



High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform

Xiaoping Liu^a, Guohua Hu^b, Yimin Chen^{b,*}, Xia Li^{b,*}, Xiacong Xu^a, Shaoying Li^c, Fengsong Pei^d, Shaojian Wang^a

^a School of Geography and Planning, Guangdong Key Laboratory for Urbanization and Geo-simulation, Sun Yat-sen University, Guangzhou 510275, PR China

^b Key Lab of Geographic Information Science (Ministry of Education), School of Geographic Sciences, East China Normal University, Shanghai 200241, PR China

^c School of Geographical Sciences, Guangzhou University, Guangzhou 510006, PR China

^d School of Urban and Environmental Sciences, Jiangsu Normal University, Xuzhou 221116, PR China

ARTICLE INFO

Keywords:

NUACI
Global urban land
Google Earth Engine

ABSTRACT

Timely and accurate delineation of global urban land is fundamental to the understanding of global environmental changes. However, most of the contemporary global urban land maps have coarse resolutions and are available for one or two years only. In this study, we developed the multi-temporal global urban land maps based on Landsat images for the 1990–2010 period with a five-year interval ('Urban land' in these maps refers to 'impervious surface', i.e., artificial cover and structures such as pavement, concrete, brick, stone and other man-made impenetrable cover types). We proposed the method of Normalized Urban Areas Composite Index (NUACI) and utilized the Google Earth Engine to facilitate the global urban land classifications from an extensive number of Landsat images. The global level's overall accuracy, producer's accuracy and user's accuracy for our mapping results are 0.81–0.84, 0.50–0.60 and 0.49–0.61, respectively. The Kappa values are 0.43–0.50 at the global level, and ~0.33 (in China) and ~0.42 (in the U.S.) at the country level. By analyzing the presented dataset, we found that the world's urban land area had increased from 450.97 ± 1.18 thousand km² in 1990 to 747.05 ± 1.50 thousand km² in 2010, reaching a global coverage of 0.63%. China, the United States and India together (14% of the world's terrestrial area in total) contributed almost 43% of the total increase of global urban land area. A free download link for these data is attached at the end of this paper.

1. Introduction

Urban land has far-reaching impacts on societies and environments located beyond the city boundaries. Accurate and timely information about global urban land is essential for the research of land cover change, hydrologic dynamics, carbon cycles and climate change (Schneider et al., 2010). A variety of urban land definitions exist in literature, resulting in the varying estimates of the world's urban land from 0.45–0.65% to 2–3% (Liu et al., 2014). To avoid confusion, the term 'urban land' in this paper refers to 'impervious surface', i.e., artificial cover and structures such as pavement, concrete, brick, stone and other man-made impenetrable cover types (Chen et al., 2015).

In response to the growing demand for global urban land data, since 2000, several maps have been developed using remote sensing techniques (Potere et al., 2009; Schneider et al., 2010). The data sources of these maps include Moderate Resolution Imaging Spectroradiometer (MODIS), Defense Meteorological Satellite Program-Operational Line-scan System (DMSP-OLS) and Landsat. Most of these maps have the

coarse spatial resolutions of 500–1000 m (Bartholomé and Belward, 2005; Elvidge et al., 2007; Schneider et al., 2010). Additionally, they are only available for a single year or two. Therefore, it is difficult to attain a clear picture of global urban land expansion over a long historical period.

As finer-resolution images, such as the Landsat series, are becoming increasingly available, it is feasible to refine the global mapping of urban land by using these image sources. There are extensive regional applications of Landsat images for urban expansion mapping. Representative examples include Bagan and Yamagata (2012) for Toyko, Li et al. (2015) for Beijing, Zhang and Weng (2016) for the Pearl River Delta, Sexton et al. (2013) and Song et al. (2016) for the Washington, D.C.–Baltimore metropolitan region, etc. In particular, Sexton et al. (2013) and Song et al. (2016) adopt an advanced 'continuous' characterization of urban land, which can better represent the fuzzy boundaries between urban land and the surrounding environments.

In this study, we present a new multi-temporal global urban land data product developed by using the Landsat images during the

* Corresponding authors at: School of Geographic Sciences, East China Normal University, 500 Dongchuan Rd., Shanghai 200241, PR China.
E-mail addresses: ymchen@geo.ecnu.edu.cn (Y. Chen), lixia@geo.ecnu.edu.cn (X. Li).

1990–2010 period, at five-year intervals. At present, there are two Landsat-based global land cover products, i.e., FROM-GLC by Gong et al. (2013) and GlobeLand30 by Chen et al. (2015), respectively. They provide the most up-to-date information of comprehensive global land cover types, although they are only available for the years of 2000 and 2010. The production of our data product aims to mitigate the shortage of high-resolution multi-temporal global urban land maps and provide reliable information for global urban research.

The production of the multi-temporal global urban land maps from Landsat images raises three critical challenges. The first one is the image selection scheme, which should ensure the optimal global coverage of images with minimal cloud cover. However, varying weather conditions from one region to another decrease the availability of acceptable images, especially for tropical and sub-tropical regions. The second challenge relates to the establishment of an efficient platform for image data processing. The third challenge is the development of adequate image classification methods with satisfactory performance. A wide range of methods have already been proposed for the automated prediction of urban land from Landsat images (Bagan and Yamagata, 2012; Li et al., 2015; Zhang and Weng, 2016). However, most of them are regional applications, whereas their applicability for global uses is usually less understood.

To confront the first and second challenges noted above, we utilized the Google Earth Engine (GEE, <https://earthengine.google.org>) to exploit its computational power as well as to facilitate the image selection process. To address the third challenge, we adopted an index-based method for the automated prediction of urban land. Previous research has revealed that index-based methods are effective and efficient for applications with different image sources, such as Landsat (Patel et al., 2015) and DMSP-OLS (Liu et al., 2015; Yu et al., 2014). These methods can also be easily adapted to applications in different geographical regions. In this study, we used the Normalized Urban Areas Composite Index (NUACI), which was proposed recently by Liu et al. (2015). NUACI articulates several important normalized indices to discriminate urban land from other non-urban land cover types. This new index has been applied to urban land mapping in China during the 2000–2010 period (available via <http://www.geosimulation.cn/ISA-China.htm>).

The NUACI-based global urban land classification is carried out in several steps. First, the world's terrestrial area is stratified based on the urban ecoregion scheme proposed by Schneider et al. (2010). This stratification approach regards local similarities with respect to the ecological, cultural and social elements of cities at the global scale. Second, urban land sample is collected to conduct the region-specific calibration for NUACI. Third, urban land is classified by segmenting the NUACI images with the region-specific thresholds. Finally, classification accuracy is evaluated by comparing the results against a global reference dataset and other mainstream global urban land products.

2. Methods

2.1. Normalized Urban Areas Composite Index (NUACI)

The development of NUACI integrates the Normalized Difference Water Index (NDWI) (McFeeters, 1996), the Normalized Difference Vegetation Index (NDVI) (Chen and Cihlar, 1996) and the Normalized Difference Built-up Index (NDBI) (Zha et al., 2003). The calculations of these indices are specified below:

$$NDWI = (Green_{band2} - NIR_{band4}) / (Green_{band2} + NIR_{band4}) \quad (1)$$

$$NDVI = (NIR_{band4} - RED_{band3}) / (NIR_{band4} + RED_{band3}) \quad (2)$$

$$NDBI = (SWIR_{band5} - NIR_{band4}) / (SWIR_{band5} + NIR_{band4}) \quad (3)$$

Fig. 1 shows the representative urban land pixels along with five other types of non-urban pixels collected from a sub-tropical city, Guangzhou, and two desert cities, Phoenix and Riyadh (2000 pixels per

type per city). It is evident that the urban land pixels form a group clustering around the regions with large NDBI values and small NDWI and NDVI values. For the desert cities, the urban land pixels are also clearly separated from the barren land pixels in the projected planes of NDVI-NDBI and NDWI-NDBI. Therefore, the indices of NDBI, NDWI and NDVI can be integrated to assist the urban land classification:

$$NUACI = U_{NTL} \times (1 - \sqrt{(NDWI_i - a_{NDWI})^2 + (NDVI_i - b_{NDVI})^2 + (NDBI_i - c_{NDBI})^2}) \quad (4)$$

where U_{NTL} represents the binary mask obtained from DMSP-OLS nighttime lights; $U_{NTL} = 0$ if pixels are classified as non-urban according to the DMSP-OLS nighttime lights threshold, and $U_{NTL} = 1$ otherwise; $NDWI_i$, $NDVI_i$ and $NDBI_i$ are the corresponding values of NDWI, NDVI and NDBI in pixel i ; and a_{NDWI} , b_{NDVI} and c_{NDBI} are the corresponding mean values of NDWI, NDVI and NDBI in the urban land sample.

The binary mask (U_{NTL}) is obtained using the DMSP-OLS nighttime lights data that are concurrent with the Landsat images. The threshold method is used to segment the DMSP-OLS nighttime lights images into 'urban' and 'non-urban,' resulting in rough estimates of the spatial extent of built-up areas. This method is similar to that used by Shi et al. (2014a,b), which regards the DMSP-OLS pixels with values > 0 as potential urban pixels. After the segmentation, 'non-urban' pixels can be excluded with certainty, whereas potential 'urban' pixels were retained for subsequent classification.

The NUACI-based classification method requires calibration for the representative centroids (the triplet $(a_{NDWI}, b_{NDVI}, c_{NDBI})$ in Eq. (4)) and the thresholds for NUACI segmentation. The centroids and thresholds may vary across different regions throughout the world due to heterogeneous physical and socioeconomic characteristics. Therefore, we divided the world map into different regions and calibrated the centroids and thresholds separately for each region. We applied the global stratification scheme of urban ecoregions developed by Schneider et al. (2010) (Fig. 2(a)). This stratification scheme takes three elements into account including a biome designation characterizing general climate and vegetation, urban topology differences and the economic level defined by per capita gross domestic product (GDP). It should be noted that the 16th urban ecoregion is covered by permanent ice and snow, and thus it was excluded in the subsequent analysis.

Within each urban ecoregion, multiple individual cities are selected according to their population size and economic status. The urban extent within each individual city is defined by buffering the city core with an average distance of 20–25 km, depending on the actual physical size of the city. For each urban ecoregion, the selected cities are randomly assigned into three different groups of the same size (Fig. 2(a)):

- (1) 'Centroid sites': For these sites, we collected urban land sample and estimated the mean NDWI, NDVI and NDBI values, which were further used to estimate the centroids.
- (2) 'Threshold sites': We generated the NUACI images using the estimated centroids. Then, we calibrated the threshold values with the reference data for these sites. The reference data were obtained by the object-based classification of Landsat images for these sites and a subsequent double-checking correction procedure with the concurrent Google Earth images. We calibrated the thresholds by experimenting with different thresholds until the one with the smallest error was found. The errors were assessed using the indicators of quantity disagreement and allocation disagreement (Pontius and Millones, 2011):

$$P_{ij} = \left(\frac{n_{ij}}{\sum_{i=1}^J n_{ij}} \right) \left(\frac{N_i}{\sum_{i=1}^J N_i} \right) \quad (5)$$

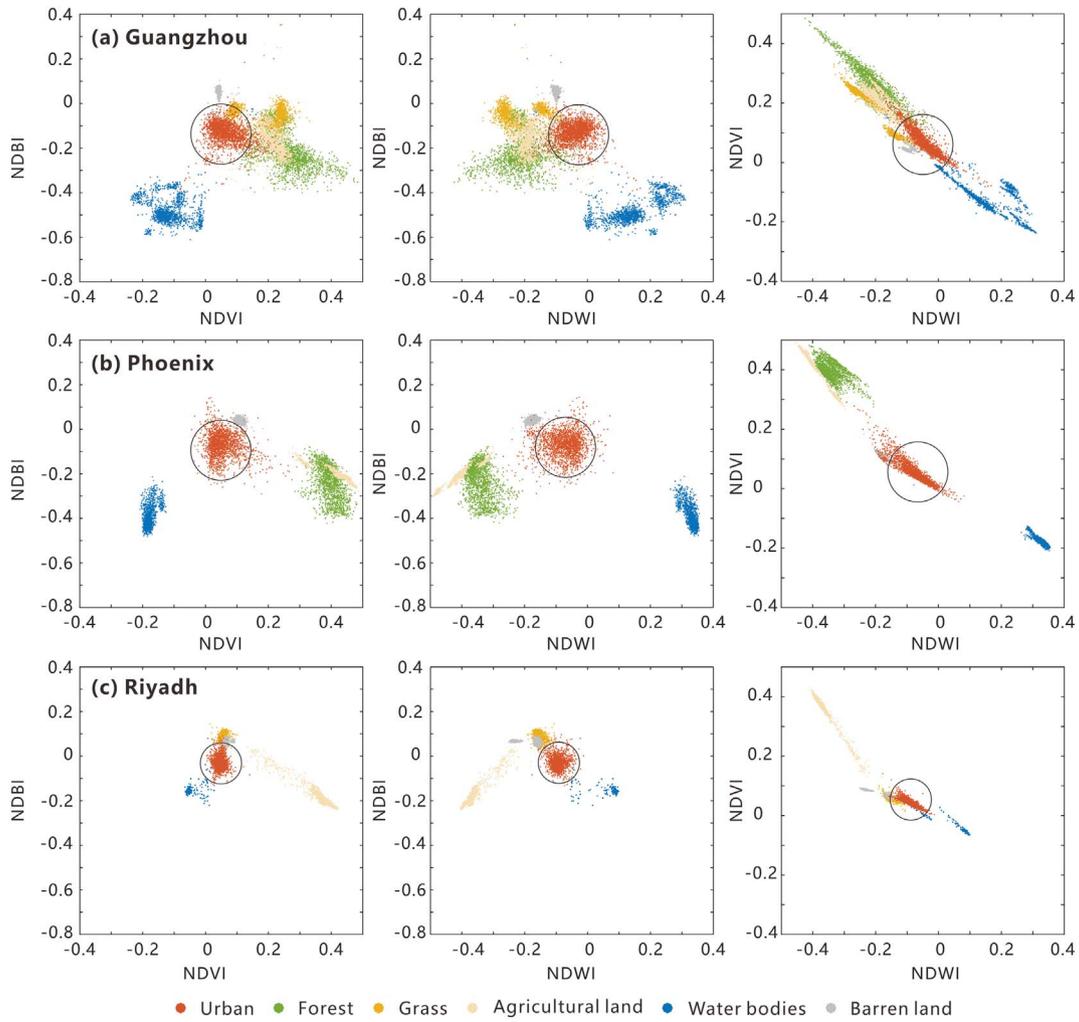


Fig. 1. Distributions of representative pixels for different land cover types in the feature space composed of NDWI, NDVI and NDBI. These pixels were collected from the Landsat images of 122/044, 037/037 and 165/043, respectively.

$$Q = \frac{1}{2} \sum_{g=1}^J q_g = \frac{1}{2} \sum_{g=1}^J \left| \left(\sum_{i=1}^J p_{ig} \right) - \left(\sum_{j=1}^J p_{gj} \right) \right| \quad (6)$$

$$A = \frac{1}{2} \sum_{g=1}^J a_g = \frac{1}{2} \sum_{g=1}^J 2 \min \left[\left(\sum_{i=1}^J p_{ig} \right) - p_{gg}, \left(\sum_{j=1}^J p_{gj} \right) - p_{gg} \right] \quad (7)$$

$$D = Q + A \quad (8)$$

where J is the number of land types, whereby $J = 2$ is applied in our analysis (i.e., ‘urban’ and ‘non-urban’); n_{ij} is the number of pixels classified as type i and referenced as type j ; N_i is the total number of pixels in type i ; q_g and a_g are the quantity disagreement and the allocation disagreement for type g , respectively; Q and A are the overall quantity disagreement and the overall allocation disagreement; the overall error D can be obtained by the sum of Q and A .

(3) ‘Testing sites’: We generated the NUACI images for these sites using the calibrated centroids, and then applied the calibrated thresholds for the urban land classification. Finally, we validated the classification against the reference data (they are obtained similarly to those mentioned in ‘Threshold sites’) using the indicators of quantity disagreement and allocation disagreement (Eqs. (5)–(8)).

2.2. Google Earth Engine Platform

Google Earth Engine is a cloud-based platform for scientific analysis and visualization of petabyte-scale geospatial datasets. It stores several decades of historical images and scientific datasets. We used the Landsat 5 TM calibrated top-of-atmosphere reflectance data as the primary image source, with the constraint cloud score of < 20 .

To utilize more cloud-free images, we extended the image selection dates to a period from two years before the target year to two years after. An example can be found via: <https://code.earthengine.google.com/980b294eac1c83d08fee10698de24286> (A Google account is required). It shows that our selection strategy can effectively address the cloud-free image shortage in tropical areas such as India and Indonesia. However, a large part of Russia has no Landsat image coverage in 1990s due to the issues of ground receiving stations, and hence there are missing data in this area for our mapping results in 1990. After the Landsat image selection procedure, the yearly average values of NDWI, NDVI and NDBI were estimated for each region. The results were used to estimate the region-specific centroids with the collected sample. Based on the estimated region-specific centroids and the nighttime lights data masks, the NUACI images can then be generated. Later, urban land pixels in each region were obtained using the region-specific NUACI segmentation thresholds.

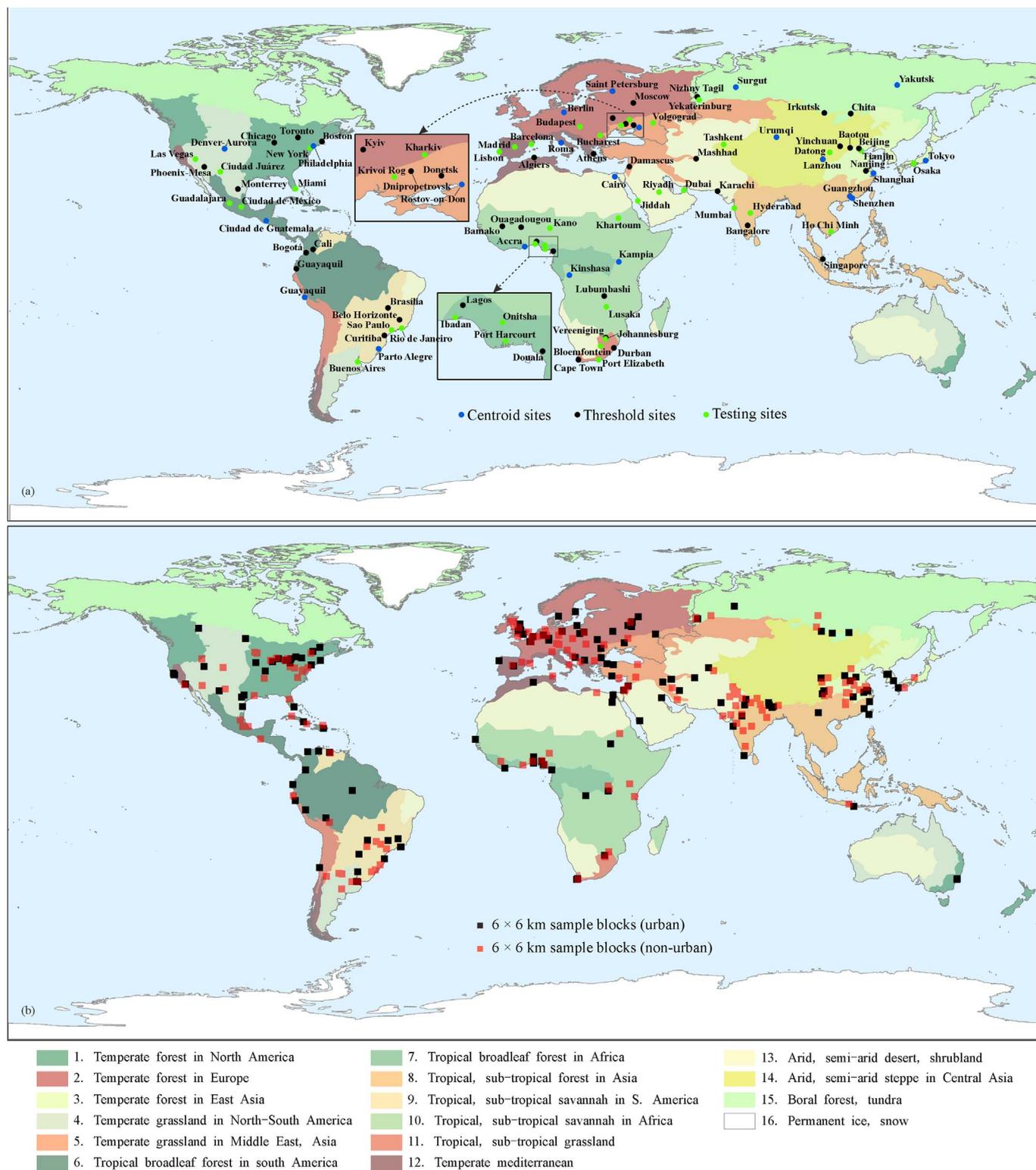


Fig. 2. The locations of (a) calibration and testing sites and (b) 6 × 6 km reference data blocks in each urban ecoregion.

2.3. Quality evaluation and area estimation

To rigorously assess the accuracy of our global urban land classifications, we prepared a multi-temporal reference dataset based on a stratified random sampling design (Olofsson et al., 2012; Pengra et al., 2015; Stehman et al., 2012). Specifically, we first partitioned the world's terrestrial areas using 6 × 6 km blocks. In the next step, we used the LandScan 2010 data and adopted the criterion of > 1000 persons/

km² (= 36,000 per block) to extract the 'urban' blocks (~15,000 'urban' blocks globally) (Olofsson et al., 2012). Then, we set the total sample size as 150 blocks according to cost and precision considerations. The 150 sample blocks first were proportionally allocated in each stratum. Subsequent manual adjustments were then applied to slightly increase the sample size for strata with too few blocks (e.g., #10). Finally, the sample blocks were randomly picked according to the adjusted sample sizes. The same allocation procedure was also applied to the sampling

Table 1
The number of sample blocks in each urban eco-region (each table entry represents equal numbers of urban and non-urban blocks).

U.E.	Proportional allocation	Manual adjustment	U.E.	Proportional allocation	Manual adjustment
1	20	18	9	9	9
2	25	27	10	2	4
3	16	14	11	5	5
4	11	9	12	11	9
5	9	8	13	9	10
6	8	8	14	5	6
7	5	7	15	4	5
8	11	11			

U.E. = urban ecoregion.

of ‘non-urban’ blocks with the same size of the ‘urban’ blocks (150) (Table 1 and Fig. 2(b)).

In each sample block, we performed manual interpretation to label the urban land distributions. For each block, we first implemented object-based image segmentation on the corresponding Landsat image, and overlaid the resulting image objects (polygons) with the concurrent Google Earth images so that the true land cover types (‘urban’ or ‘non-urban’) in the objects could be interpreted. An example is shown in Fig. 3. The polygons with red outlines are resulting image objects using the Landsat image. The yellow polygons are examples of the labeled ‘urban’ objects according to manual interpretation with Google Earth images. All the labeled objects are checked by independent working groups that do not participate in the image manual interpretation, and corrections are also made to the object labels if needed.

In the image selection procedure, we chose the Google Earth images with dates as close as possible to those of the Landsat images. The image selection follows the priorities of similar dates, the same season, the same year and adjacent years. For example, given a sample block with the date of June 4th, 1990, if there are no available Google Earth images with similar dates, the period of May to July or in the whole year of 1990, then other images (if available) in the adjacent years of 1991, 1992 or 1993 are considered. However, there are regions where historical Google Earth images are not available. In these cases, the sample blocks are labeled ‘missing reference images.’ If a sample block is labeled ‘missing reference images’ for more than two target years, then this block is excluded from the reference dataset and a new block

is randomly selected following the same sampling procedures discussed above.

By using this reference dataset, we performed a per-pixel accuracy assessment to validate our urban land classifications using the metrics of overall accuracy (OA), producer’s accuracy (PA), user’s accuracy (UA) and Cohen’s Kappa. We also estimated the standard deviations and the confidence intervals at 95% significance level for these metrics (Congalton and Green, 2008). Moreover, this reference dataset is also used as a basis for making a comparative analysis between our mapping results and other contemporary global urban land products, including GlobeLand30, MODIS500m, GLC2000, GRUMP and IMPSA.

Because the global urban land classifications contain errors, urban land area estimates calculated using the simple pixel counting method are inadequate. Therefore, we followed the ‘good practices’ suggested by Olofsson et al. (2014) and estimated the urban land areas using the error matrices obtained from the reference data blocks. The area estimation is based on \hat{p}_{urban} , the proportion of urban land area estimated from the reference data, and its estimated standard error $S(\hat{p}_{urban})$. Denoting the mapped urban land area as A_{urban} , the estimated urban land area with 95% confidence intervals is then estimated as: $\hat{A}_{urban} \pm 1.96 \times S(\hat{A})$, where $\hat{A}_{urban} = A_{urban} \times \hat{p}_{urban}$ and $S(\hat{A}) = A_{urban} \times S(\hat{p}_{urban})$.

3. Results and discussion

3.1. NUACI images

We collected 15,000 urban pixels (1000 pixels per urban ecoregion) to estimate the region-specific centroids for generating the NUACI images. Fig. 4 shows the histograms of NDWI, NDVI and NDBI for these pixels. All distributions except some cases (NDBI in #7, #14 and #15) are roughly approximated to a thin Gaussian shape, implying that these pixels are tightly clustered in the feature space. The distribution shapes of the NDBI in #7, #14 and #15 are analogous to a weighted-sum of two Gaussian distributions, indicating that these urban pixels are grouped into two clusters along the NDBI component. Therefore, these pixels can be separated from other land cover types using proper segmentation thresholds.

Fig. 4 also suggests that most of the NDWI and NDBI mean values are negative, while most of the NDVI mean values are positive. Moreover, these values are close to 0, indicating that urban pixels in different urban ecoregions are clustered at roughly the same location in the

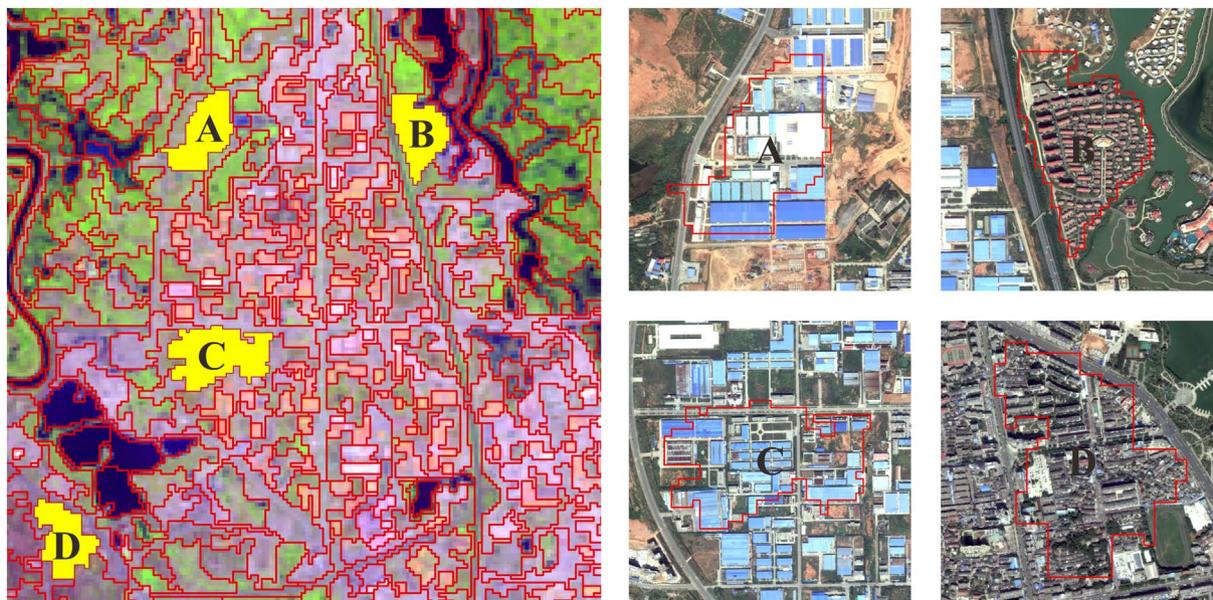


Fig. 3. An example of the production of reference data based on image segmentation and manual interpretation.

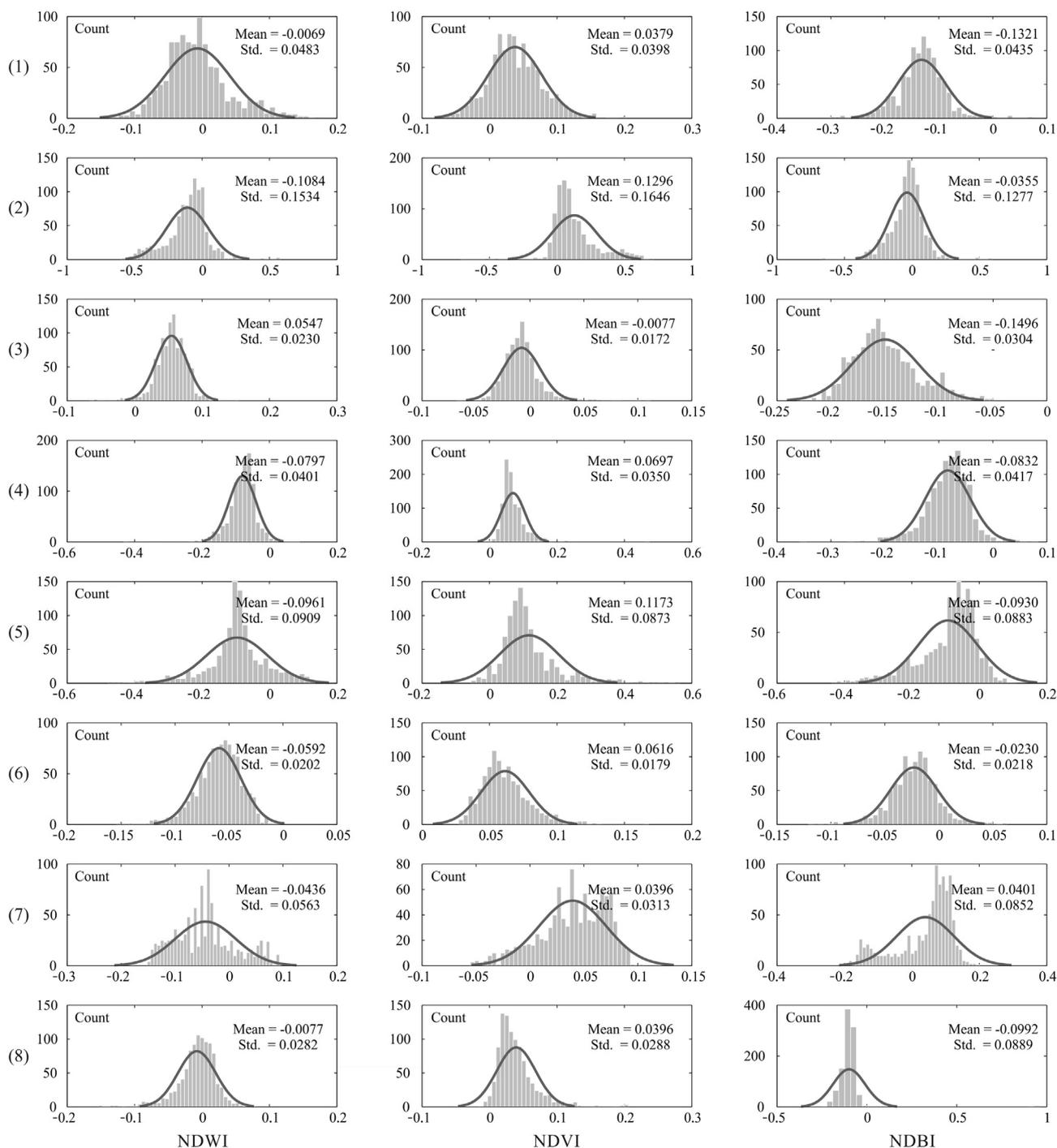


Fig. 4. The statistical distributions of the urban land sample with respect to NDWI, NDVI and NDBI, taking the results pertaining to 2000 as an example. The number shown on the left corresponds to the index of the urban ecoregion in Fig. 2(a).

feature space. The standard deviations of these three indices are consistently small across different urban ecoregions, particularly in regions #1, #6, #11 and #13. Therefore, the urban pixels in these four regions are the most homogeneous, and we can expect a steep trough in the segmentation threshold selection for these four regions (as confirmed by Fig. 6).

A subset of the generated multi-temporal NUACI images for some representative cities are shown in Fig. 5. These results successfully reflect the artificial elements and constructions (e.g., residential blocks and road networks) even in the areas with mixed land cover types. One can easily inspect the urban land expansion in these cities, such as Tokyo and Shanghai.

3.2. Urban land classification

As noted earlier, we obtained the NUACI thresholds at the ‘threshold sites’ of each urban ecoregion. Using the year 2000 as an example, Fig. 6 shows the relationships between the overall error ($Q + A$) and the varying NUACI thresholds in each of the urban ecoregions. The histograms of the NDVI, NDBI and NDWI components for regions #1, #6, #11 and #13 are aggregative, implying that the urban land pixels of these four regions are the most homogeneous. In contrast, in the regions of #7, #14 and #15, for example, the threshold curves inherit no steep troughs since the histograms of these regions are relatively dispersive. Table 2 suggests that the calibrated thresholds vary from

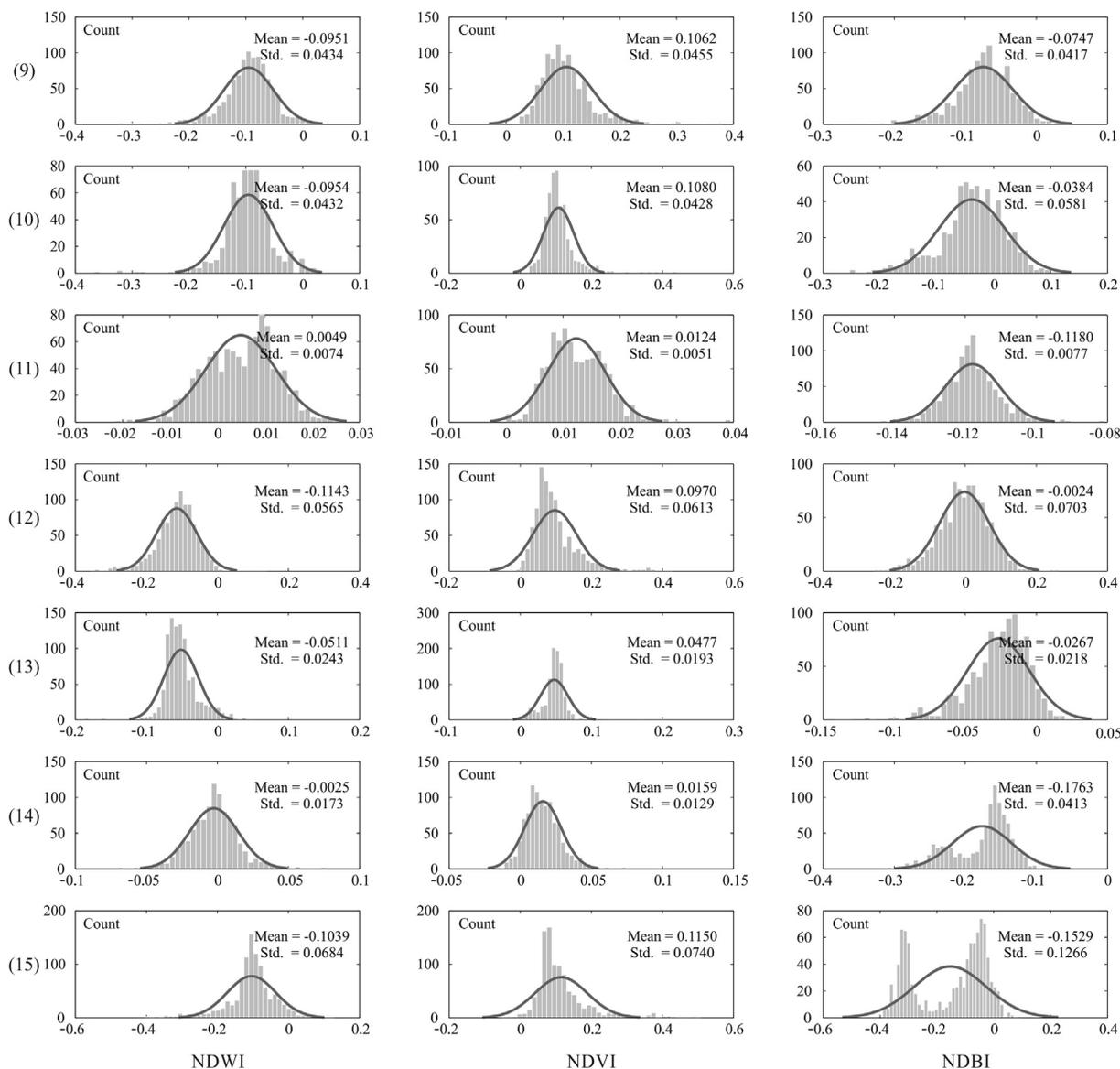


Fig. 4. (continued)

0.53 to 0.87, reflecting the heterogeneous geographical conditions of different urban ecoregions. The thresholds in the urban ecoregions #3 and #13 are the largest (0.87 and 0.80, respectively), whereas those in the urban ecoregions #9 and #14 are the smallest (0.57 and 0.53, respectively). The differences for the other urban ecoregions are relatively small (between 0.61 and 0.76).

Fig. 7 shows the urban land classifications for the representative cities in the year 2000. The results not only show the continuous urbanized areas but also successfully include the small, scattered settlements surrounded by different land cover types such as bare soil, shrubland, harvested farmland and desert, which are easily confused with urban land.

Some limitations of the proposed approach can also be found according to these results. For example, the water surfaces around Moscow are misclassified as urban land. Similar errors are also found in the results for Nanjing, where some aquaculture land along the Yangtze River is misclassified as urban land. These errors are probably caused by inadequate NUACI threshold values for segmentation. Refinements for the NUACI thresholds are required to further improve the results for these situations. Additionally, the proposed approach is not robust to mountain shadows. This limitation is clearly revealed in the results for Monterrey, in which mountain shadows cause some classification

errors.

3.3. Validation and comparison

Table A1 in the appendix provides the full validation statistics (including overall accuracy, producer's accuracy, user's accuracy, Kappa, and the 95% confidence intervals) of our mapping results. Both the producer's and user's accuracies are large (over 0.80 or even up to 0.95) for the non-urban type. For urban land, however, the global producer's accuracy increases from 0.50 in 1990 to 0.60 in 2010, while the global user's accuracy also increases from 0.49 in 1990 to 0.61 in 2010. Smaller producer's accuracy and user's accuracy (0.2–0.4) for urban land are found in several urban ecoregions, such as #3, #5, #8, #10 and #11, for the years 1990, 1995 and 2000. Therefore, major omission errors and commission errors exist in these urban ecoregions. The producer's accuracies for urban land in urban ecoregions #13 and #14 are also relatively small (< 0.4) in the years 1990 and 1995, implying that a larger number of urban pixels in these regions are misclassified as non-urban type.

The global Kappa values for the mapping results are 0.43 (1990), 0.43 (1995), 0.48 (2000), 0.49 (2005) and 0.50 (2010). Urban ecoregions #6, #9 and #12 have the greatest Kappa values (> 0.60),

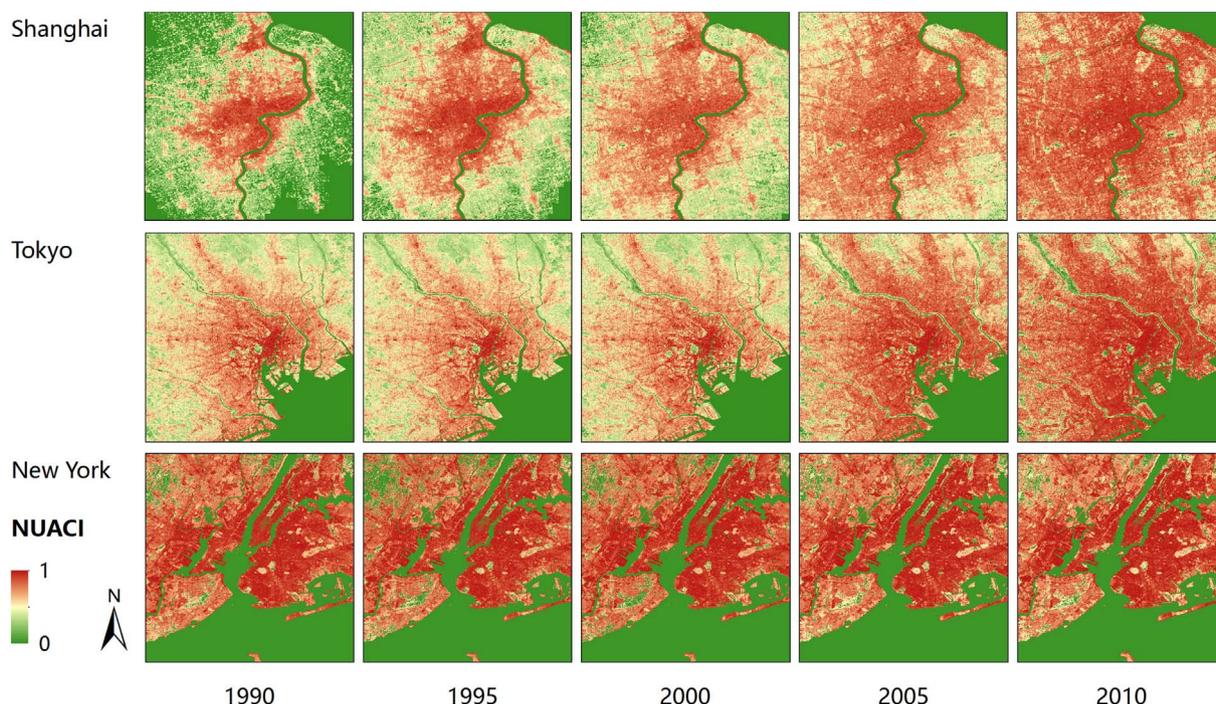


Fig. 5. The generated multi-temporal NUCI images in the representative cities.

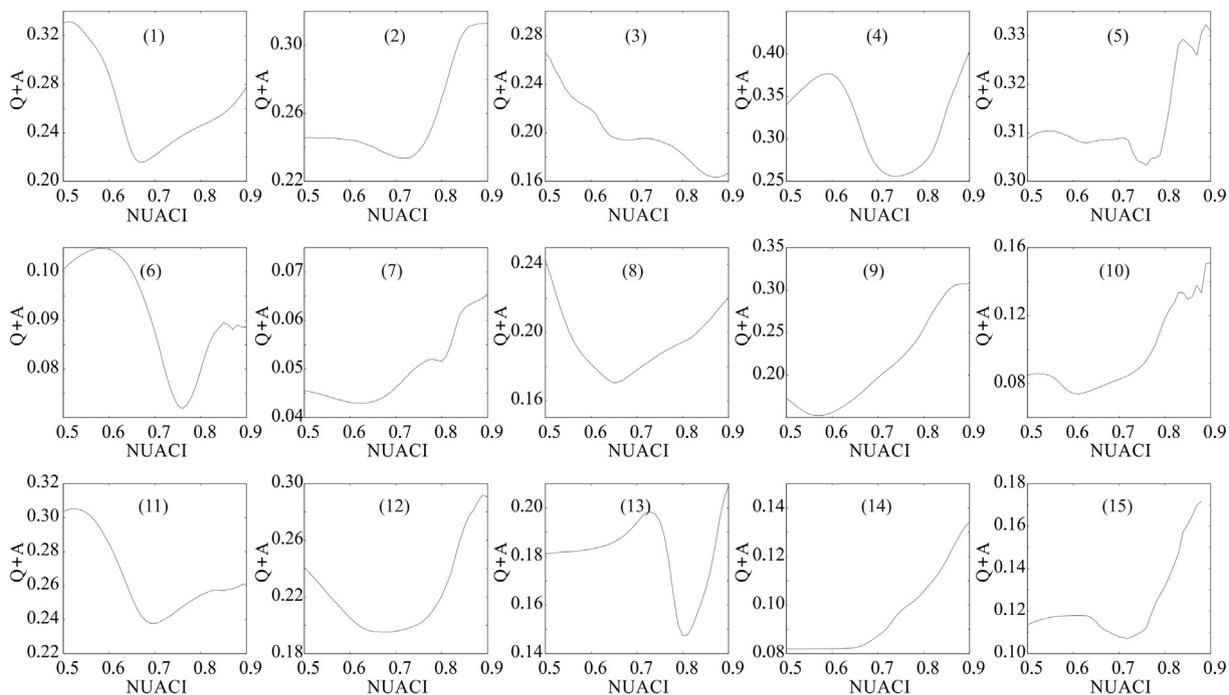


Fig. 6. The relationships between the overall error ($Q + A$) and the varying NUCI thresholds (based on the 2000 data).

whereas #11 and #13 have relatively small Kappa values of approximately or < 0.20 . The other urban ecoregions have Kappa values ranging from 0.30 to 0.50. These results indicate that our urban land classifications in most of the urban ecoregions have fair (0.21–0.40) and moderate agreement (0.41–0.60) with the reference data according to the guidelines suggested by [Viera and Garrett \(2005\)](#).

We also compared our results in the year 2000 with GlobeLand30 (2000), MODIS500m (circa 2001–2002), GLC2000 (2000), GRUMP (1995) and IMPSA (2000–2001). [Fig. 8](#) shows the comparisons in six representative cities. These cities range from mega cities (e.g., New York and Guangzhou), middle-size cities (e.g., Guadalajara) to small

cities (e.g., Haifeng). When visually compared with the referenced Landsat image, our mapping results and GlobeLand30 demonstrate the accurate representation of urban land with the finest spatial details. MODIS500 m also provides a good delineation of urban land, albeit less detailed due to the coarse spatial resolution of 500 m. The urban land products of GLC2000, GRUMP and IMPSA exhibit unstable quality in the selected cities. GLC2000 tends to overestimate the distribution of urban land in Lublin, where the urban land mixed with the other non-urban land types is depicted as a continuous urbanized area. For cities such as São Paulo, GLC2000 dramatically underestimates the extent of urban land. Moreover, GLC2000 have more serious omission errors for

Table 2
The NUACI thresholds and their classification errors in the year 2000.

Urban Ecoregion	1	2	3	4	5	6	7	8
Optimal threshold	0.67	0.72	0.87	0.74	0.76	0.76	0.63	0.65
Q + A (calibrating)	0.22	0.23	0.16	0.26	0.31	0.07	0.04	0.17
Q + A (testing)	0.28	0.16	0.16	0.25	0.29	0.16	0.05	0.16
	9	10	11	12	13	14	15	
Optimal threshold	0.57	0.61	0.70	0.68	0.80	0.53	0.72	
Q + A (calibrating)	0.15	0.07	0.24	0.20	0.15	0.08	0.11	
Q + A (testing)	0.17	0.09	0.09	0.14	0.21	0.15	0.11	

the cities of Adana, Guangzhou and Haifeng. GRUMP and IMPSA show similar drawbacks, i.e., over- or under-estimation of urban land. GRUMP completely misrepresents the urban areas in most of the selected cities except for New York and São Paulo, while IMPSA mainly omits the urban land at the outskirts of the selected cities.

Cohen's Kappa was also computed for the selected global urban land products based on the 300 6-km reference data blocks for the year 2000. The full results can be found in Table A2 in the appendix. Our results and GlobeLand30 have the greatest Kappa at both regional and global levels. The Kappa values for our results are relatively larger than those for GlobeLand30 in six urban ecoregions (#1, #2, #4, #7, #9 and #12) but are smaller in the urban ecoregions of #3, #8, #10, #11, #13, #14 and #15. At the global level, however, the average Kappa value for our results (0.48) is larger than that of GlobeLand30 (0.4309). Compared with the MODIS500m data (0.43), our results have larger Kappa values in most urban ecoregions except for #3, #5, #10, #11, #13 and #14. The comparison of Kappa values also reveals that our approach outperforms the other urban land products of GLC2000 (0.30), GRUMP (0.21) and IMPSA (0.38).

Next, we compared the areas of urban land estimated by the selected global urban land products (x-axis), including ours, against those obtained from the 300 reference data blocks (y-axis). It is assumed that more accurate mapping results are those with smaller root mean square errors (RMSE) and larger R^2 values. Fig. 9 shows that our mapping

results fit tightly to the 1:1 line ($R^2 = 0.88$; slope = 0.99). The slope values for almost all the urban land products are < 1.0 , indicating an overestimation of the urban land area. In particular, the GRUMP data are the least accurate due to the most serious overestimation, reflected by the small slope value of 0.34 ($R^2 = 0.35$).

Furthermore, we evaluated the quality of our urban land classifications by comparing them with two regional-scale, multi-temporal land cover datasets covering the countries of China and the U.S. We used these two datasets as the reference data for 1990, 1995 (only available in China's data), 2000, 2005 and 2010. Next, we randomly selected 25 cities from China and the U.S., respectively, as the evaluation sites (Fig. 10).

In China's 25 selected cities, our classifications yield mean Kappa values of 0.23 (1990), 0.32 (1995), 0.32 (2000), 0.36 (2005) and 0.42 (2010), while in the U.S., mean Kappa values in the selected cities are 0.36 (1990), 0.41 (2000), 0.43 (2005) and 0.46 (2010). For China, our classification produces large Kappa values for cities such as Beijing, Shenzhen, Qiqihaer and Longyan. However, the Kappa values vary over time for Huzhou, Zhoukou and Dali. For the U.S. cities, large Kappa values (> 0.60) were noted for Cleveland, Oxnard and Sarasota, whereas the Kappa values decline in cities such as Las Vegas and Des Moines. Overall, the Kappa values for our urban land classifications are between 0.43 and 0.50 at the global level and ~ 0.33 (China: [0.23, 0.42]; 1990–2010) and ~ 0.42 (U.S: [0.36, 0.46]; 1990–2010) at the country level.

The presented mapping results also have several limitations. There are missing data in the mapping results for the 1990s due to either the lack of Landsat images in part of the world (e.g., Russia) or the shortage of high-resolution images for obtaining reference data. These problems impair the completeness and quality of the presented data product. The binary mask also needs improvement, because the nighttime lights data may not effectively detect artificial infrastructure such as interstate highways and paved settlements that are usually not lit up in the night. Moreover, the accuracies of our classifications are relatively small in tropical areas and arid areas. The binary urban and non-urban classification is also a simplification of reality and cannot effectively address the mixed pixels problem of Landsat images in suburban regions. The

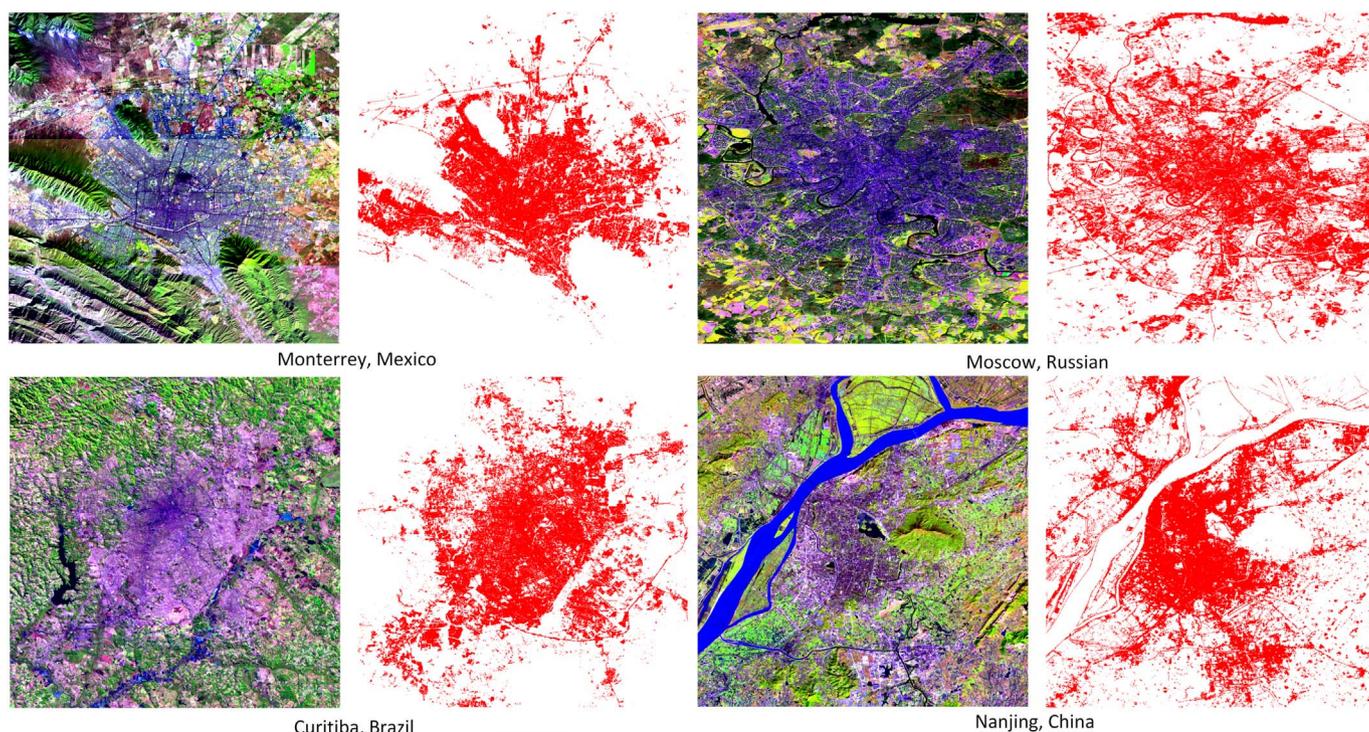


Fig. 7. Urban land classifications for the representative cities in the year 2000.

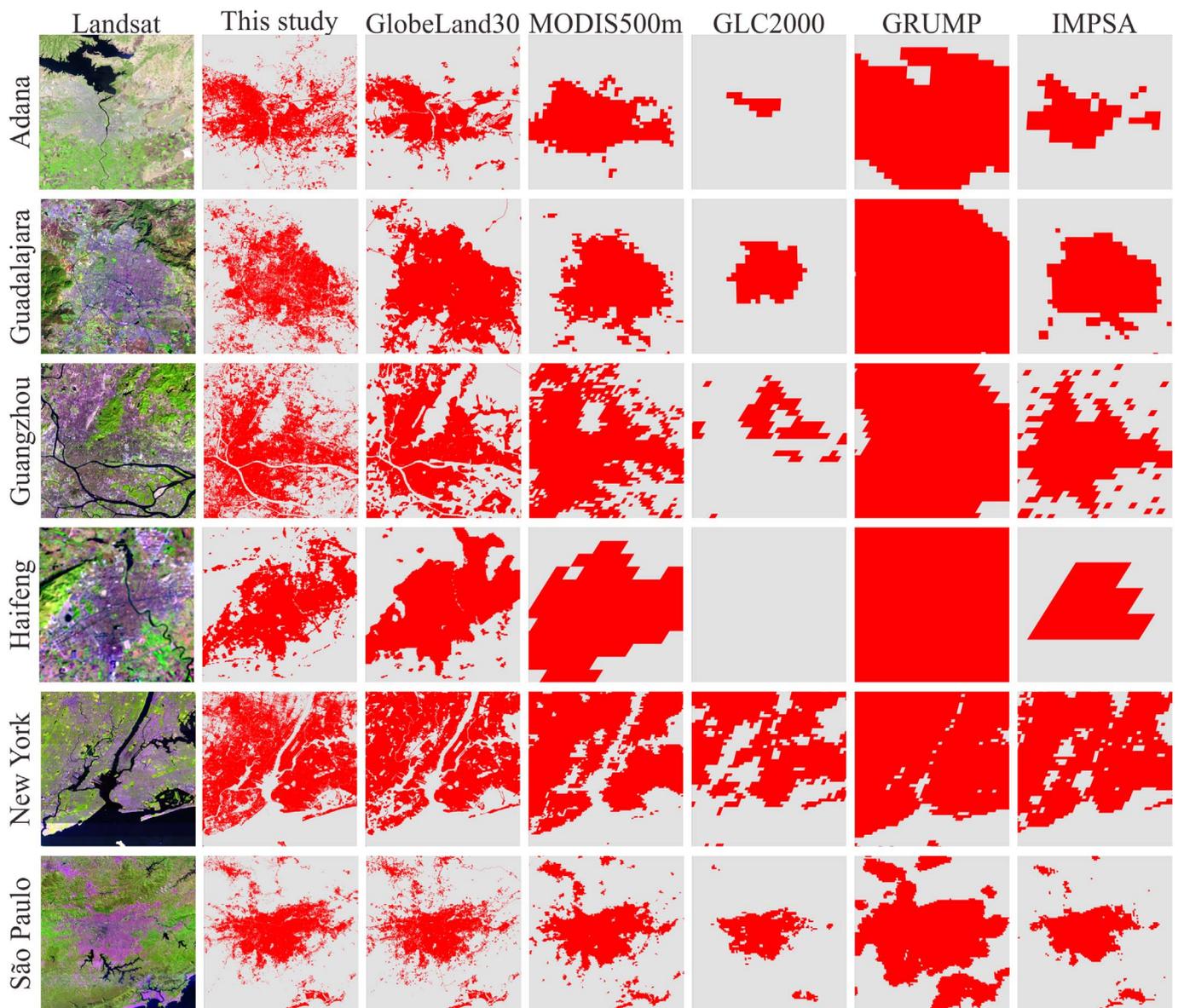


Fig. 8. Comparing the urban land classifications obtained by NAUCI-based classifications and other contemporary products.

development of sub-pixel classification methods is a feasible approach to address this problem.

On the other hand, the reference dataset produced by manual interpretation is not error-free (Foody et al., 2016). The imperfections in the sample size, mislabeling objects and mixed pixels may induce uncertainties in the classification accuracies assessment and area estimation (Foody, 2009, 2010, 2013). Nevertheless, manual interpretation is a feasible method to produce the reference data regarding the budget's condition and precision considerations. We have also released our reference dataset and users can access it through the link: <http://www.geosimulation.cn/GlobalUrbanLand.html>.

3.4. Assessment of global urban land expansion from 1990 to 2010

The presented new multi-temporal global urban land data facilitate the characterization of global urban land expansion from 1990 to 2010. Fig. 11 shows the extent and expansion of urban land in several representative cities around the world. Table 3 shows estimated global urban land areas with their 95% confidence intervals from 1990 to 2010. By analyzing these data, we found that the world's urban land area increased to 747.05 ± 1.50 thousand km² during the past two

decades. The global coverage of urban land changed from 0.46% in 1990 to 0.63% in 2010 (excluding Antarctica and Greenland). These results are close to those reported in some previous studies, such as Gong et al. (2013) (0.66%), Zhou et al. (2015) (0.50%) and Liu et al. (2014) (0.45%). Table 4 summarizes the world's 10 countries with the most rapid urban land expansion. In total, these 10 countries contributed 57.61% of the world's total increase of urban land area from 1990 to 2010. In particular, China (21.49%), the United States (14.21%) and India (7.30%) together (14% of the world's terrestrial area in total) already contribute almost 43% of the world's total expansion. Overall, these are the results for a preliminary analysis of global urban land expansion using our multi-temporal global urban land data. These data can also be applied to other urban research, such as the projection of future urban growth at regional and global scales (Chen et al., 2014; Gao and Bryan, 2017; Li et al., 2017; Liu et al., 2017).

4. Conclusions

In this paper, we presented the production of a new dataset of multi-temporal 30-m global urban land maps. The estimated global level

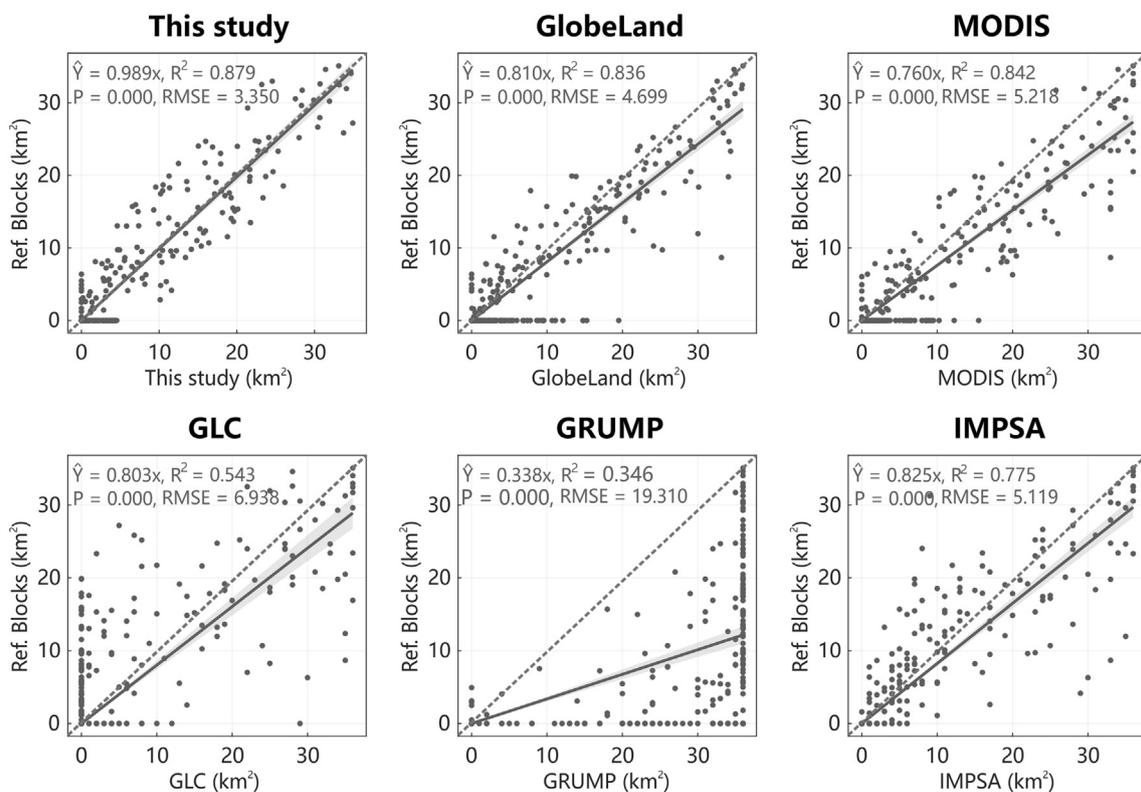


Fig. 9. Scatter graphs of the 300 reference data blocks, in which each graph shows the relationship between the estimated urban land areas in each data product and those obtained from the reference data blocks. The broken line is the 1:1 line, while the solid line represents the fit line.

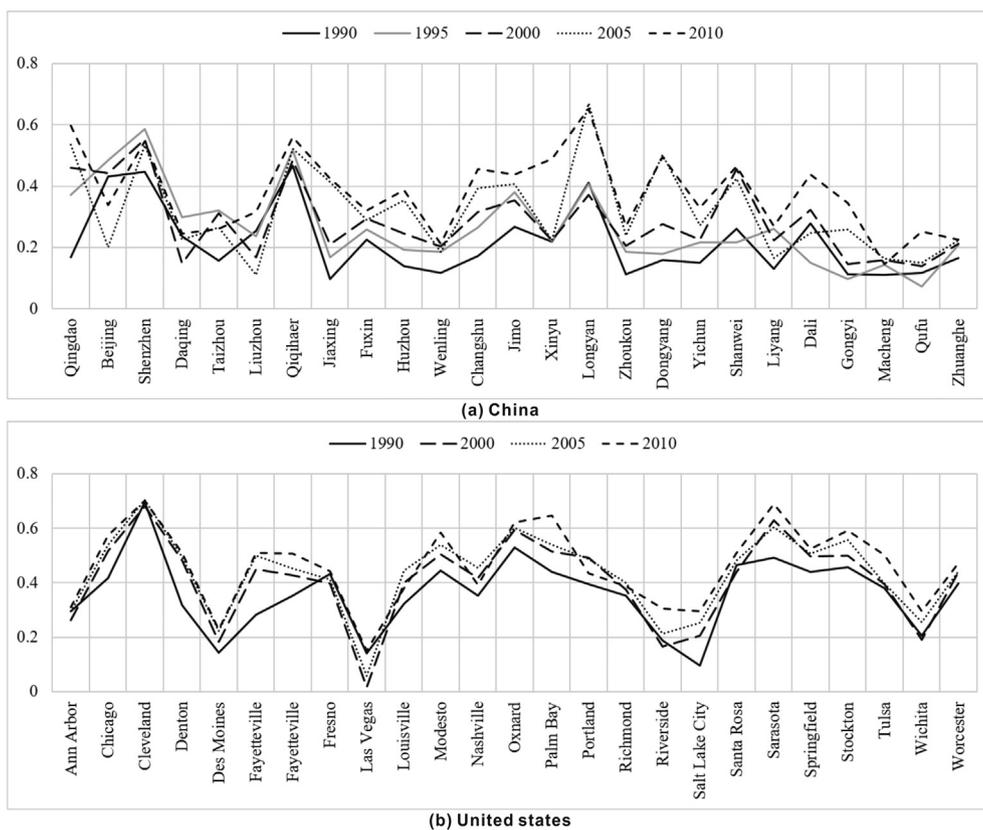


Fig. 10. Kappa values over time using the data from China's Land-use Database and the National Land Cover Database of the U.S.

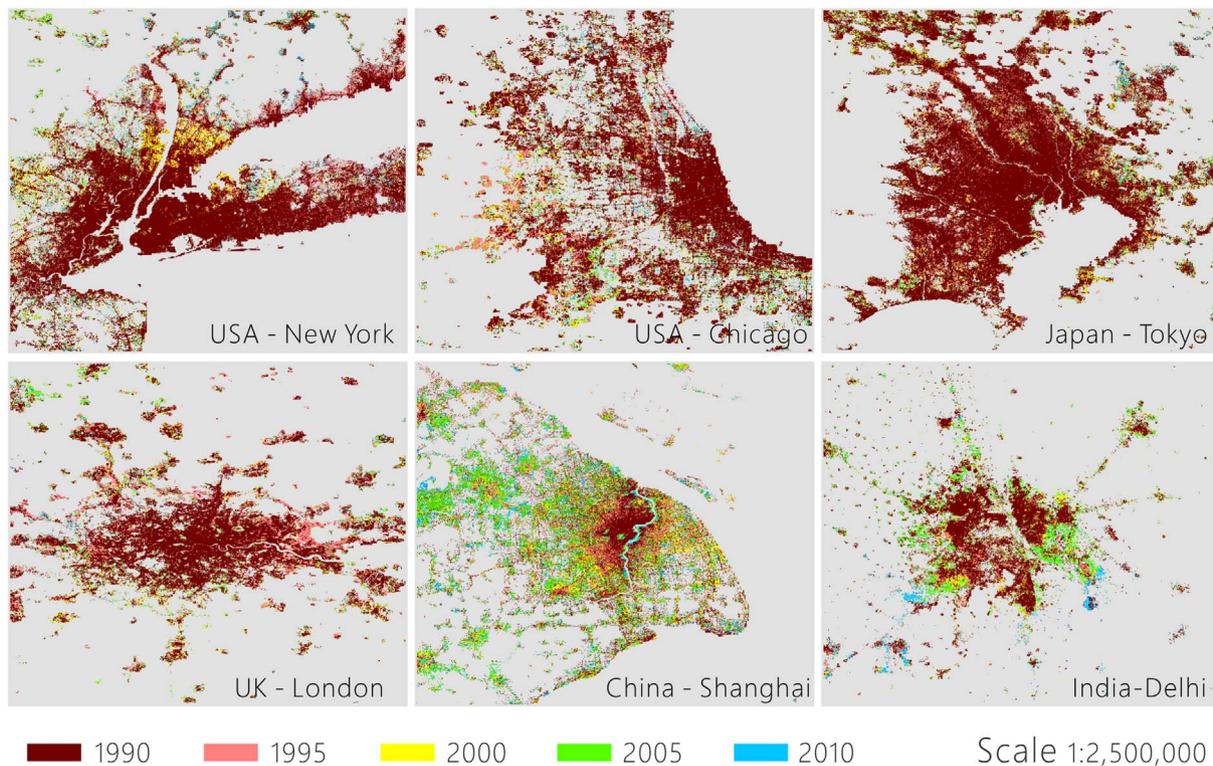


Fig. 11. Urban land expansion from 1990 to 2010 in the representative cities.

Table 3
World's estimated urban land areas from 1990 to 2010 with their 95% confidence intervals (thousand km²).

	1990	1995	2000	2005	2010
Africa	17.52 (± 0.21)	18.57 (± 0.20)	28.53 (± 0.25)	43.65 (± 0.39)	41.20 (± 0.30)
Asia-Oceania	196.50 (± 1.14)	244.77 (± 1.24)	310.40 (± 1.43)	310.58 (± 1.35)	393.59 (± 1.58)
Europe	100.36 (± 0.58)	100.83 (± 0.58)	126.45 (± 0.61)	169.39 (± 0.89)	158.12 (± 0.70)
N. America	124.67 (± 0.49)	142.88 (± 0.56)	149.84 (± 0.55)	180.76 (± 0.72)	171.03 (± 0.59)
S. America	30.09 (± 0.20)	33.64 (± 0.21)	36.56 (± 0.21)	40.94 (± 0.22)	44.93 (± 0.21)
Total	450.97 (± 1.18)	504.11 (± 1.26)	596.12 (± 1.34)	655.70 (± 1.41)	747.05 (± 1.50)

Table 4
The 10 countries with the greatest rates of urban land expansion from 1990 to 2010 (thousand km²).

	1990	2010	% of World's total change 1990–2010
China	51.45 (± 0.71)	115.09 (± 1.03)	21.49%
United States	97.72 (± 0.43)	139.78 (± 0.54)	14.21%
India	20.60 (± 0.53)	42.23 (± 0.66)	7.30%
Brazil	13.89 (± 0.15)	21.89 (± 0.17)	2.70%
France	6.40 (± 0.17)	13.88 (± 0.22)	2.53%
Germany	11.74 (± 0.16)	18.91 (± 0.21)	2.42%
Japan	16.83 (± 0.34)	24.00 (± 0.33)	2.42%
Russia	28.56 (± 0.21)	34.49 (± 0.22)	2.00%
United Kingdom	10.49 (± 0.12)	15.13 (± 0.14)	1.57%
South Africa	2.95 (± 0.06)	5.80 (± 0.09)	0.96%

overall accuracy, producer's accuracy and user's accuracy for this dataset are 0.81–0.84, 0.50–0.60 and 0.49–0.61, respectively. The Kappa values are 0.43–0.50 at the global level, and ~0.33 (in China) and

~0.42 (in the U.S.) at the country level. By analyzing this dataset, we found that the world's total urban land area increased from 450.97 ± 1.18 thousand km² in 1990 to 747.05 ± 1.50 thousand km² in 2010. The world's three countries with the greatest rates for urban land expansion, namely, the United States, China and India (14% of the world's terrestrial area in total), contributed almost 43% of the total urban land area increase. In 2010, the global coverage of urban land has reached approximately 0.63%. The presented global urban land maps can be accessed through the webpage: <http://www.geosimulation.cn/GlobalUrbanLand.html>. Researchers from all around the world are encouraged to use and evaluate this new 30-m resolution dataset. Our future efforts will be devoted to the continuous update and refinement of this dataset.

Acknowledgement

We sincerely thank Prof. Ronald McRoberts and the anonymous reviewers for their useful comments and suggestions. The authors of Xiaoping Liu and Guohua Hu contribute equally to this paper. This research was supported by the National Key R&D Program of China (Grant No. 2017YFA0604401, 2017YFA0604402 and 2017YFA0604404), the National Natural Science Foundation of China (Grant No. 41601420), the Key National Natural Science Foundation of China (Grant No. 41531176), the Guangdong Natural Science Foundation (Grant No. 2015A030310288), and the Fundamental Research Funds for the Central Universities (16lgpy03, 15lgjc38).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.rse.2018.02.055>.

References

Bagan, H., Yamagata, Y., 2012. Landsat analysis of urban growth: how Tokyo became the world's largest megacity during the last 40 years. *Remote Sens. Environ.* 127,

- 210–222.
- Bartholomé, E., Belward, A., 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. *Int. J. Remote Sens.* 26, 1959–1977.
- Chen, J.M., Cihlar, J., 1996. Retrieving leaf area index of boreal conifer forests using Landsat TM images. *Remote Sens. Environ.* 55, 153–162.
- Chen, Y., Li, X., Liu, X., Ai, B., 2014. Modeling urban land-use dynamics in a fast developing city using the modified logistic cellular automaton with a patch-based simulation strategy. *Int. J. Geogr. Inf. Sci.* 28, 234–255.
- Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., 2015. Global land cover mapping at 30 m resolution: a POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* 103, 7–27.
- Congalton, R.G., Green, K., 2008. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. CRC Press.
- Elvidge, C.D., Tuttle, B.T., Sutton, P.C., Baugh, K.E., Howard, A.T., Milesi, C., Bhaduri, B., Nemani, R., 2007. Global distribution and density of constructed impervious surfaces. *Sensors* 7, 1962–1979.
- Foody, G.M., 2009. Sample size determination for image classification accuracy assessment and comparison. *Int. J. Remote Sens.* 30, 5273–5291.
- Foody, G.M., 2010. Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sens. Environ.* 114, 2271–2285.
- Foody, G.M., 2013. Ground reference data error and the mis-estimation of the area of land cover change as a function of its abundance. *Remote Sens. Lett.* 4, 783–792.
- Foody, G.M., Pal, M., Rocchini, D., Garzon-Lopez, C.X., Bastin, L., 2016. The sensitivity of mapping methods to reference data quality: training supervised image classifications with imperfect reference data. *ISPRS Int. J. Geo-Inf.* 5, 199.
- Gao, L., Bryan, B.A., 2017. Finding pathways to national-scale land-sector sustainability. *Nature* 544, 217–222.
- Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., 2013. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. *Int. J. Remote Sens.* 34, 2607–2654.
- Li, X., Gong, P., Liang, L., 2015. A 30-year (1984–2013) record of annual urban dynamics of Beijing City derived from Landsat data. *Remote Sens. Environ.* 166, 78–90.
- Li, X., Chen, G., Liu, X., Liang, X., Wang, S., Chen, Y., Pei, F., Xu, X., 2017. A new global land-use and land-cover change product at a 1-km resolution for 2010 to 2100 based on human–environment interactions. *Ann. Am. Assoc. Geogr.* 107, 1040–1059.
- Liu, Z., He, C., Zhou, Y., Wu, J., 2014. How much of the world's land has been urbanized, really? A hierarchical framework for avoiding confusion. *Landsc. Ecol.* 29, 763–771.
- Liu, X., Hu, G., Ai, B., Li, X., Shi, Q., 2015. A Normalized Urban Areas Composite Index (NUACI) based on combination of DMSP-OLS and MODIS for mapping impervious surface area. *Remote Sens.* 7, 17168–17189.
- Liu, X., Liang, X., Li, X., Xu, X., Ou, J., Chen, Y., Li, S., Wang, S., Pei, F., 2017. A future land use simulation model (FLUS) for simulating multiple land use scenarios by coupling human and natural effects. *Landsc. Urban Plan.* 168, 94–116.
- 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.
- Olofsson, P., Stehman, S.V., Woodcock, C.E., Sulla-Menashe, D., Sibley, A.M., Newell, J.D., Friedl, M.A., Herold, M., 2012. A global land-cover validation data set, part I: Fundamental design principles. *Int. J. Remote Sens.* 33, 5768–5788.
- Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* 148, 42–57.
- Patel, N.N., Angiuli, E., Gamba, P., Gaughan, A., Lisini, G., Stevens, F.R., Tatem, A.J., Trianni, G., 2015. Multitemporal settlement and population mapping from Landsat using Google Earth Engine. *Int. J. Appl. Earth Obs. Geoinf.* 35, 199–208.
- Pengra, B., Long, J., Dahal, D., Stehman, S.V., Loveland, T.R., 2015. A global reference database from very high resolution commercial satellite data and methodology for application to Landsat derived 30m continuous field tree cover data. *Remote Sens. Environ.* 165, 234–248.
- Pontius, R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *Int. J. Remote Sens.* 32, 4407–4429.
- Potere, D., Schneider, A., Angel, S., Civco, D.L., 2009. Mapping urban areas on a global scale: which of the eight maps now available is more accurate? *Int. J. Remote Sens.* 30, 6531–6558.
- Schneider, A., Friedl, M.A., Potere, D., 2010. Mapping global urban areas using MODIS 500-m data: new methods and datasets based on 'urban ecoregions'. *Remote Sens. Environ.* 114, 1733–1746.
- Sexton, J.O., Song, X.-P., Huang, C., Channan, S., Baker, M.E., Townshend, J.R., 2013. Urban growth of the Washington, DC–Baltimore, MD metropolitan region from 1984 to 2010 by annual, Landsat-based estimates of impervious cover. *Remote Sens. Environ.* 129, 42–53.
- Shi, K., Yu, B., Huang, Y., Hu, Y., Yin, B., Chen, Z., Chen, L., Wu, J., 2014a. Evaluating the ability of NPP-VIIRS nighttime light data to estimate the gross domestic product and the electric power consumption of China at multiple scales: a comparison with DMSP-OLS data. *Remote Sens.* 6, 1705–1724.
- Shi, K., Huang, C., Yu, B., Yin, B., Huang, Y., Wu, J., 2014b. Evaluation of NPP-VIIRS night-time light composite data for extracting built-up urban areas. *Remote Sens. Lett.* 5, 358–366.
- 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.
- Stehman, S.V., Olofsson, P., Woodcock, C.E., Herold, M., Friedl, M.A., 2012. A global land-cover validation data set, II: augmenting a stratified sampling design to estimate accuracy by region and land-cover class. *Int. J. Remote Sens.* 33, 6975–6993.
- Viera, A.J., Garrett, J.M., 2005. Understanding interobserver agreement: the kappa statistic. *Fam. Med.* 37, 360–363.
- Yu, B., Shu, S., Liu, H., Song, W., Wu, J., Wang, L., Chen, Z., 2014. Object-based spatial cluster analysis of urban landscape pattern using nighttime light satellite images: a case study of China. *Int. J. Geogr. Inf. Sci.* 28, 2328–2355.
- Zha, Y., Gao, J., Ni, S., 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens.* 24, 583–594.
- Zhang, L., Weng, Q., 2016. Annual dynamics of impervious surface in the Pearl River Delta, China, from 1988 to 2013, using time series Landsat imagery. *ISPRS J. Photogramm. Remote Sens.* 113, 86–96.
- Zhou, Y., Smith, S.J., Zhao, K., Imhoff, M., Thomson, A., Bond-Lamberty, B., Asrar, G.R., Zhang, X., He, C., Elvidge, C.D., 2015. A global map of urban extent from nightlights. *Environ. Res. Lett.* 10, 054011.