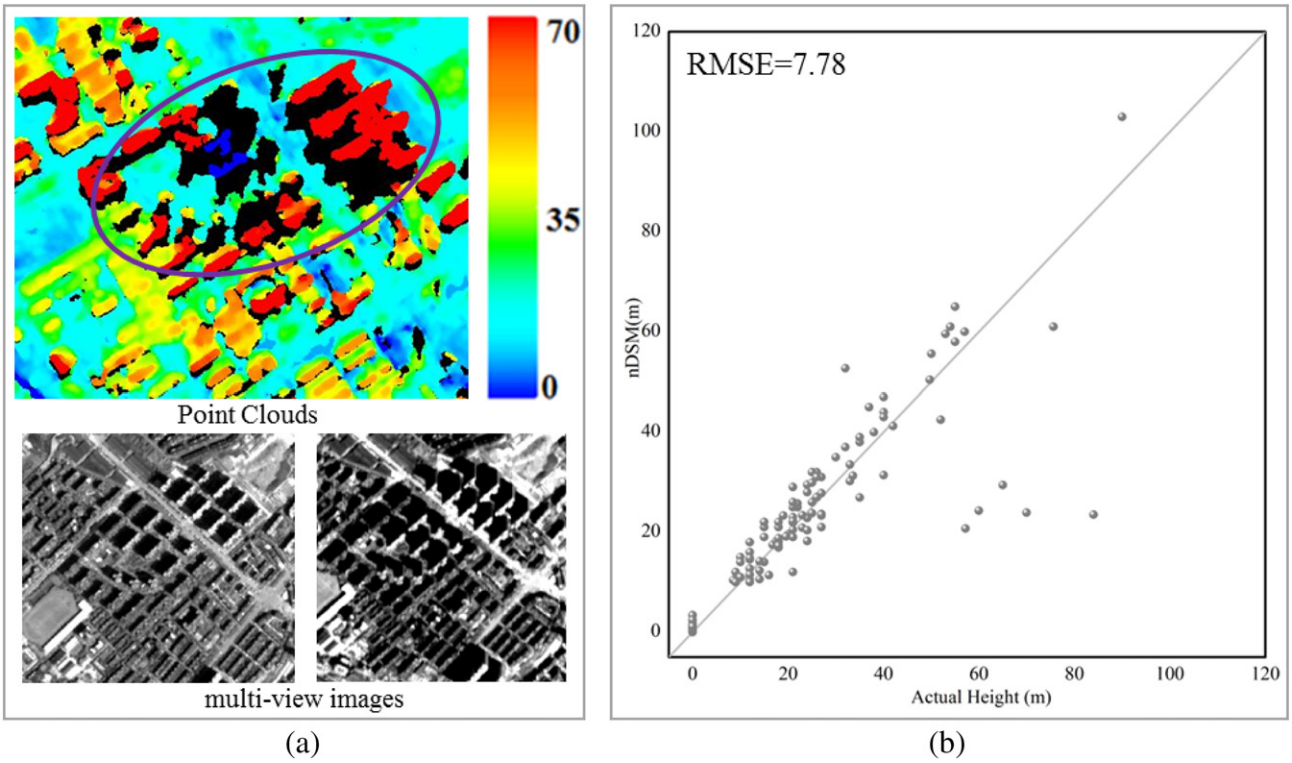


Fig. 15. Accuracy assessment of the nDSM: (a) an example of point clouds in a high-rise building area, with the purple circle outlining the matching errors; (b) scatter plot to depict the accuracy of the nDSM, where the vertical and horizontal axes represent the nDSM and actual height, respectively.

quality for the stereo pairs since the multi-view images can be acquired in almost identical illumination and weather conditions. Moreover, the base to height ratio of the ZY-3 satellite is designed to be 0.85–0.95, which is appropriate for producing DSMs. Therefore, ZY-3 stereo images are convenient for generating orthographic images and DSMs.

Image orthorectification is essential for reducing topographic effects. For multi-temporal change analysis using high spatial resolution data, a major difficulty or obstacle lies in the multi-temporal heterogeneity of the data, e.g., spatial mis-registration, parallax distortion for high architectures, and different viewing angles (Stumpf et al., 2014). Courtesy of the multi-view images of ZY-3, orthorectification can be conducted with stereo-photogrammetrically derived DSMs, which have a higher spatial resolution and are hence able to produce accurate and high spatial resolution orthographic images for time-series comparison and analysis.

DSMs are able to provide height information. Height information plays an essential role for urban land-cover mapping and 3D city modeling (Bergen et al., 2009; Chen et al., 2009; Heiden et al., 2012; Voltersen et al., 2014). However, in our study, although the nDSM could be employed to identify elevated land covers (i.e., buildings and trees), its direct application for precisely retrieving actual height values of the urban canopy was limited due to its accuracy. We compared the estimated building height derived from the nDSM and the actual height obtained from the urban planning department. The results (Fig. 15(b)) show that the generated nDSM was not accurate enough to estimate the building height (RMSE = 7.78 m). In particular, a number of buildings with a height larger than 50 m were underestimated. Fig. 15(a) shows an example of point clouds in a high-building area (height coded by color), where the purple circle outlines examples of such errors. These high buildings created large disparity and occlusion in the epipolar image space, which is usually problematic for narrow baseline matching algorithms. For Chinese cities, the buildings for residential and commercial use usually have a large height (for maximizing space utility) and are often mixed with low and old buildings, which is a typical urban landscape, as shown in Fig. 1. This complex 3D urban landscape increases the difficulty of multi-view image matching.

5.2. Performance of the sample migration strategy

The sample number and mapping performance in different stages of the sample migration were investigated (as shown in Table 10). To guarantee a fair comparison between the mapping performances, training samples with the same number (500 per class) were randomly selected from the generated sample sets at different stages. From the table, it is interesting to see that constraint (b) (i.e., reliable) plays a key role in improving the classification since the addition of this condition can significantly raise the accuracy in both study areas, i.e., the Kappa coefficient increased from 0.61 to 0.81 in Beijing, and from 0.83 to 0.89 in Wuhan.

In the sample migration method, a rigorous CVA threshold is adopted in order to make sure that the migrated samples are unchanged. In order to validate the sensitivity of mapping performance to selection of the CVA threshold, classification using different strict CVA thresholds (i.e., 0.001, 0.005, 0.01, 0.05, 0.1 and 0.2) were discussed. As can be seen in Table 11, with such strict threshold values,

Table 10
Sample number and mapping performance in different stages of the sample migration (OA = overall accuracy).

	Beijing			Wuhan		
	(a)	(a) + (b)	(a) + (b) + (c)	(a)	(a) + (b)	(a) + (b) + (c)
Sample #	3625346	726306	4568	2888179	1206213	4345
OA	69.4	84.5	85.4	85.9	91.1	92.0
Kappa	0.61	0.81	0.82	0.83	0.89	0.90

Note: (a) unchanged, (b) reliable, and (c) representative.

Table 11
Mapping performance using different CVA thresholds (OA = overall accuracy).

	Beijing						Wuhan					
Threshold	0.001	0.005	0.01	0.05	0.1	0.2	0.001	0.005	0.01	0.05	0.1	0.2
OA	86.0	85.4	85.3	84.4	84.8	84.7	91.6	92.0	91.5	90.5	90.6	90.4
Kappa	0.83	0.82	0.82	0.81	0.81	0.81	0.90	0.90	0.90	0.89	0.89	0.88

in both study areas, mapping performance is high and stable with Kappa coefficient larger than 0.81.

It should be noted that, in this study, the cloud-free images are chosen. For cloud-covered images, the cloud covered samples cannot satisfy the "unchanged" condition of the sample migration strategy, and hence will be not considered as training samples at time 2.

5.3. Scaling analysis for the grid size

The grid size for the hot-spot detection is an important parameter. When the specified grid size is small, some grids will be of a hybrid nature and some may consist partly of streets or buildings. A small grid size can capture the details, but it may be inadequate to characterize the neighborhood extent (e.g., the spatial pattern and arrangement of the land-cover classes). On the other hand, however, a larger grid can describe adequate contextual information, but it may fail to preserve the details. In this regard, the effect of the grid sizes was examined in terms of both detection accuracy and detail preservation. The maximum value of the grid size was set to 204 × 204 pixels, which is roughly equal to the average size of the city blocks generated by OpenStreetMap. In this research, four grid sizes—24 × 24, 51 × 51, 102 × 102, and 204 × 204 pixels—were considered and compared.

For the detection accuracy, the correctness of the results at different grid sizes is shown in Fig. 16(a). Mutual information entropy (Susaki et al., 2014), which is a direct measure of the difference between the entropy for image pairs, was used to evaluate the loss of details (Fig. 16(b)). A base image is required for calculation of the mutual information. The results with the smallest grid size (i.e., 24 pixels) were selected as the base image since they had the largest information entropy. The scaling analysis indicates that larger grid sizes (e.g., 51, 102, and 204 pixels) can generate higher correctness for the hot-spot change detection. The correctness sharply increased from 24 to 51, and became relatively stable when the size was larger than 51. On the other hand, as expected, larger grid sizes lead to smaller mutual information

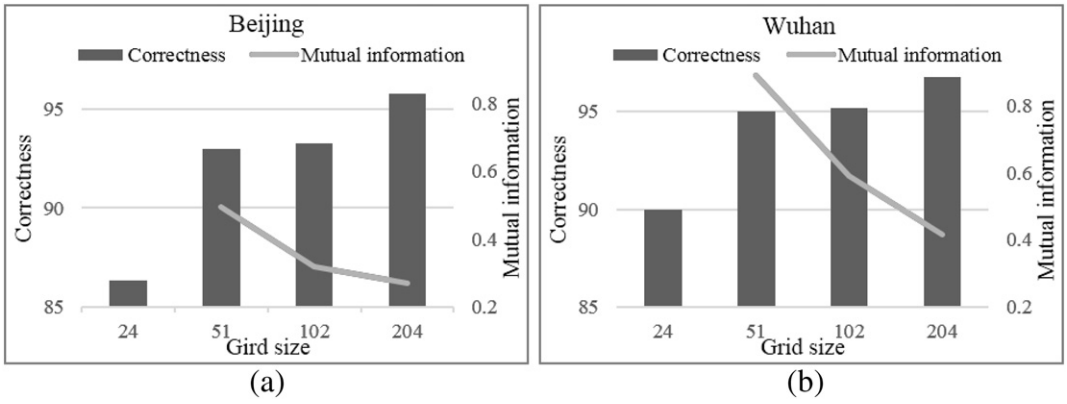


Fig. 16. Comparison between the change detection accuracy (correctness) and detail (mutual information entropy) at different grid sizes for (a) Beijing and (b) Wuhan. Note that the results with the smallest grid size (i.e., 24 pixels) were selected as the base image for calculation of the mutual information entropy.

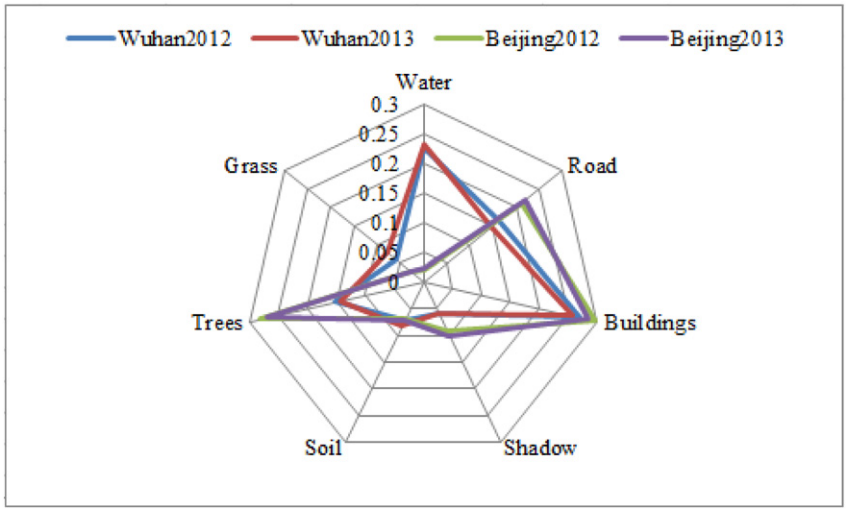


Fig. 17. Spider chart illustrating the proportions of the different land-cover types derived from ZY-3 in Wuhan and Beijing.

Table 12
Acquisition times of the ZY-3 and Landsat images for Beijing and Wuhan.

	Acquisition time (Beijing)		Acquisition time (Wuhan)	
ZY-3	2012-10-11	2013-11-28	2012-04-22	2013-08-12
Landsat	2012-10-24	2013-11-28	2012-03-14	2013-08-08

entropy, which signifies loss of details. Therefore, in this study, the grid size was set to 51 pixels in terms of both detection accuracy and detail preservation.

5.4. Comparison between Landsat and ZY-3 for detecting subtle urban changes

The proportions of the multi-temporal land-cover compositions, obtained from the high-resolution ZY-3 data, are depicted in a spider chart form in Fig. 17. Here, it can be clearly seen that only a small amount of changes occurred in both cities during 2012–2013. These subtle changes may be missed by coarse- or moderate-resolution data. We therefore chose Landsat images from Beijing and Wuhan, whose acquisition times were quite close to ZY-3 (Table 12), in order to compare their performances for monitoring subtle urban changes. For the Landsat change detection, the widely used method, CVA method (Yu et al., 2016) was adopted. Radiometric normalization between the multi-temporal data was performed before implementing the CVA algorithm. Specifically, the pseudo-invariant features method was employed for the radiometric normalization (Davies et al., 2016).

The hot-spot change detection at the grid level was used for comparison, and the change intensity maps are shown in Fig. 18. It can be seen that the results generated by Landsat are subject to a large number of false alarms and "salt-and-pepper" noise, whereas the results of ZY-3 are much better able to highlight the changes. Thereafter, the intensity maps of both sensors were binarized with an automatic thresholding strategy—minimum cross-entropy (Li and Lee, 1993)—and their accuracies were evaluated based on field investigation and visual interpretation of the high-resolution images (ZY-3) and Google Earth. The quantitative accuracy comparison is provided in Table 13, where (A) and (B) indicate the number of correctly detected changed grids among all the detected ones, respectively, for Landsat and ZY-3. In the case of Landsat, the correctness values are quite low, with only 34.5% (Beijing) and 28.8% (Wuhan) of the grids being successfully identified, and most of the detected hot spots are false alarms. The results are consistent with the phenomenon observed in Fig. 18. On the other hand, ZY-3 correctly extracts 93.0% (Beijing) and 95.0% (Wuhan) hot spots among all the detected ones. Moreover, it should be noted that only 40 (Beijing) and 221 (Wuhan) changed grids are successfully identified with the Landsat data, whereas 384 (Beijing) and 1885 (Wuhan) grids are correctly detected with ZY-3. This indicates that ZY-3 is able to monitor subtle urban changes more effectively, with much smaller commission errors.

The performances of ZY-3 and Landsat for change detection were further compared by a cross-checking. Among the 384 (Beijing) and 1885 (Wuhan) changed grids that are correctly identified by ZY-3, only 45 (11.7%) and 223 (11.8%) are captured by Landsat. On the other hand, however, 57.5% (Beijing) and 73.8% (Wuhan) grids, which

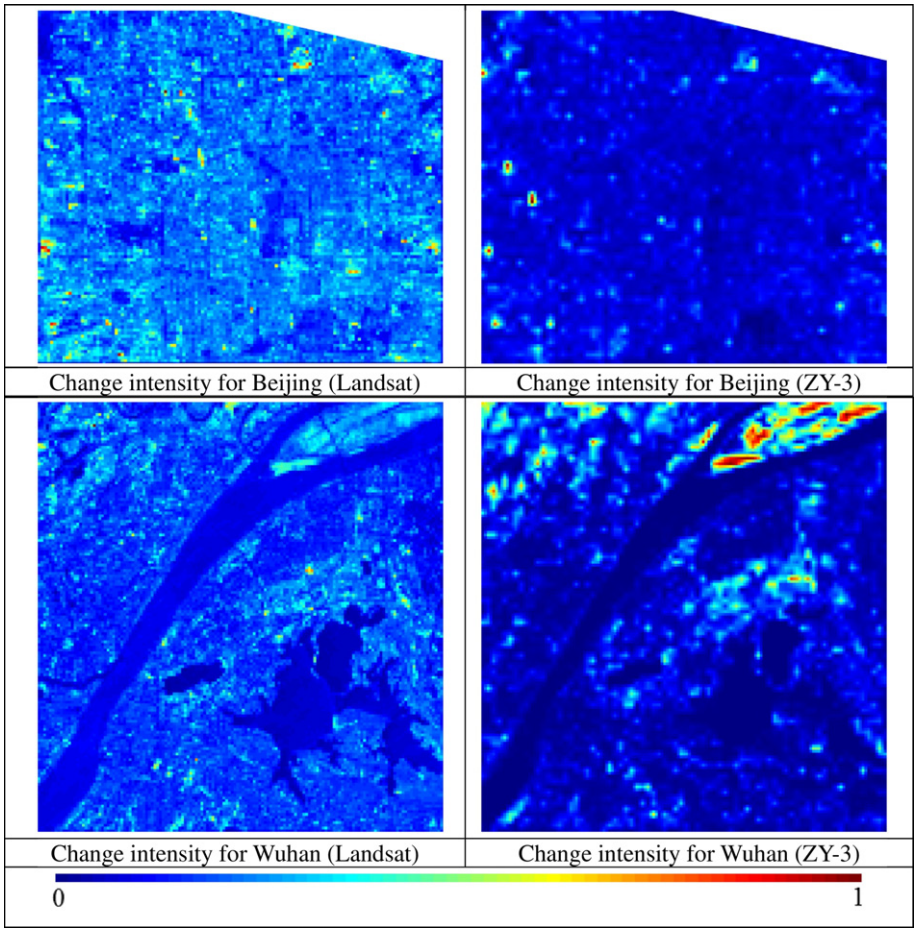


Fig. 18. Hot-spot change intensity derived from ZY-3 and Landsat data in Wuhan and Beijing.

Table 13
Comparison between ZY-3 and Landsat for the grid-based hot-spot change detection.

	Beijing		Wuhan	
	Accuracy (#)	Accuracy (%)	Accuracy (#)	Accuracy (%)
(A) Landsat	40/116	34.5	221/768	28.8
(B) ZY-3	384/413	93.0	1885/1984	95.0
(C) Landsat checked by ZY-3	45/384	11.7	223/1885	11.8
(D) ZY-3 checked by Landsat	23/40	57.5	163/221	73.8

are correctly extracted from Landsat, are identified by ZY-3. This is an evidence that a high-resolution sensor, e.g., ZY-3 (2.5 m), is essential for precisely detecting subtle urban changes, and that Landsat data, with a spatial resolution of 30 m, are not sensitive to most of the subtle changes occurring in urban areas.

The change detection results generated by both sensors are visually compared in Fig. 19, where the red and blue colors represents the hot spots only extracted by the Landsat and ZY-3 data, respectively, and the yellow color represents the hot spots identified by both sensors. In the result for Beijing, cases (a) and (b) correspond to building construction. It can be observed that although both sensors identify this changed area, the result generated by ZY-3 is more complete. Case (c) shows a typical example of building demolition, which was successfully marked by ZY-3 but missed by Landsat. Case (d) demonstrates a subtle urban change, where an architecture was constructed during 2012 and 2013. It is not surprising that this kind of change can only be captured by the high-resolution ZY-3 sensor. In addition, several examples are related to the false alarms generated by Landsat, e.g., cases (e) and (f), which were caused by the difference in the illumination. Concerning the result for Wuhan, both sensors can identify large-scale changes, e.g., cases (a) and (b). However, the Landsat data fail to detect slight changes, such as (c) and (d), associated with small-area removal and construction of buildings. Likewise, a number of false alarms can be observed from the Landsat result, induced by different illumination (case (e)) or variation of the roof color (case (f)).

6. Conclusions

This research focused on the subtle variations occurring in rapidly changing urban environments, which cannot be captured by coarse- or moderate-resolution remote sensing data. Such small changes are critically important for urban planning, environmental assessment, economic or demographic surveying, etc. Undoubtedly, high-resolution images are imperative for the monitoring of subtle urban changes. However, few studies have addressed urban change detection using high-resolution data, due to their multi-temporal heterogeneity, such as spatial mis-registration, parallax distortion for high architectures, and different viewing angles. In this study, we filled this gap and resolved these difficult problems by the use of the multi-view satellite data obtained by ZY-3, which is China's first civilian high-resolution three-line array stereo satellite (launched in January 2012). A notable advantage of the multi-view ZY-3 satellite data is their ability to generate multi-temporal

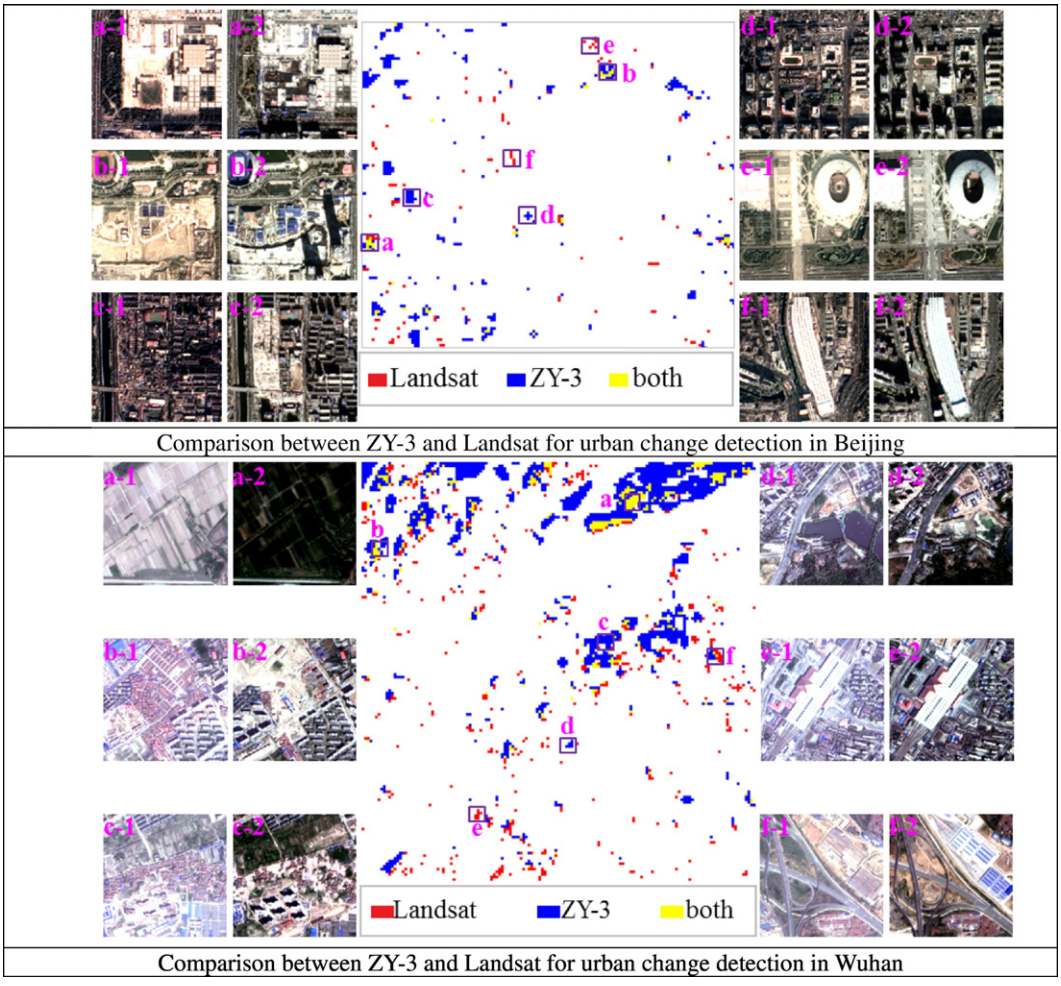


Fig. 19. Visual comparison between ZY-3 and Landsat for urban change detection.

orthographic images, which can solve the aforementioned problems, thus making it possible to monitor urban changes at a high spatial resolution.

In this research, we presented a multi-level framework for monitoring subtle urban changes using multi-view ZY-3 satellite data. This new framework was applied in two typical rapidly changing urbanized regions—Beijing and Wuhan—to test its generality and effectiveness. Some general conclusions can be made as follows:

1) The generation of orthographic images is a key step for avoiding the geometrical differences between multi-temporal high-resolution images, and for applying high-resolution data to change detection.
2) From the perspective of image interpretation, the presented method, integrating pixel- and object-based multiple features and rules, is effective for the classification of high-resolution images (85–87% for Beijing and 91–92% for Wuhan) and, hence, can guarantee the accuracy of the PCC. In addition, multi-temporal land-cover mapping can be efficiently performed with the sample migration strategy, achieving 83–85% change detection accuracy at the pixel level and 93–95% correctness at the grid level.
3) Pixel-level processing is able to indicate the change trajectories, but it suffers from "salt-and-pepper" noise. Thus, the results derived from the pixel level are further aggregated into the block level, which can highlight the urban changes and reveal the urban development patterns. The results demonstrated that although both cities are megacities of China, Beijing and Wuhan showed different urban change patterns. Wuhan remains at the period of rapid urban development, with a large amount of infrastructure construction. However, in Beijing, due to the policy of preserving the historical style of the Old City, most of the changed grids were distributed at the urban fringes, and the land development in the downtown (urban core) was restricted. Subsequently, the landscape-level change analysis indicated that the building patches in both cities became smaller and more fragmented. We also found that the minimum distance between building patches decreased by ~1.0 m during 2012–2013.
4) We found that Landsat data (30 m) were not capable of monitoring the subtle urban changes, due to the low detection accuracy (~30%), compared to the ~90% achieved by ZY-3 (2.5 m). This implies that high-resolution remote sensing images are indispensable for precise urban change analysis.

Acknowledgements

The work presented in this paper was supported by the National Key Research and Development Program of China under grant 2016YFB0501403, by the China Science Fund for Excellent Young Scholars under grant 41522110, and by the Foundation for the Author of National Excellent Doctoral Dissertation of PR China (FANEDD) under grant 201348.

References

Aguirre-Gutiérrez, J., Seijmonsbergen, A.C., Duivenvoorden, J.F., 2012. Optimizing land cover classification accuracy for change detection, a combined pixel-based and object-based approach in a mountainous area in Mexico. Appl. Geogr. 34, 29–37.
Aytekin, Ö., Ulusoy, İ., 2011. Automatic segmentation of VHR images using type information of local structures acquired by mathematical morphology. Pattern Recogn. Lett. 32, 1618–1625.
Bai, X., Chen, J., Shi, P., 2011. Landscape urbanization and economic growth in China: positive feedbacks and sustainability dilemmas. Environ. Sci. Technol. 46, 132–139.
Bai, X., Shi, P., Liu, Y., 2014. Realizing China's urban dream. Nature 509, 158–160.
Baus, P., Kováč, Urban P., Eva, K., Ivana, K., Jozef, K., 2014. Identification of interconnections between landscape pattern and urban dynamics—Case study Bratislava, Slovakia. Ecol. Indic. 42, 104–111.
Beijing Municipal Statistics Bureau, 2013. Beijing Statistical Yearbook. China Statistics Press, Beijing.
Benediktsson, J.A., Pesaresi, M., Amason, K., 2003. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. IEEE Trans. Geosci. Remote Sens. 41, 1940–1949.
Bergen, K., Goetz, S., Dubayah, R., Henebry, G., Hunsaker, C., Imhoff, M., Nelson, R., Parker, G., Radeloff, V., 2009. Remote sensing of vegetation 3-D structure for biodiversity and habitat: review and implications for lidar and radar spaceborne missions. J. Geophys. Res. Biogeosci. 2005–2012, 114.
Berger, C., Voltersen, M., Eckardt, R., Eberle, J., Heyer, T., Salepci, N., Hese, S., Schmullius, C., Tao, J., Auer, S., 2013. Multi-modal and multi-temporal data fusion: outcome of the 2012 GRSS data fusion contest. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 6, 1324–1340.
Bruzzone, L., Carlin, L., 2006. A multilevel context-based system for classification of very high spatial resolution images. IEEE Trans. Geosci. Remote Sens. 44, 2587–2600.
Chen, Y., Su, W., Li, J., Sun, Z., 2009. Hierarchical object oriented classification using very high resolution imagery and LIDAR data over urban areas. Adv. Space Res. 43, 1101–1110.
Cheng, J., Masser, I., 2003. Urban growth pattern modeling: a case study of Wuhan city, PR China. Landsc. Urban Plan. 62, 199–217.
Cheng, J., Zhou, J., 2015. Urban growth in a rapidly urbanized mega city: Wuhan. Urban Development Challenges, Risks and Resilience in Asian Mega Cities. Springer, pp. 301–322.
ChinaDaily, 2011. China vows tough measures to curb illegal land use. In. http://usa.chinadaily.com.cn/business/2011-01/21/content_11898511.html.
Davies, K.P., Murphy, R.J., Bruce, E., 2016. Detecting historical changes to vegetation in a Cambodian protected area using the Landsat TM and ETM + sensors. Remote Sens. Environ. 187, 332–344.
Dell'Acqua, F., Gamba, P., Ferrari, A., Palmason, J., Benediktsson, J., Arnason, K., 2004. Exploiting spectral and spatial information in hyperspectral urban data with high resolution. IEEE Geosci. Remote Sens. Lett. 1, 322–326.
Dewan, A.M., Yamaguchi, Y., 2009. Land use and land cover change in Greater Dhaka, Bangladesh: using remote sensing to promote sustainable urbanization. Appl. Geogr. 29, 390–401.
Du, P., Xia, J., Zhang, W., Tan, K., Liu, Y., Liu, S., 2012. Multiple classifier system for remote sensing image classification: a review. Sensors 12, 4764–4792.
Du, P., Liu, S., Xia, J., Zhao, Y., 2013. Information fusion techniques for change detection from multi-temporal remote sensing images. Inf. Fusion 14, 19–27.
El-Kawy, O.A., Rød, J., Ismail, H., Suliman, A., 2011. Land use and land cover change detection in the western Nile delta of Egypt using remote sensing data. Appl. Geogr. 31, 483–494.
Faryadi, S., Taheri, S., 2009. Interconnections of urban green spaces and environmental quality of Tehran. Int. J. Environ. Res. 3, 199–208.
Fraser, C.S., Hanley, H.B., 2003. Bias compensation in rational functions for IKONOS satellite imagery. Photogramm. Eng. Remote. Sens. 69, 53–57.
Haralick, R.M., Shanmugam, K., Dinstein, I.H., 1973. Textural features for image classification. IEEE Trans. Syst. Man Cybern. 610–621.
He, C., Shi, P., Li, J., Chen, J., Pan, Y., Li, J., Zhuo, L., Ichinose, T., 2006. Restoring urbanization process in China in the 1990s by using non-radiance-calibrated DMSP/OLS nighttime light imagery and statistical data. Chin. Sci. Bull. 51, 1614–1620.
Heiden, U., Heldens, W., Roessner, S., Segl, K., Esch, T., Mueller, A., 2012. Urban structure type characterization using hyperspectral remote sensing and height information. Landsc. Urban Plan. 105, 361–375.
Hirschmuller, H., 2008. Stereo processing by semi-global matching and mutual information. IEEE Trans. Pattern Anal. Mach. Intell. 30, 328–341.
Huang, H., Wei, Y.D., 2014. Intra-metropolitan location of foreign direct investment in Wuhan, China: institution, urban structure, and accessibility. Appl. Geogr. 47, 78–88.
Huang, X., Zhang, L., 2010. Comparison of vector stacking, multi-SVMs fuzzy output, and multi-SVMs voting methods for multiscale VHR urban mapping. IEEE Geosci. Remote Sens. Lett. 7, 261–265.
Huang, X., Zhang, L., 2011. A multidirectional and multiscale morphological index for automatic building extraction from multispectral GeoEye-1 imagery. Photogramm. Eng. Remote. Sens. 77, 721–732.
Huang, X., Zhang, L., 2012a. Morphological building/shadow index for building extraction from high-resolution imagery over urban areas. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 5, 161–172.
Huang, X., Zhang, L., 2012b. A multiscale urban complexity index based on 3D wavelet transform. For spectral-spatial feature extraction and classification: an evaluation on the 8-channel Worldview-2 imagery. Int. J. Remote Sens. 33, 2641–2656.
Huang, X., Zhang, L., 2013. An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery. IEEE Trans. Geosci. Remote Sens. 51, 257–272.
Huang, X., Lu, Q., Zhang, L., 2014a. A multi-index learning approach for classification of high-resolution remotely sensed images over urban areas. ISPRS J. Photogramm. Remote Sens. 90, 36–48.
Huang, X., Lu, Q., Zhang, L., Plaza, A., 2014b. New postprocessing methods for remote sensing image classification: a systematic study. IEEE Trans. Geosci. Remote Sens. 52, 7140–7159.
Huang, X., Liu, H., Zhang, L., 2015. Spatiotemporal detection and analysis of urban villages in mega city regions of China using high-resolution remotely sensed imagery. IEEE Trans. Geosci. Remote Sens. 53, 3639–3657.
Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ. 83, 195–213.
Hüllermeier, E., Brinker, K., 2008. Learning valued preference structures for solving classification problems. Fuzzy Sets Syst. 159, 2337–2352.
Hur, M., Nasar, J.L., Chun, B., 2010. Neighborhood satisfaction, physical and perceived naturalness and openness. J. Environ. Psychol. 30, 52–59.
Hussain, M., Chen, D., Cheng, A., Wei, H., Stanley, D., 2013. Change detection from remotely sensed images: from pixel-based to object-based approaches. ISPRS J. Photogramm. Remote Sens. 80, 91–106.
Johnson, R.D., Kasischke, E.S., 1998. Change vector analysis: a technique for the multispectral monitoring of land cover and condition. Int. J. Remote Sens. 19, 411–426.

Koroso, N.H., van der Molen, P., Tuladhar, A.M., Zevenbergen, J.A., 2013. Does the Chinese market for urban land use rights meet good governance principles? Land Use Policy 30, 417–426.
Kressler, F.P., Steinnocher, K.T., 1999. Detecting land cover changes from NOAA-AVHRR data by using spectral mixture analysis. Int. J. Appl. Earth Obs. Geoinf. 1, 21–26.
Lasanta, T., Vicente-Serrano, S.M., 2012. Complex land cover change processes in semiarid Mediterranean regions: an approach using Landsat images in northeast Spain. Remote Sens. Environ. 124, 1–14.
Leitão, A.B., Miller, J., Ahern, J., McGarigal, K., 2012. Measuring Landscapes: A Planner's Handbook. Island Press.
Leslie, E., Sugiyama, T., Ierodiaconou, D., Kremer, P., 2010. Perceived and objectively measured greenness of neighbourhoods: are they measuring the same thing? Landsc. Urban Plan. 95, 28–33.
Li, W., 2016. Failure by design – national mandates and agent control of local land use in China. Land Use Policy 52, 518–526.
Li, C.H., Lee, C., 1993. Minimum cross entropy thresholding. Pattern Recogn. 26, 617–625.
Li, J., Song, C., Cao, L., Zhu, F., Meng, X., Wu, J., 2011. Impacts of landscape structure on surface urban heat islands: a case study of Shanghai, China. Remote Sens. Environ. 115, 3249–3263.
Li, X., Li, W., Middel, A., Harlan, S.L., Brazel, A.J., Turner Ii, B.L., 2016. Remote sensing of the surface urban heat island and land architecture in Phoenix, Arizona: combined effects of land composition and configuration and cadastral–demographic–economic factors. Remote Sens. Environ. 174, 233–243.
Lin, G.C., 2007. Reproducing spaces of Chinese urbanisation: new city-based and land-centred urban transformation. Urban Stud. 44, 1827–1855.
Liu, T., Yang, X., 2015. Monitoring land changes in an urban area using satellite imagery, GIS and landscape metrics. Appl. Geogr. 56, 42–54.
Liu, Y., Yu, X., Huang, J.X., An, A., 2011. Combining integrated sampling with SVM ensembles for learning from imbalanced datasets. Inf. Process. Manag. 47, 617–631.
Liu, Z., He, C., Zhang, Q., Huang, Q., 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.
Liu, Y., Luo, T., Liu, Z., Kong, X., Li, J., Tan, R., 2015. A comparative analysis of urban and rural construction land use change and driving forces: implications for urban–rural coordination development in Wuhan, Central China. Habitat Int. 47, 113–125.
Long, H., 2014. Land use policy in China: introduction. Land Use Policy 40, 1–5.
Lu, D.-d., Yao, S.-m., Li, G.-p., Liu, H., Gao, X.-l., 2007. Comprehensive analysis of the urbanization process based on China's conditions. Econ. Geogr. 6, 002.
Lunetta, R.S., Knight, J.F., Ediriwickrema, J., Lyon, J.G., Worthy, L.D., 2006. Land-cover change detection using multi-temporal MODIS NDVI data. Remote Sens. Environ. 105, 142–154.
Ma, T., Zhou, C., Pei, T., Haynie, S., Fan, J., 2012. Quantitative estimation of urbanization dynamics using time series of DMSP/OLS nighttime light data: a comparative case study from China's cities. Remote Sens. Environ. 124, 99–107.
Mangai, U.G., Samanta, S., Das, S., Chowdhury, P.R., 2010. A survey of decision fusion and feature fusion strategies for pattern classification. IETE Tech. Rev. 27, 293–307.
McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int. J. Remote Sens. 17, 1425–1432.
McGarigal, K., Marks, B.J., 1995. Spatial pattern analysis program for quantifying landscape structure. Gen. Tech. Rep. PNW-GTR-351. US Department of Agriculture, Forest Service, Pacific Northwest Research Station.
McGarigal, K., Cushman, S.A., Neel, M.C., Ene, E., 2002. FRAGSTATS: spatial pattern analysis program for categorical maps. URL:. www.umass.edu/landeco/research/fragstats/fragstats.html.
Merrell, P., Akbarzadeh, A., Wang, L., Mordohai, P., 2007. Real-time visibility-based fusion of depth maps. Proceedings, pp. 1–8.
Mertes, C.M., Schneider, A., Sulla-Menashe, D., Tatem, A., Tan, B., 2015. Detecting change in urban areas at continental scales with MODIS data. Remote Sens. Environ. 158, 331–347.
Mura, M.D., Benediktsson, J.A., Waske, B., Bruzzone, L., 2010. Morphological attribute profiles for the analysis of very high resolution images. IEEE Trans. Geosci. Remote Sens. 48, 3747–3762.
Pacifici, F., Del Frate, F., 2010. Automatic change detection in very high resolution images with pulse-coupled neural networks. IEEE Geosci. Remote Sens. Lett. 7, 58–62.
Pacifici, F., Chini, M., Emery, W.J., 2009. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. Remote Sens. Environ. 113, 1276–1292.
Pagot, E., Pesaresi, M., 2008. Systematic study of the urban postconflict change classification performance using spectral and structural features in a support vector machine. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 1, 120–128.
Pelletier, C., Valero, S., Inglada, J., Champion, N., Dedieu, G., 2016. Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. Remote Sens. Environ. 187, 156–168.
Peña-Barragán, J.M., Ngugi, M.K., Plant, R.E., Six, J., 2011. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. Remote Sens. Environ. 115, 1301–1316.
Peng, S., Piao, S., Ciais, P., Friedlingstein, P., Ottle, C., Bréon, F.o.-M., Nan, H., Zhou, L., Myneni, R.B., 2011. Surface urban heat island across 419 global big cities. Environ. Sci. Technol. 46, 696–703.
People'sDaily, 2010. Hoarded land put under the spotlight again. In. http://en.people.cn/90001/90778/90860/7111427.html.
Plaza, A., Benediktsson, J.A., Boardman, J.W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., 2009. Recent advances in techniques for hyperspectral image processing. Remote Sens. Environ. 113, S110–S122.
Qin, R., 2014. Change detection on LOD 2 building models with very high resolution spaceborne stereo imagery. ISPRS J. Photogramm. Remote Sens. 96, 179–192.
Qin, R., Fang, W., 2014. A hierarchical building detection method for very high resolution remotely sensed images combined with DSM using graph cut optimization. Photogramm. Eng. Remote. Sens. 80, 873–883.
Rajadell, O., Garcia-Sevilla, P., Dinh, V.C., Duin, R.P.W., 2014. Improving hyperspectral pixel classification with unsupervised training data selection. IEEE Geosci. Remote Sens. Lett. 11, 656–660.
Rogge, D., Rivard, B., Zhang, J., Sanchez, A., Harris, J., Feng, J., 2007. Integration of spatial–spectral information for the improved extraction of endmembers. Remote Sens. Environ. 110, 287–303.
Rothermel, M., Wenzel, K., Fritsch, D., Haala, N., 2012. Sure: photogrammetric surface reconstruction from imagery. Proceedings LC3D Workshop, Berlin, pp. 1–9.
Salomons, E.M., Pont, M.B., 2012. Urban traffic noise and the relation to urban density, form, and traffic elasticity. Landsc. Urban Plan. 108, 2–16.
Schneider, A., 2012. Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach. Remote Sens. Environ. 124, 689–704.
Schneider, A., Friedl, M.A., Potere, D., 2009. A new map of global urban extent from MODIS satellite data. Environ. Res. Lett. 4, 044003.
Serna, A., Marcotegui, B., 2014. Detection, segmentation and classification of 3D urban objects using mathematical morphology and supervised learning. ISPRS J. Photogramm. Remote Sens. 93, 243–255.
Sesnie, S.E., Gessler, P.E., Finegan, B., Thessler, S., 2008. Integrating Landsat TM and SRTM-DEM derived variables with decision trees for habitat classification and change detection in complex neotropical environments. Remote Sens. Environ. 112, 2145–2159.
Sexton, J.O., Song, X.-P., Huang, C., Channan, S., Baker, M.E., Townshend, J.R., 2013. 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. 129, 42–53.
Shalaby, A., Tateishi, R., 2007. Remote sensing and GIS for mapping and monitoring land cover and land-use changes in the northwestern coastal zone of Egypt. Appl. Geogr. 27, 28–41.
Shin, H.B., 2010. Urban conservation and revalorisation of dilapidated historic quarters: the case of Nanluoguxiang in Beijing. Cities 27, S43–S54.
Song, X.-P., Sexton, J.O., Huang, C., 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.
Stow, D.A., Chen, D.M., 2002. Sensitivity of multitemporal NOAA AVHRR data of an urbanizing region to land-use/land-cover changes and misregistration. Remote Sens. Environ. 80, 297–307.
Stumpf, A., Malet, J.-P., Allemand, P., Ulrich, P., 2014. Surface reconstruction and landslide displacement measurements with Pléiades satellite images. ISPRS J. Photogramm. Remote Sens. 95, 1–12.
Susaki, J., Kajimoto, M., Kishimoto, M., 2014. Urban density mapping of global megacities from polarimetric SAR images. Remote Sens. Environ. 155, 334–348.
Tarantino, C., Adamo, M., Lucas, R., Blonda, P., 2016. Detection of changes in semi-natural grasslands by cross correlation analysis with WorldView-2 images and new Landsat 8 data. Remote Sens. Environ. 175, 65–72.
Tatem, A.J., Noor, A.M., Von Hagen, C., Di Gregorio, A., Hay, S.I., 2007. High resolution population maps for low income nations: combining land cover and census in East Africa. PLoS One 2, e1298.
Taubenböck, H., Wiesner, M., Felbier, A., Marconcini, M., Esch, T., Dech, S., 2014. New dimensions of urban landscapes: the spatio-temporal evolution from a polynuclei area to a mega-region based on remote sensing data. Appl. Geogr. 47, 137–153.
Tewkesbury, A.P., Comber, A.J., Tate, N.J., Lamb, A., Fisher, P.F., 2015. A critical synthesis of remotely sensed optical image change detection techniques. Remote Sens. Environ. 160, 1–14.
Tian, J., Reinartz, P., 2013. Fusion of multi-spectral bands and DSM from Worldview-2 Stereo imagery for building extraction. Urban Remote Sensing Event (JURSE), 2013 Joint. IEEE, pp. 135–138.
Trias-Sanz, R., Stamon, G., Louchet, J., 2008. Using colour, texture, and hierarchial segmentation for high-resolution remote sensing. ISPRS J. Photogramm. Remote Sens. 63, 156–168.
Tuia, D., Pasolli, E., Emery, W., 2011. Using active learning to adapt remote sensing image classifiers. Remote Sens. Environ. 115, 2232–2242.
UN, 2014. World Urbanization Prospects: The 2014 Revision. United Nations, New York.
Voltersen, M., Berger, C., Hese, S., Schmullius, C., 2014. Object-based land cover mapping and comprehensive feature calculation for an automated derivation of urban structure types at block level. Remote Sens. Environ. 154, 192–201.
Wang, T., Zhang, G., Li, D., Tang, X., Jiang, Y., Pan, H., Zhu, X., Fang, C., 2014. Geometric accuracy validation for ZY-3 satellite imagery. IEEE Geosci. Remote Sens. Lett. 11, 1168–1171.
Wen, D., Huang, X., Zhang, L., Benediktsson, J.A., 2016. A novel automatic change detection method for urban high-resolution remotely sensed imagery based on multiindex scene representation. IEEE Trans. Geosci. Remote Sens. 54, 609–625.
Woźniak, M., Graña, M., Corchado, E., 2014a. A survey of multiple classifier systems as hybrid systems. Inf. Fusion 16, 3–17.
Woźniak, M., Graña, M., Corchado, E., 2014b. A survey of multiple classifier systems as hybrid systems. Inf. Fusion 3–17.
Wuhan Municipal Statistics Bureau, 2013. Wuhan Statistical Yearbook. China Statistics Press, Beijing.
Yager, R.R., 1992. On the specificity of a possibility distribution. Fuzzy Sets Syst. 50, 279–292.
Yu, B., Liu, H., Wu, J., Hu, Y., Zhang, L., 2010. Automated derivation of urban building density information using airborne LiDAR data and object-based method. Landsc. Urban Plan. 98, 210–219.

Yu, W., Zhou, W., Qian, Y., Yan, J., 2016. A new approach for land cover classification and change analysis: integrating backdating and an object-based method. Remote Sens. Environ. 177, 37–47.

Yuan, F., Sawaya, K.E., Loeffelholz, B.C., Bauer, M.E., 2005. Land cover classification and change analysis of the twin cities (Minnesota) metropolitan area by multitemporal Landsat remote sensing. Remote Sens. Environ. 98, 317–328.

Zhang, C., 2015. Applying data fusion techniques for benthic habitat mapping and monitoring in a coral reef ecosystem. ISPRS J. Photogramm. Remote Sens. 104, 213–223.

Zhang, L., Huang, X., Huang, B., Li, P., 2006. A pixel shape index coupled with spectral information for classification of high spatial resolution remotely sensed imagery. IEEE Trans. Geosci. Remote Sens. 44, 2950–2961.

Zhang, L., Zhang, L., Tao, D., Huang, X., 2012. On combining multiple features for hyperspectral remote sensing image classification. IEEE Trans. Geosci. Remote Sens. 50, 879–893.

Zhao, L., Tang, J., Yu, X., Li, Y., Mi, S., Zhang, C., 2012. Content-based remote sensing image retrieval using image multi-feature combination and SVM-based relevance feedback. Lecture Notes in Electrical Engineering.

Zhong, P., Wang, R., 2007. A multiple conditional random fields ensemble model for urban area detection in remote sensing optical images. IEEE Transactions on Geoscience & Remote Sensing 45, 3978–3988.

Zhou, D., Zhao, S., Liu, S., Zhang, L., Zhu, C., 2014. Surface urban heat island in China's 32 major cities: spatial patterns and drivers. Remote Sens. Environ. 152, 51–61.