Remote Sensing of Environment 205 (2018) 253–275

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Using Landsat and nighttime lights for supervised pixel-based image classification of urban land cover

Ran Goldblatt[a,*], Michelle F. Stuhlmacher[b], Beth Tellman[b], Nicholas Clinton[c], Gordon Hanson[a], Matei Georgescu[b], Chuyuan Wang[b], Fidel Serrano-Candela[d], Amit K. Khandelwal[e], Wan-Hwa Cheng[b], Robert C. Balling Jr[b]

[a] School of Global Policy and Strategy, UC San Diego, 9500 Gilman Drive La Jolla, CA 92093-0519, USA
[b] School of Geographical Sciences and Urban Planning, Arizona State University, 976 S Forest Mall, Tempe, AZ 85281, USA
[c] Google, Inc., 1600 Amphitheatre Pkwy, Mtn. View, CA 94043, USA
[d] Laboratorio Nacional de Ciencias de la Sostenibilidad, Universidad Nacional Autónoma de México, Apartado Postal 70-275 Ciudad Universitaria, UNAM 04510 México, D.F. Circuito exterior s/n anexo al Jardín Botánico Exterior, Mexico
[e] Columbia Business School, Columbia University, New York, NY 10027, USA

ARTICLE INFO

Keywords:
Urbanization
Built-up land cover
Nighttime light
Image classification
Google Earth Engine

ABSTRACT

Reliable representations of global urban extent remain limited, hindering scientific progress across a range of disciplines that study functionality of sustainable cities. We present an efficient and low-cost machine-learning approach for pixel-based image classification of built-up areas at a large geographic scale using Landsat data. Our methodology combines nighttime-lights data and Landsat 8 and overcomes the lack of extensive ground-reference data. We demonstrate the effectiveness of our methodology, which is implemented in Google Earth Engine, through the development of accurate 30 m resolution maps that characterize built-up land cover in three geographically diverse countries: India, Mexico, and the US. Our approach highlights the usefulness of data fusion techniques for studying the built environment and is a first step towards the creation of an accurate global-scale map of urban land cover over time.

1. Introduction

Urbanization has been a fundamental trend of the past two centuries and a key force shaping the development of the modern world. Between 1950 and 2014, the share of the global population living in urban areas increased from 30% to 54%, and in the next few decades is projected to expand by an additional 2.5 billion urban dwellers, primarily in Asia and Africa (Seto et al., 2011; UN, 2014). Urban population growth is accompanied by a dramatic increase in the land area incorporated in cities (Seto et al., 2011). While urbanization in rapidly growing nations is helping lift hundreds of millions of people out of poverty, it is also creating immense societal challenges by increasing greenhouse-gas emissions, destabilizing fragile ecosystems, and creating new demands on public services and infrastructure that impose significant burdens on the environment (Ban et al., 2015). Timely and reliable information on the extent of urban areas is fundamental for the support of sustainable urban development and management (Ban et al., 2015; Jacob and Ban, 2015). Despite the importance of understanding the drivers of urban growth, we are still unable to quantify the magnitude and pace of urbanization in a consistent manner at high resolution and global scale (Ban et al., 2015; Giri et al., 2013).

The revolution in geospatial data has transformed how we study cities. Previous approaches leveraged household surveys but these are expensive to collect, produced infrequently, and subject to measurement problems. Since the 1970s, however, terrestrial Earth-observation data have been continuously collected in various spectral, spatial and temporal resolutions. As improved satellite imagery becomes available, new remote-sensing methods and machine-learning approaches have been developed to convert terrestrial Earth-observation data into meaningful information about the nature and pace of change of urban landscapes and human settlements (Ban et al., 2015; Chen et al., 2015; CIESIN, 2005; Gaughan et al., 2013; Pesaresi et al., 2016; Potere et al., 2009; Seto et al., 2011; Taubenböck et al., 2012).

The availability of satellite data has triggered the development of new methods to map global land cover using remotely-sensed data such as Landsat (Chen et al., 2015; Gaughan et al., 2013; Goldblatt et al., 2016; Patel et al., 2015), MODIS (Moderate Resolution Imaging Spectroradiometer) (Schneider et al., 2009, 2010; Wan et al., 2015), DMSP-

* Corresponding author.
E-mail addresses: rgoldblatt@ucsd.edu, ranrang@gmail.com (R. Goldblatt).

https://doi.org/10.1016/j.rse.2017.11.026
Received 20 February 2017; Received in revised form 12 November 2017; Accepted 29 November 2017

OLS (Elvidge et al., 2014; Liu et al., 2012; Xiao et al., 2014; Zhang and Seto, 2013) and other spaceborne High-Resolution (HR), Very-High-Resolution (VHR) and Synthetic-aperture radar (SAR) radar sensors (Ban et al., 2015; Gamba et al., 2011; Jacob and Ban, 2015). Recent studies have developed automated and semi-automated classification procedures to map global land cover at a 30 m resolution with high accuracy (Ban et al., 2015; Chen et al., 2015). Because Landsat satellites have been collecting data from Earth since 1972, Landsat data are often used for analysis of urban change (Patel et al., 2015), and are ideal for land cover mapping (Woodcock et al., 2008). Nighttime light data are also associated with developed land (Elvidge et al., 2014; Levin and Duke, 2012; Sutton, 2003) and can be used to infer the extent of urban areas (Bagan and Yamagata, 2015; Small and Elvidge, 2013; Zhang and Seto, 2013), as well as economic activity at the local, regional and national levels (Elvidge et al., 2014; Henderson et al., 2003; Keola et al., 2015). Sensors on board the Operational Line-scan System of the Defense Meteorological Satellite Program (DMSP-OLS) have captured artificial lighting since the early 1990's. A pixel's nighttime light value that exceeds a specified threshold, which may vary across regions or countries, signifies urban development (Henderson et al., 2003; Liu et al., 2016; Small and Elvidge, 2013; Su et al., 2015; Wei et al., 2014; Zhou et al., 2014, 2015). However, inference using nighttime-light data are often inaccurate, particularly in low-density urban areas (Zhang and Seto, 2013). DMSP-OLS can also exaggerate the extent of urban areas (Henderson et al., 2003; Small et al., 2005), while overlooking small or developing settlements. In addition, the extent and intensity of lit areas cannot directly delimit urban regions due to the “blooming” effect (Imhoff et al., 1997) and “saturation” of pixels (Hsu et al., 2015). Blooming refers to the identification of lit areas as consistently larger than the settlements with which they are associated (Small et al., 2005); saturation occurs when pixels in bright areas, such as in city centers, reach the highest possible digital number (DN) value (i.e., 63) and no further details can be recognized (Hsu et al., 2015).

Until recently, most remote sensing studies focused on local settings (Herold, 2009). Mapping land cover at a national or regional scale is challenging because of the lack of high-resolution global imagery, the heterogeneous and complex spectral characteristics of land, and the small and fragmented spatial configuration of many cities (Chen et al., 2015; Herold, 2009). In the case of mapping urbanization, existing maps of urban land show considerable disagreement on the location and extent of urbanization (Potere et al., 2009; Seto et al., 2011) and are limited across space and time. These inconsistencies arise in part because the delineation of urban land depends on the input data (Schneider et al., 2010), which may capture different dimensions of urbanization, such as built-up land cover or land use and population density (Bagan and Yamagata, 2014; Stevens et al., 2015; Tatem et al., 2007).

1.1. Detecting urbanization using machine learning

Urban areas can be detected in satellite imagery using various machine-learning approaches (e.g., supervised, unsupervised and semi-supervised). These approaches typically rely on reference data that mark urban features, either for training or validation. Reference data are fundamental not only for mapping urbanization across space, but also for classification over time (Boucher and Seto, 2009). Some of the reference datasets used for classification include Landsat-based urban maps (Potere et al., 2009), census-based population databases (Stevens et al., 2015), hand-labeled examples (Goldblatt et al., 2016), and data collected via crowd-source platforms, such as OpenStreetMap (OSM) (Belgiu and Drăguţ, 2014; Estima and Painho, 2015). However, because they are expensive to collect, reference datasets for large geographic scales are scarce (Miyazaki et al., 2011). Due to the scarcity of ground-reference data, it is often necessary to exploit existing global coarse datasets and classification products to create accurate higher-resolution maps of urban areas (Kasimu et al., 2009; Trianni et al., 2015). Moreover, mapping land cover at a global scale and with high precision requires effective, efficient and operational approaches to deal with a very large volume of data. For example, it is estimated that over 10,000 Landsat satellite images are required to cover the entire Earth at 30 m resolution (Chen et al., 2015). Until recently, the majority of studies that analyze urbanization have been limited in their geographic scale because of the lack of extensive high-resolution satellite data, scarcity of ground-reference data, and computational constraints. Emerging cloud-based computational platforms now allow for scaling analysis across space and time. Google Earth Engine (GEE) is one platform that leverages cloud-computing services to achieve planetary-scale utility. GEE has been previously used to map population (Patel et al., 2015; Trianni et al., 2015), urban areas (Goldblatt et al., 2016), and surface water (Pekel et al., 2016). This paper contributes to this literature by developing a machine-learning methodology for supervised high-resolution image classification of built-up areas using GEE's cloud-based computational platform.

1.2. Research objective and contribution

The use of nighttime remotely-sensed data to map urbanization is not new to the literature. Remotely-sensed data on artificial lighting has long been considered an economical way to map urbanization and development across the globe (Elvidge et al., 2009). By utilizing the distribution of vegetation land cover, the combination of nighttime and daytime data increases the heterogeneity of urban and suburban land cover (e.g., distinguishing between built-up land cover and vegetation in urban areas) and improves the characterization of inter-urban variability in nighttime luminosity (Zhang et al., 2013). This, in turn, improves the ability to detect urban features (Lu et al., 2008; Ma et al., 2014) including sub-pixel fractional urban land cover (Huang et al., 2016). Several spectral indices that combine nighttime light and vegetation spectral characteristics have been developed, including the Vegetation Adjusted NTL Urban Index (VANUI) (Jing et al., 2015; Zhang et al., 2013), the Normalized Difference Urban Index (NDUI) (Zhang et al., 2015) and the Normalized Difference Spectral Vector (NDSV) (Trianni et al., 2015). These indices increase the separability between urban and non-urban land cover.

We develop a methodology that combines nighttime and daytime remotely-sensed data. We collect training examples automatically using DMSP-OLS data, and use them for classification of built-up areas with daytime Landsat 30 m spatial resolution imagery. Previous studies that combine nighttime and daytime data have either been limited in their spatial application (i.e., the ability to generalize the method and to apply it over regions with heterogeneous land cover) or spatial resolution (i.e., many of the existing approaches, for example those that rely on MODIS or DMSP-OLS, are limited in their spatial resolution). In this study, we adopt a hexagonal tessellation mapping approach to handle large variation across regions (where we refer to each hexagon in the hexagonal grid as a hex-cell). We collect training examples from each hex-cell and classify the hex-cell as an independent unit of analysis.

Our methodology can be applied across heterogeneous land cover and across time and, crucially, does not rely on expensive hand-labeled examples. It requires minimal manual adjustments for training and classification, and does not require adjustments to local parameters. This feature makes the approach scalable across space and time. Importantly, the methodology is time-invariant and can be applied whenever Landsat and DMSP-OLS data are coincident. The robustness of the methodology lies in our approach to sample training examples (i.e., according to the relative intensity of the emitted light at night and the distribution of vegetation land cover) and on the per-hex-cell classification, which allows us to account for regional variations in the land cover. Finally, we assess the accuracy of the methodology using an extensive dataset of 84,564 hand-labeled polygons characterizing built-up (BU) and not built-up (NBU) pixels for each of the three study areas,

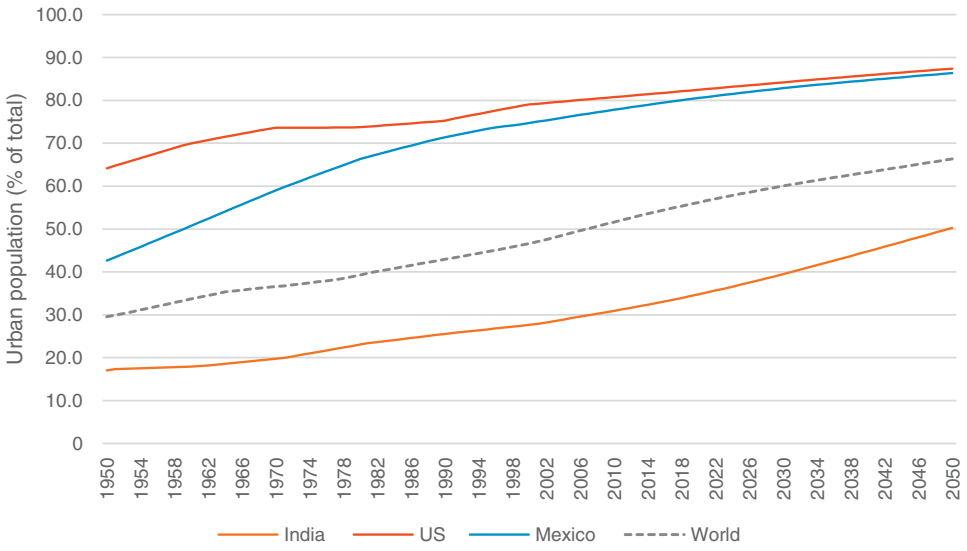

Fig. 1. Annual changes in share of urban population in India, US and Mexico compared to world average (UN, 2014).

which are described below.

1.3. Study areas

To illustrate our methodology and its applicability in heterogeneous and diverse geographic conditions, we map built-up land cover in three countries that are characterized by distinct geographic conditions (e.g., land cover, topography, climate, soil, landform, and fauna): India, Mexico, and the United States (US).

1.3.1. India

The share of India's population living in urban areas in 2015 was 33%, which is much lower than the corresponding values of 79% and 82% for Mexico and the US, respectively. India is, however, urbanizing at a relatively rapid rate (Fig. 1). For example, between 2010 and 2015, India's average annual rate of change of the urban population was 1.14%, compared to 0.36% and 0.21% in Mexico and in the US, respectively. By 2050, half of India's population is likely to be urban.

1.3.2. Mexico

Mexico has gone through three major phases of urbanization. In the first phase, 1900–1940, urban growth slowly incorporated 10% to 20% of the population. In the second phase, 1940–1980, rapid urban expansion, particularly in Mexico City, increased the share of urban population to 55%. Since the 1980s, more dispersed moderate urban growth increased the urban share of the population to over 70% (*Consejo Nacional de Población*, 2012). The current phase of urbanization is characterized by informal urbanization on the city periphery, representing 65% of all new housing construction in Mexico City, and is even higher for small to mid-sized cities (Connolly, 2014). Informal settlements tend to be marginalized in terms of lower socio-economic development, access to services like water and electricity, and are more vulnerable to risks like water scarcity and flooding (Aguilar, 2008; Aguilar and Guerrero, 2013; Consejo Nacional de Población, 2012; Eakin et al., 2016).

1.3.3. US

In the context of the US, the initial urban growth occurred from 1790 to 1890 and the country has become increasingly urban since (US Census Bureau, 2012). In 1910, the Census Bureau defined an urban area as one with a population above 2500, and the 1920 census marked the first time that 50% of the US population lived in an urban area (US Census Bureau, 2016).

2. Conceptual framework: Infused DMSP-OLS/Landsat methodology

Our methodology combines Earth-observation datasets from two domains: DMSP-OLS and Landsat. Using a subset of the data, we identify the optimal parameters to the classifiers and then classify the built-up land cover in the entire area of each country. The classification procedure involves five steps (Fig. 2): (1) Divide each country into a uniform hexagonal grid; (2) Pre-process Landsat 8 images; (3) Extract labeled examples from DMSP-OLS; (4) Perform supervised image classification; and (5) Validate and test. We next describe these steps in detail.

2.1. Divide each country into a uniform hexagonal grid ("mapping zones")

Mapping heterogeneous land cover at a large geographic scale requires partitioning the region of interest into a finite number of relatively homogenous sub-regions that are characterized by similar landform, soil, vegetation, spectral reflectance, and image footprints (Homer et al., 2004). This practice is often referred to as 'zone mapping' (Homer and Gallant, 2001). The partition can be according to different criteria, such as land cover and land use, socio-political definition, or size (Hunsaker et al., 1994; O'Neill et al., 1996; Turner, 1989), or by means of an artificial grid system where each element in the grid is treated as an independent region of interest. Examples of such grids in practice include the Global Grid (GG) system (Theobald, 2016) and the ISEA DGGs of Gong et al. (2013) for global land cover mapping. There are many advantages for using an artificial grid system over administrative division (e.g., census areas or political boundaries) or any other thematic division (e.g., ecological or climatic regions). First, administrative divisions are often delineated according to criteria that are unrelated to physical characteristics (e.g., the boundary between Washington, D.C., and Maryland). Second, the geographic scale and the size of the administrative division varies within and between countries (e.g., the division into census tracts and blocks). Unlike a uniform global artificial grid system, an administrative division does not allow for a multi-geographic-scale classification approach using recursively-divided units of analysis.

Here we partition each of our three countries of interest into an equal-area hexagonal grid (or a hexagonal tessellation). Hex-cells in a hexagonal tessellation are arranged in a contiguous global lattice, which is the most compact arrangement of many equal circles (Birch et al., 2007). Hexagonal grids are advantageous because they are characterized by elements that do not have gaps or overlaps and the

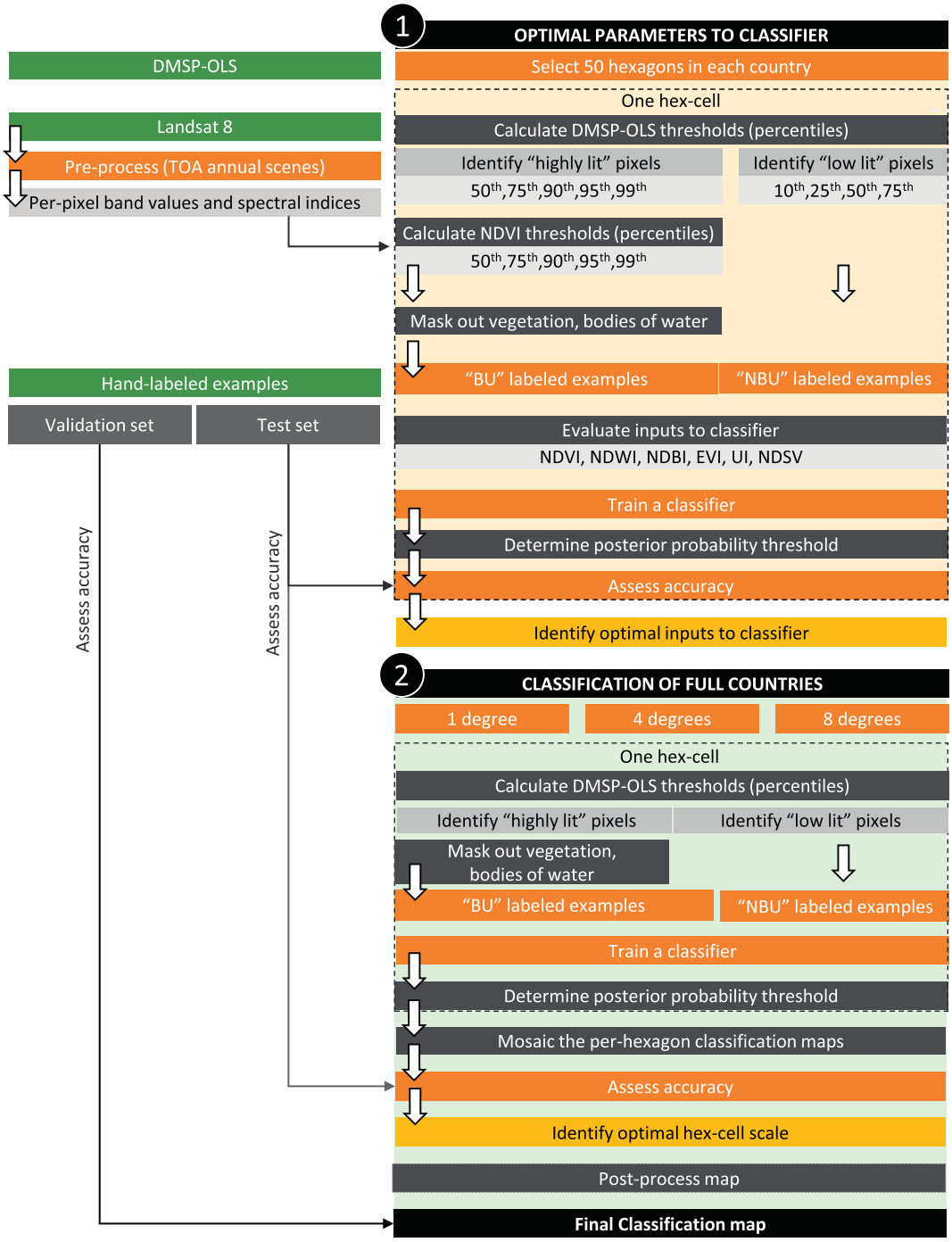

Fig. 2. Schematic illustration of the methodology.

center-to-center distances between adjacent grid cells are approximately equal. Moreover, hex-cells have a topology that is symmetrical, invariant, and of equal area, and can be recursively partitioned into smaller divisions of grids (Richards et al., 2000) (Fig. 3, right). Since the size of each hex-cell in the grid may affect the accuracy of the classification, we use a hexagonal tessellation that in successive classifications treats each pixel as belonging to a hex-cell of a different size. We examine the following sizes of hex-cells (different center-to-center distances): a distance of 1 decimal degree (1-degree), 4 decimal degrees (4-degrees) and 8 decimal degrees (8-degrees) from center to center (see Fig. 3 for illustration).

2.2. Pre-process Landsat 8 images (the classifier's inputs)

We use Landsat 8 data as classifier inputs (predictors). We apply a standard Top-of-Atmosphere (TOA) calibration on all USGS Landsat 8 Raw Scenes in one year. Since processed DMSP-OLS is only available up through 2013, our analysis begins that year. We assign a "cloud score" to each pixel and select the lowest possible range of cloud scores. The cloud score is a simple cloud-likelihood score, intended to compare multiple looks at the same point for 'relative' cloud likelihood (across time). The score, ranging between 0 and 100, is determined according to the brightness of a pixel, its temperature, and a computed per-pixel Normalized Difference Snow Index (NDSI). Similar to Trianni et al. (2015), we retrieve only those scenes with less than 10% cloud

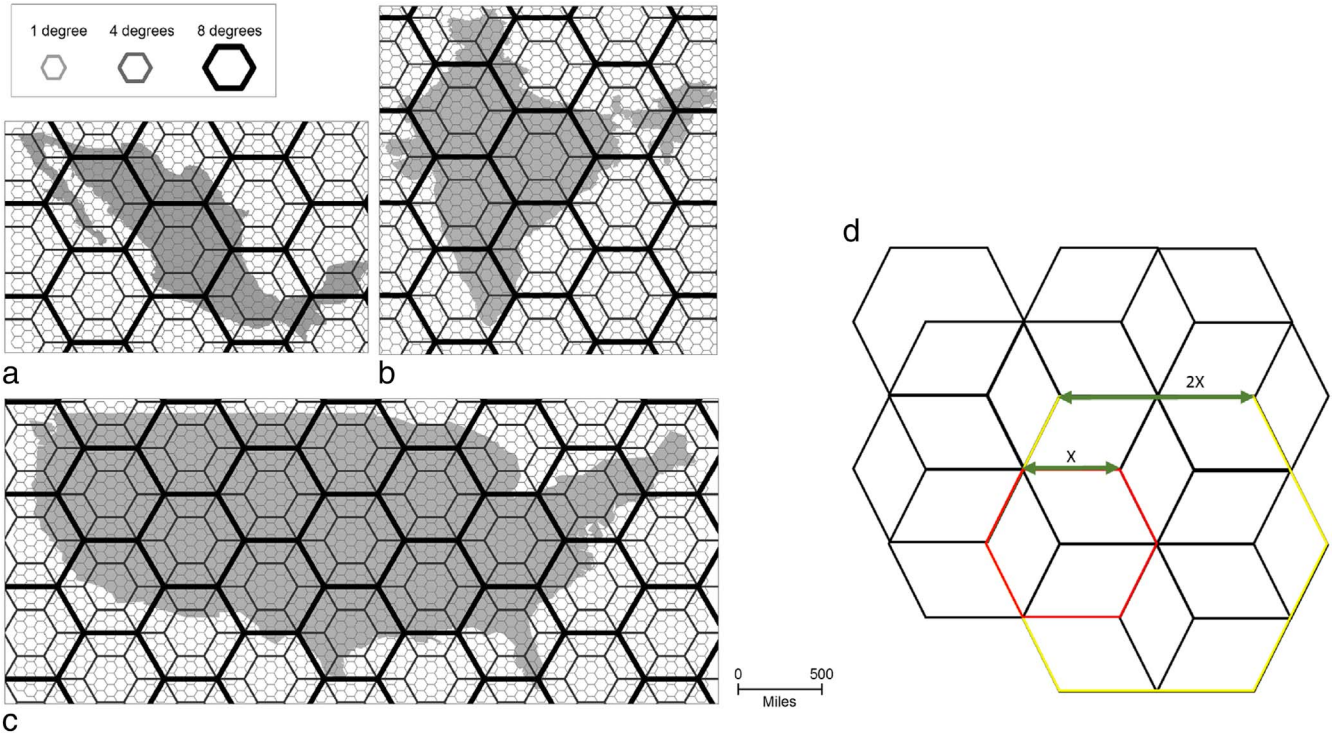

Fig. 3. The three examined hexagon levels: 1, 4, and 8° from center to center, for (a) Mexico, (b) India, and (c) the US (left); The cells in the grid can be recursively partitioned into smaller divisions creating a hierarchical grid; (d) each pixel is classified multiple times, each time as part of a different hexagonal scale.

coverage.

We compute per-band median values from the accepted pixels. Per-pixel median values are insensitive to extreme values, such as very bright or dark pixels, in a stack of scenes (Flood, 2013). The decimal value of each pixel is scaled to an 8-bit range (0–255) for display purposes. We also calculate six additional per-pixel spectral indices and add them as additional inputs to the classifier. These indices are commonly used to identify bodies of water (Normalized Difference Water Index, NDWI), built-up areas (Normalized Difference Built-up Index, NDBI; Urban Index, UI, NDSV), and vegetation (Normalized Difference Vegetation Index, NDVI; Enhanced Vegetation Index, EVI). NDWI (McFeeters, 1996) is commonly used to identify open water or to measure vegetation water content. This index and subsequent versions of it (e.g., the Modification of Normalized Difference Water Index (NDWI)) (Xu, 2006) have been shown to correlate negatively with urban areas and impervious surface (Chen et al., 2006; Xu, 2008). NDVI (Tucker, 1979) and its variant, EVI (Huete et al., 2002), produce spectral indices that tend to negatively correlate with urban areas. Previous studies have used these indices to identify urban areas and to measure urban characteristics such as heat islands, changes in net primary productivity, and urban form (Chen et al., 2006; Clinton and Gong, 2013; Estoque and Murayama, 2015; Li et al., 2013; Myint et al., 2011; Xu, 2008; Yuan and Bauer, 2007; Zhao et al., 2016). UI (Kawamura et al., 1996) and NDBI (Zha et al., 2003) are both used to identify built up areas and correlate positively with urban land cover and impervious surface (Chen et al., 2006; Estoque and Murayama, 2015; Li et al., 2013). NDSV (Normalized Difference Spectral Vector) (Angiuli and Trianni, 2014; Trianni et al., 2015) is used to identify consistent human settlement and artificial materials. This index, which is composed out of all possible normalized indexes combining two bands with the same spatial resolution, has been shown to produce robust and globally consistent measures of urban land cover (Angiuli and Trianni, 2014). We describe these indices in detail in Appendix A.

2.3. Extract "built-up" and "not built-up" labeled examples from DMSP-OLS

DMSP-OLS sensors detect visible and near-infrared (VNIR) emissions at night. Because highly lit pixels are associated with man-made structures that emit light, we assume that pixels with DN values that exceed a threshold represent areas with built-up land cover and man-made structures. In this study we use DMSP-OLS (the "stable light" band of the 'F182013' satellite) to identify highly lit pixels. The "stable light" band discards ephemeral events, such as fires, but DMSP-OLS sensors are sensitive to persistent lighting, including from other sources, such as gas flares or volcanoes. Elvidge et al. (2009)'s Global Gas Flaring Estimates showed that while in India and Mexico there are no gas flares on land, the US has several gas flares on land, primarily in remote locations in Texas and Wyoming. We use their data to remove pixels that were classified as "built-up" from these locations during post-processing.

We sample "built-up" examples from highly lit DMSP-OLS pixels. We characterize a pixel as highly lit if its DN value exceeds a given threshold. We define this threshold as a percentile value (the 50th, 75th, 90th, 95th or the 99th percentile) of all DMSP-OLS pixels within a hex-cell. Additionally, we sample "not-built-up" examples from low-lit pixels. We characterize a DMSP-OLS pixel as low-lit if its DN value is lower than a given threshold (which we define as the 10th, 25th, 50th or the 75th percentile of all DMSP-OLS pixels within a hex-cell). We do not sample examples from DMSP-OLS pixels that are located in the intermediate buffer between highly-lit and low-lit pixels(Fig. 4).

Due to the spatial resolution of DMSP-OLS and the blooming effect, areas identified as highly lit may potentially include non-built-up land cover. We therefore exclude these types of land cover from lit pixels based on Landsat's per-pixel NDVI and NDWI values. We characterize a pixel as vegetation if its NDVI value is higher than a given threshold. Due to the heterogeneity of the land cover, we determine this threshold separately for each hex-cell. We calculate the 50th, 75th, 90th, 95th and the 99th percentile NDVI value of all pixels within the hex-cell and remove from the DMSP-OLS highly-lit pixels all Landsat pixels whose

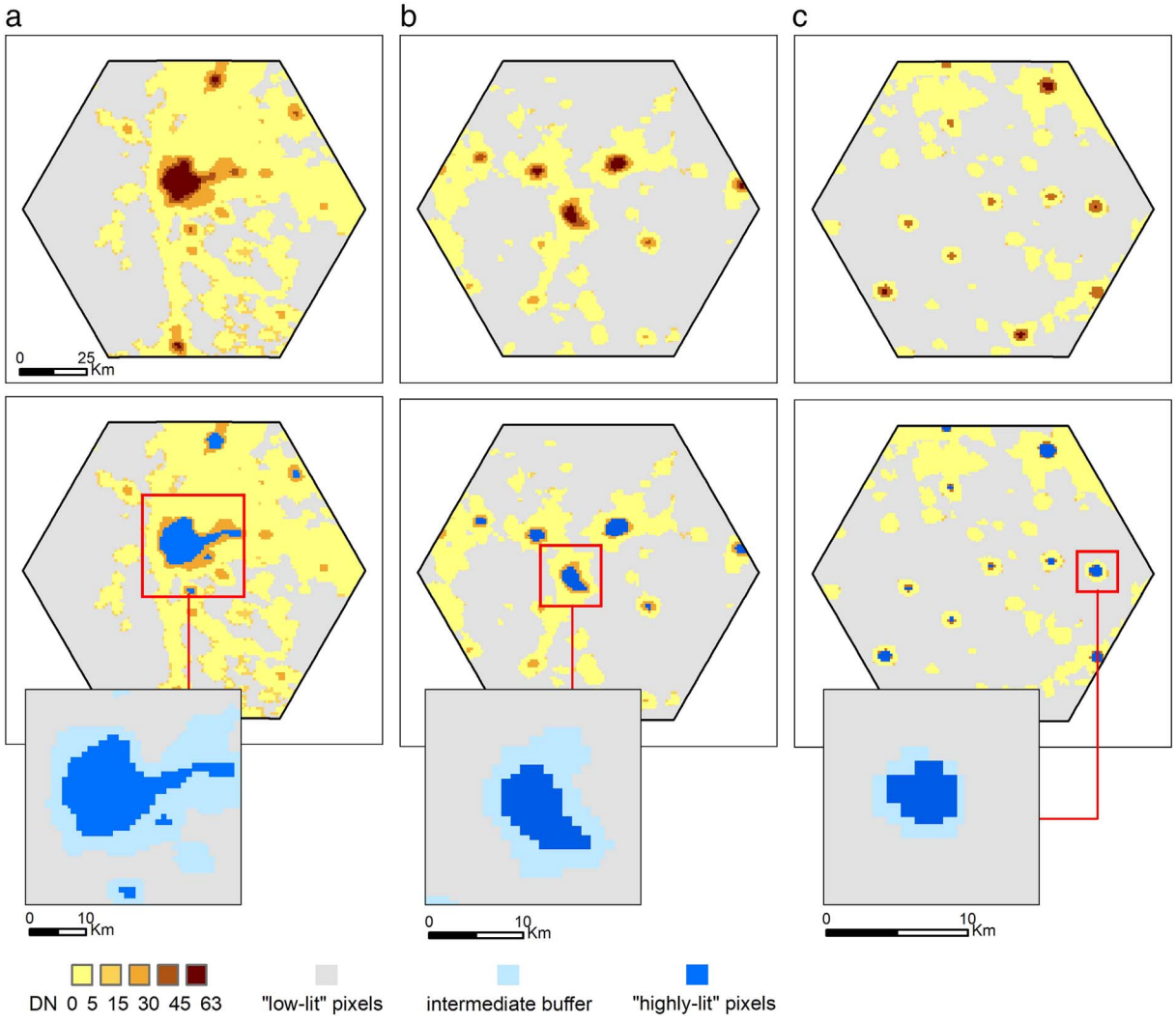

Fig. 4. The method to identify highly-lit and low-lit pixels for our classification training data. A pixel is characterized as highly-lit and as lowly-lit if its DN value exceeds and is below a given threshold, respectively. The threshold is defined as a percentile value of all DMSP-OLS pixels in the hex-cell. This example presents three 1-degree hex-cells in India. The thresholds to characterize highly-lit pixels are (a) 27, (b) 30 and (c) 17; The threshold to characterize low-lit pixels is 9. Note, that we do not sample examples from the intermediate buffer between highly-lit and low-lit pixels (marked in light blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

NDVI value exceeds this threshold. Additionally, we remove bodies of water from DMSP-OLS highly-lit pixels. We characterize a pixel as body of water if its NDWI value is positive (see Gao, 1996).

Finally, we randomly sample 100,000 pixels in each hex-cell. We create a point at the center of each Landsat pixel and associate each point with the spectral values of the Landsat composite and derived spectral indices. Each example includes a label (i.e., highly-lit or not) and the spectral values from Landsat. These training examples are used to build a classification model for each hex-cell, independently.

2.4. Perform supervised image classification

We train and classify each hex-cell using a Random Forest classifier. Random Forests are tree-based classifiers that include k decision trees and p randomly chosen predictors for each recursion. When classifying an example, its variables are run through each of the k tree predictors, and the k predictions are merged by voting on the most popular class. The learning process of the forest involves some level of randomness. Each tree is trained over a random sample of examples from the training set and each node's binary question in a tree is selected from a random subset of p input variables. We use Random Forest because previous studies find that its performance is superior to other classifiers (Goldblatt et al., 2016), especially when applied to large geographic scales and when the data are noisy and high dimensional (Gislason et al., 2006; Guan et al., 2013). Random Forests are computationally less intensive than other tree ensemble methods (Jean et al., 2016; Rodriguez-Galiano et al., 2012) and can effectively incorporate many covariates with minimum tuning and supervision (Stevens et al., 2015). We set the number of trees in the Random Forest classifier to 20. Previous studies show mixed results as for the optimal number of trees in the decision tree. The number ranges between 10 trees (Zhang et al., 2012) and 100 trees (Rodriguez-Galiano et al., 2012). Goldblatt et al. (2016) show that although the performance of Random Forest improves as the number of trees increases, this pattern holds only up to 10 trees; the performance remains stable with up to 50 or even 100 decision trees.

We map each hex-cell as an independent unit of analysis. We only sample and classify hex-cells that include at least one DMSP-OLS lit pixel (i.e., DN value above 0). These hex-cells, however, (i.e., the 1-degree and 4-degrees hexagonal grids) may also include isolated rural and suburban settlements that do not emit much light, which we still wish to capture. To ensure that they are not omitted, pixels in the excluded hex-cells are mapped according to the classifier trained with larger hex-cells (i.e., the 8-degree hexagonal grid). We clip the classification at the 8-degrees hex-cells to the extent of the smaller hex-cells that were excluded from the analysis.

The output of the classification is a per-pixel posterior probability value (i.e., the probability that a pixel is built-up). We characterize a pixel as "built-up" if the probability it is "built-up" is higher than a given threshold. Previous studies determine this threshold based on visual assessment (Schneider et al., 2010) or according to the natural breaks in the data values (Mertes et al., 2015). Here, we determine this threshold for each hex-cell using the Otsu algorithm, a clustering-based image thresholding method (Otsu, 1979). Otsu is a nonparametric and unsupervised method for automatic threshold selection that was originally developed for picture segmentation. The method uses a discriminant criterion to identify an optimal threshold that maximizes the between-class variance. Specifically here, the threshold is chosen to maximize the variance between "built-up" and "not-built-up" classes:

$$\sum_{k=1}^{2} (\overline{p}_k - \overline{p})^2 \quad (1)$$

where the threshold partitions the pixels into classes k = 1 (built-up) and k = 2 (not built-up). In the remote-sensing domain, Otsu has been used for classification of mangrove forests (Chen et al., 2013) and urban areas (Sirmacek and Unsalan, 2010).

Finally, we tile the classifications of the hex-cells to the extent of each country and post process the classification maps. We clip them to the extent of the countries' borders and remove misclassified built-up pixels from gas flaring areas using the Elvidge et al. (2009) database.

2.5. Validate and test

In each country, we assess the accuracy of the classification using hand-labeled examples. We manually label these examples (polygons, 30 m by 30 m in size) as "built-up" or as "not built-up". We define a polygon (30 m by 30 m in size) as "built-up" if the majority of its area (50% or greater) is paved or covered by human-made surfaces and used for residential, industrial, commercial, institutional, transportation, or other non-agricultural purposes. Otherwise, it is defined as "not built-up". Fig. 5 presents examples of "built-up" and "not built-up" labeled polygons. Previous studies, including Goldblatt et al. (2016), Potere et al. (2009) and Schneider et al. (2010), propose a similar definition for urban. For India, we use Goldblatt et al.'s (2016) ground-reference dataset, which includes 21,030 examples (30 m by 30 m in size) labeled as "built-up" or as "not built-up" that span the entire country (4682 polygons labeled as built-up and 16,348 labeled as not built-up). The polygons are manually labeled by two graduate students (each polygon is labeled by one student) who were provided with extensive training and supervised by the researchers. Each student was provided with an equal proportion of samples. The students labeled each polygon either as "built-up" or as "not built-up" based on a visual interpretation of the most recent available high-resolution satellite image in Google Earth (typically from 2014 to 2015).

For the US and Mexico, we construct a manually labeled ground-reference dataset of 20,000 examples (polygons, 30 m by 30 m in size) for each country. To ensure representative "built-up" and "not built-up" examples, we construct examples using a random stratified sampling procedure based on the intensity of nighttime light (the stable light band of DMSP-OLS). In each country, we identify the DMSP-OLS pixels whose value is 63 (the highest possible value). Then, we calculate a 5-pixel radius circular kernel (a buffer of approximately 5 km). In each country, we select a random sample of points: 15,000 from the DMSP-OLS pixels with a DN value 63 and 5000 from the 5 km buffer zone. A 30 × 30 m polygon is drawn around each sample point. We overlay these polygons onto high-resolution imagery in Google Earth. The US and Mexico examples are labeled by six undergraduate students (three students per country), each polygon is labeled either by one (in Mexico) or by three students (in the US). The students were provided with extensive training and were supervised by the researchers. The students were instructed to flag ambiguous polygons where the definition of the polygon as built-up or not built up is not absolutely clear; we exclude these polygons from the dataset. We randomly assign half of the dataset in each country as the test set (to determine optimal parameters to the classifier) and half as the validation set (to validate the final classified map). Between 22% and 27% of the polygons are labeled as built-up (Table 1).

Previous studies (Liu and Zhou, 2004; Rwanga and Ndambuki, 2017; Schneider et al., 2009; Zhang and Seto, 2011) rely on hand-

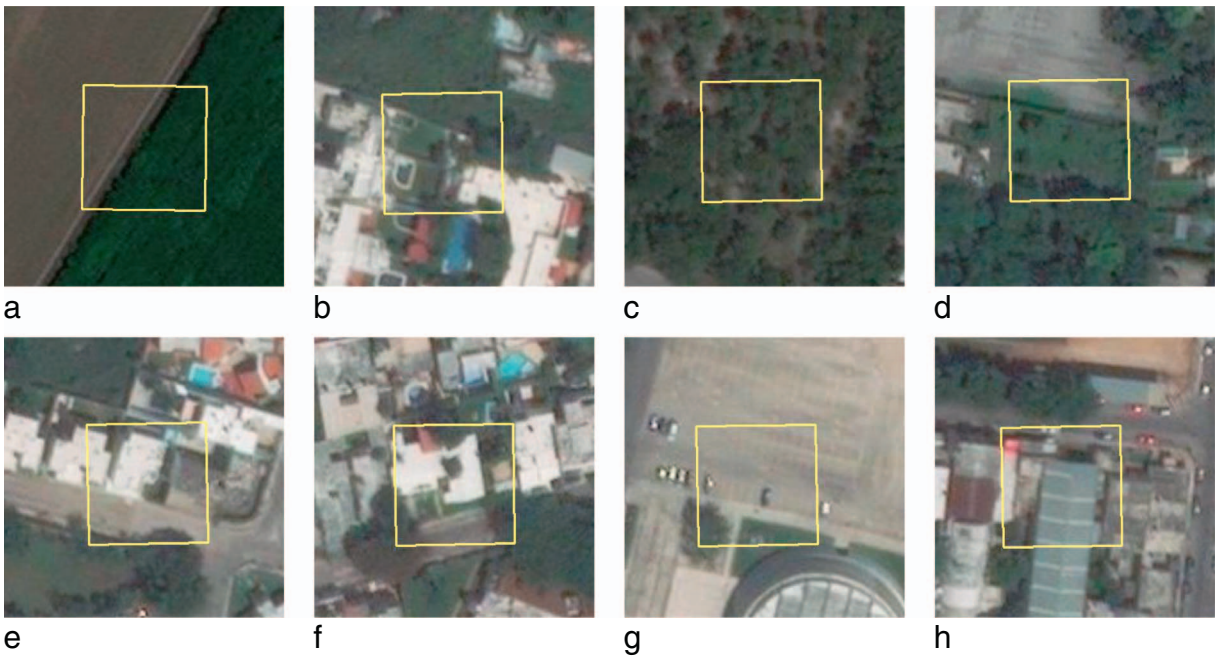

Fig. 5. Ground-reference labeled examples. Each example (a polygon 30 m × 30 m in size) is labeled as "not built-up" (top) and as "built-up" (bottom). To illustrate, the examples in Fig. 5a, b and d are labeled as "not built-up" (the majority of the polygon in Fig. 5a is an agriculture field, the majority of the polygon in Fig. 5b is a backyard covered with grass, and the majority of the polygon in Fig. 5d is covered with grass). The examples in Fig. 5e and g are labeled as "built-up" (the majority of the polygon in Fig. 5e is a residential building and a paved road; the majority of the polygon in Fig. 5g is a paved parking lot).

Table 1
the distribution of the built-up and not built-up points for each country.

	BU points	NBU points	Total
Hand-labeled polygons			
India	4682 (22.26%)	16,348 (77.74%)	21,030
US	4386 (21.6%)	15,898 (78.4%)	20,284
Mexico	5477 (27.4%)	14,523 (72.6%)	20,000
Tomnod-labeled polygons			
India	2276 (23%)	7602 (77%)	9878
US	3305 (30.7%)	7456 (69.3%)	10,761
Mexico	829 (31.8%)	1782 (68.2%)	2611
			84,564

labeled examples for accuracy assessment of classification maps. However, a key challenge in any machine learning task is “class noise” which may be caused by a subjective interpretation of the target label (Foody, 2002). To assure the validity of our hand-labeled examples, we add an additional source of reference data to the final accuracy of our classification. We utilize DigitalGlobe's Tomnod platform to label 23,250 examples (referred to as Tomnod examples). The examples are labeled as “built-up” or as “not built-up” according to Tomnod's CrowdRank algorithm (Wood and Har-Noy, 2013), a measure designed to estimate crowd consensus and to assure that a sufficient number of labeling individuals reach a consensus for each labeled examples. The rank is also used to verify the accuracy and reliability of individual contributors. Together with the hand-labeled examples, the dataset we use for test and for validation consists of 84,564 labeled examples in total (Table 1).

We use half of the hand-labeled dataset (the test set) to assess alternative parameters for the classifiers, and the other half (the validation set), in addition to the Tomnod examples, to evaluate and report its performance and the accuracy of the final classification map. We estimate the performance of the classifiers with several performance estimators. We refer to the class “built-up” (BU) as positive and to the class “not built-up” (NBU) as negative: (1) Producer's accuracy, referred to as True-Positive Rate (TPR), the percentage of actual BU examples classified correctly as BU; (2) True-Negative Rate (TNR), the percentage of actual NBU examples classified correctly as NBU; (3) Balanced Accuracy, the average of TPR and TNR; and (4) User's accuracy, the percentage of correctly classified BU examples of all examples predicted as BU.

$$\text{Producer's accuracy (TPR)} = TP/(TP + FN) \quad (2)$$

$$TNR = TN/(TN + FP) \quad (3)$$

$$\text{Balanced Accuracy} = (TPR + TNR)/2 \quad (4)$$

$$\text{User's accuracy} = TP/(TP + FP) \quad (5)$$

where TP is the number of the actual BU examples predicted to be BU; TN is the NBU examples predicted as NBU; FN is the actual BU examples predicted as NBU and FP is the actual NBU examples predicted as BU.

2.5.1. Optimal parameters to the classifier

We use half of the hand-labeled dataset (the test set) to classify 50 1-degree hex-cells per country and to assess alternative parameters for the classifiers. We select hex-cells for which we have the largest number of hand labeled examples (at least 35 examples). This accuracy assessment procedure is designed to identify (1) the optimal percentile values to characterize a DMSP-OLS pixel as highly-lit and as lowly-lit; (2) the optimal percentile value to characterize a Landsat pixel as containing vegetation (according to NDVI); and (3) the optimal input features to the classifier. We proceed in the following steps, also show in Fig. 6:

a. **Define a Landsat pixel as vegetation**: First, we define highly-lit and low-lit DMSP-OLS pixels as pixels whose DN values are above and below a constant percentile value (the 99th percentile) of all DMSP-OLS pixels in a hexagon, respectively. We remove from highly-lit pixels bodies of water (i.e. Landsat pixels with a positive NDWI value) and vegetation. We define a Landsat pixel as vegetation if its NDVI value exceeds either the 50th, 75th, 90th, 95th or the 99th percentile of the NDVI value of all Landsat pixels in a hex-cell. We perform classification with Landsat spectral bands, together with five spectral indices (NDVI, NDWI, NDBI, EVI, UI) as inputs to the classifier. Based on these results, we determine the optimal percentile value to characterize a Landsat pixel as vegetation.
b. **Classify highly-lit pixels**: We define a DMSP-OLS pixel as highly-lit if its DN value exceeds either the 50th, 75th, 90th, 95th or the 99th percentile value of all DMSP-OLS pixels within a hex-cell (below this value, the pixel is defined as low-lit). We remove bodies of water (i.e., a positive NDWI value) and vegetation (according to the optimal NDVI percentile value found in step (a)) and perform classification with Landsat's band values together with five spectral indices (NDVI, NDWI, NDBI, EVI, UI) as input (we do not include NDSV here). These results indicate the optimal percentile threshold to characterize a DMSP-OLS pixel as highly lit. Pixels below the threshold are classified as lowly lit.
c. **Determine inputs to the classifier**: We define a DMSP-OLS pixel as highly-lit according to procedure in (b) and remove bodies of water (i.e., a positive NDWI value) and vegetation (according to the definition of a Landsat's vegetation pixel from step (a)). We assess the performance of the classifiers with five combinations of inputs: (1) only Landsat band values (without additional spectral indices); Landsat bands together with (2) NDVI; (3) NDBI; (4) NDVI, NDWI, NDBI, EVI, UI; and (5) NDSV. The result of this step indicate the optimal inputs to the classifier.
d. **Classify lowly-lit pixels**: In the procedure above, we define a pixel as lowly lit if it is not highly lit. However, this definition may result in mislabeled pixels as “not built-up”, especially in proximity to highly-lit pixels in the core of settlements. In the final step of the parameter evaluation, we sample “not built-up” examples only from DMSP-OLS pixels whose DN value is below either the 10th, 25th, 50th and 75th percentile value of all DMSP-OLS pixels in a hex-cell. We perform classification using the optimal parameters found in steps (a) through (c). Based on these results, we determine the threshold to identify a DMSP-OLS pixel as lowly-lit.

3. Results

3.1. Optimal parameters to classifier

As explained above, we first select in each country 50 1-degree hex-cells in order to identify the optimal parameters to the classifier. We classify these hex-cells and assess the accuracy of the classification with the test set (the results of the accuracy assessment with each of the examined parameter are presented in Appendix B).

The results show variations between countries in the optimal parameters to the classifiers. First, we find that the countries differ in the threshold to characterize a DMSP-OLS pixel as highly lit. In the US and in Mexico the optimal threshold is the 90th percentile value of all DMSP-OLS pixels in a hex-cell, while in India it is the 95th percentile. We find that the optimal threshold to define a pixel as lowly lit is the 75th percentile value in the three countries. Additionally, the countries differ in the optimal threshold to define a Landsat pixel as vegetation. In India and the US, the best performance is achieved when Landsat pixels are defined as vegetation if their NDVI value exceeds the 90th percentile of all pixels in a hex-cell, while in Mexico this threshold is lower (the 75th percentile). We remove vegetation Landsat pixels from highly-lit DMSP-OLS pixels. Moreover, the results show variations between countries in the optimal inputs to the classifiers. Classification with NDVI, NDWI, NDBI, EVI, UI as inputs to the classifier results in the best performance (determined by the highest balanced accuracy) for the

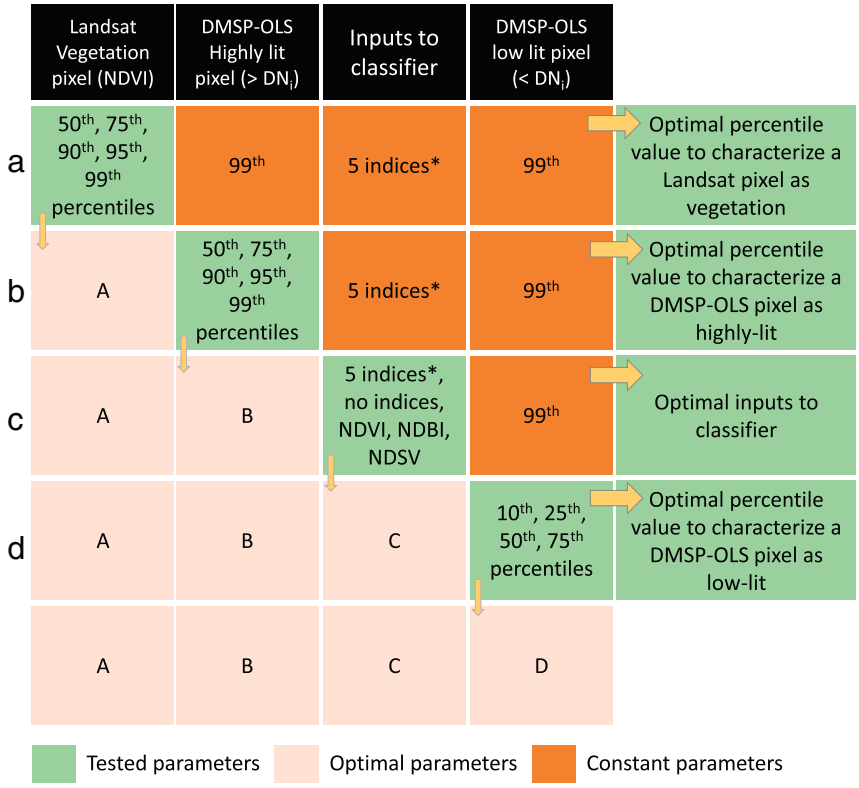

Fig. 6. The procedure to identify the optimal parameters to the classifier. We select in each country 50 hexagon (1-decimal-degree from center to center) and identify (1) the optimal percentile value to characterize a DMSP-OLS pixel as highly-lit and as lowly-lit; (2) the optimal percentile value to characterize a Landsat pixel as containing vegetation (according to NDVI); and (3) the optimal input features to the classifier. Each experiment (B-D) relies on the optimal parameters found in the previous experiments. (* The five indices are NDVI, NDBI, UI, NDWI, EVI).

Table 2
The optimal parameters to the classifiers, per country (India, US and Mexico): percentile NDVI value to characterize a Landsat pixel as vegetation, percentile DN value to characterize a DMSP-OLS pixel as highly-lit or low-lit and inputs to the classifier.

	NDVI threshold (percentiles)	Highly-lit pixels (percentiles)	Inputs to classifier	Low-lit pixels (percentiles)
India	90th	95th	NDSV	75th
US	90th	90th	NDVI, NDWI, NDBI, EVI, UI	75th
Mexico	75th	90th	NDVI, NDWI, NDBI, EVI, UI	75th

Table 3
Performance measures for classification of 50 1-degree hex-cells (classification with the optimal parameters to the classifier): True positive rate (TPR), True negative rate (TNR), balanced accuracy and User's accuracy.

	TPR (producer's accuracy)	TNR	Balanced accuracy	User's accuracy
India	77.5%	84.1%	80.8%	60.1%
US	82.5%	75.1%	78.8%	67.3%
Mexico	88.7%	78.9%	83.8%	67.6%

US and Mexico, while classification with NDSV as input to the classifier results in the best performance of the classifiers for India. With these parameters (Table 2), classification of the 50 hex-cells yields an accurate classification of built-up land cover, indicated by a high balanced accuracy of 79%, 81% and 84% for classification in the US, India and Mexico, respectively (Table 3). User's accuracy ranges between 60% in India to 67% and 68% in the US and Mexico, and the TNR and TPR measures are above 75% in all countries.

3.2. The Otsu method

Random forests predict for each new example the probability it is a positive example (in this case, a built-up pixel). This is a posterior probability that ranges between 0 and 1. A pixel is defined as built-up if the probability it is built-up exceeds a given threshold, which we identify for each hex-cell using the Otsu method. The Otsu method has shown to provide satisfactory results for thresholding an image when the histogram of the pixels values is characterized by a non-normal distribution, especially bimodal (Ng, 2006).

As clearly illustrated in Fig. 7 (the distribution of the posterior probability values of all pixels within 12 hexagons), the distribution of the posterior probability values is symmetric (U shaped), with two high peaks in each histogram (hex-cell) that are separated by deep valleys. These conditions allow for an effective performance of the Otsu threshold methods in separating the left side of the histogram (low posterior probability values) from the right side of the histogram (high posterior provability values).

To summarize, classification of 50 hex-cells in each country shows variations in the optimal parameters to the classifier. The countries differ in the optimal definition of DMSP-OLS pixels as highly-lit, as well as in the definition of a Landsat pixel as vegetation and in the optimal inputs to the classifiers. Based on these results, we next map the built-up land cover in the entire area of each country.

3.3. Classification of the entire area of each country

In the experiments described above, we treated each hex-cell (1-degree distance from center to center) as the unit of analysis for training and for classification. To determine the optimal scale of the hex-cells in the hexagonal grid division, we next classify the entire area of each country using three sizes of hexagonal grid divisions: 1, 4, and 8 decimal degrees from center to center. We calculate the thresholds to characterize highly-lit and lowly-lit DMSP-OLS pixels within each (1, 4 or 8°) hex-cell. As shown in Fig. 8 (the spatial distribution of the DN threshold to determine a highly-lit pixel, per hex-cell scale and per

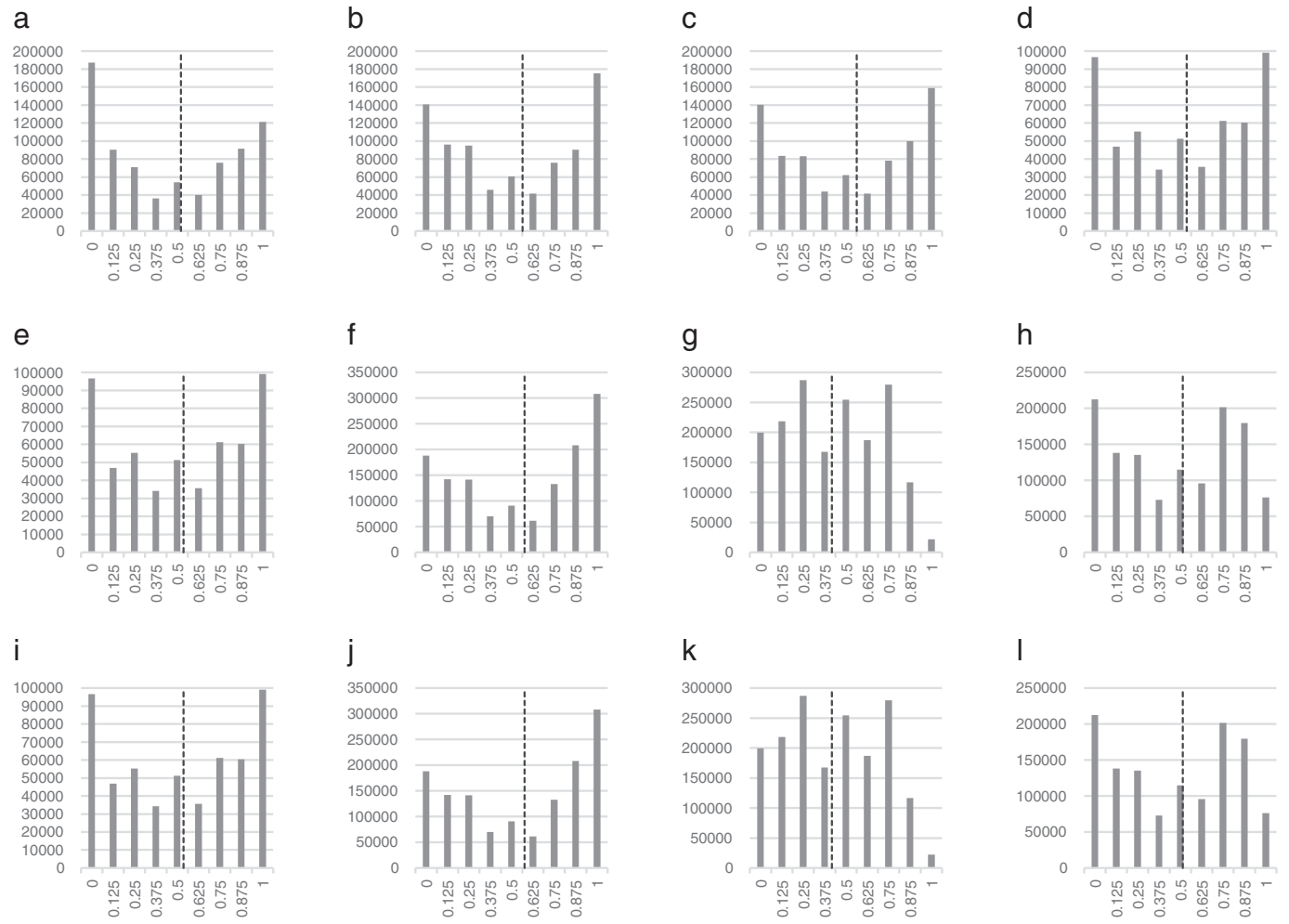

Fig. 7. The histograms of the posterior probability values of all pixels in 12 1-degree hex-cell in India (a–d), US (e–h) and Mexico (i–l). The dashed line indicates the posterior probability threshold according to Otsu method.

country), there is large variation across and within countries. This variation is also clearly illustrated in Fig. 9, the distribution histogram of the DN threshold to determine highly-lit and low-lit pixels.

Next, we mask out vegetation and bodies of water from highly-lit pixels, select built-up and not built-up examples, and perform supervised image classification with NDVI, NDWI, NDBI, EVI, UI (in the US and Mexico) and NDSV (in India) as additional inputs to the classifier. We classify each hex-cell as an independent unit of analysis and predict, for each pixel, the probability (posterior probability) it is built-up. The mean value of the posterior probability ranges between 0.29 and 0.43 and the distribution of the posterior probability threshold within each country is skewed left. Fig. 10 presents the distribution of the per hex-cell posterior probability threshold in each country, for each hex-cell scale.

As shown in Table 4, there are also variations between countries in the optimal hex-cell scale. Classification with the smallest level of hex-cell (1-degree) yields the best performance in India and Mexico, while classification with 4-degrees hex-cell scale yields the best performance in the US (though in Mexico, classification with 1-degree hex-cell only marginally exceeds classification with 4-degree hex-cells, by 0.4%). Of the three countries, we find the highest accuracy performance for classification in the US, indicated by a balanced accuracy of 82% and Producer's accuracy of 73% (for classification with 4-degree hex-cells). The balanced accuracy in India and in Mexico is a bit lower (around 77%) and the Producer's accuracy ranges between 61% to 63% (in India and Mexico, respectively, for classification with 1-degree hex-cells). When classification is performed with 8-degree hex-cells, the balanced accuracy falls by up to 4.6% and the Producer's accuracy drops more significantly by up to 13.4% (for example, from 63.2% to 49.8% in India). The countries also differ in the level of hex-cell that results in the highest User's accuracy rate. For example, we find the highest User's accuracy in India for classification with hex-cell level of 4-degree while in Mexico for classification with hex-cell level of 1-degree (70.4% and 76.7%, respectively).

3.4. Assessment of DMSP-OLS characterization of highly- and low-lit pixels

To evaluate the accuracy of our characterization of highly-lit and low-lit DMSP-OLS pixels, we perform an additional per-hex-cell accuracy assessment. In each country, we adopt the optimal parameters to the classifier (i.e., hex-cell scale, threshold to characterize a Landsat pixel as vegetation and threshold to characterize a DMSP-OLS pixel as highly-lit or as lowly-lit) to identify highly-lit DMSP-OLS pixels (i.e., locations of built-up examples) and lowly-lit pixels (i.e., locations of not built-up examples). We compare this characterization with our validation set. This analysis is intended to assess our definition of highly-lit and low-lit regions from which we sample built-up and not built-up examples (we relate to the characterization of DMSP-OLS pixels as the predicted class). We find that the balanced accuracy ranges between 67.2% in Mexico to 77% and 79.3% in India and the US, respectively. Furthermore, when vegetation and bodies of water are not excluded from highly-lit pixels the balanced accuracy drops in all countries, by

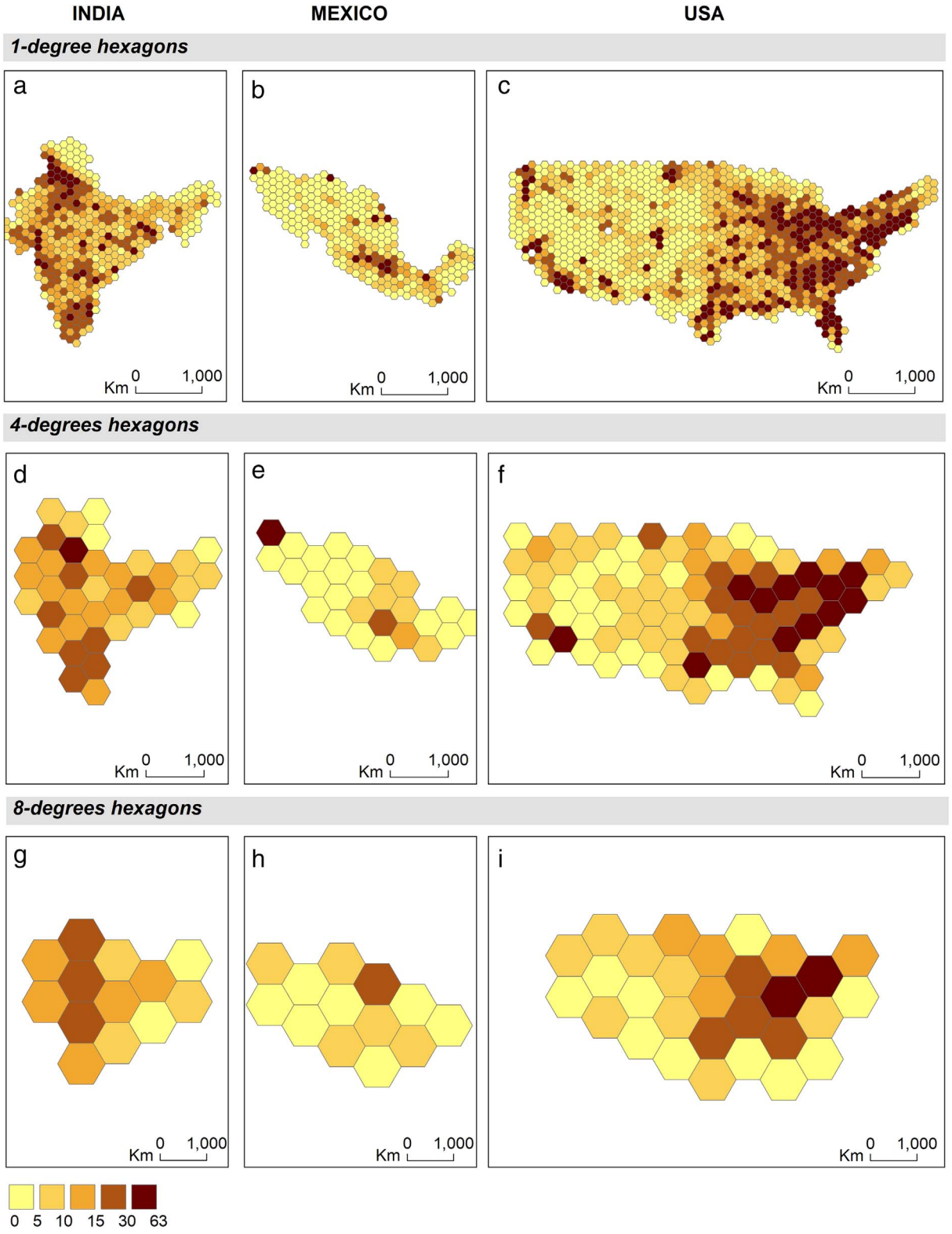

Fig. 8. The threshold to identify a pixel as highly-lit, per level of hex-cells (1, 4, and 8°) and per country (India, Mexico and the US).

up to 7.5%. By their nature, DMSP-OLS data are relatively "noisy" and suffer from many limitations (e.g., coarse spatial resolution, blooming effect), yet by relying on local thresholds to characterize highly-lit and lowly-lit pixels as well as by removing vegetation land cover and bodies of water, we improve the identification of built-up and not built-up locations.

3.5. Accuracy assessment with validation set

The results indicate that the smallest level of hex-cells (1-degree) results in the best performance of the classifiers in India and in Mexico, while classification with 4-degrees hex-cells results in the best performance in the US. Before evaluating these classification maps with the validation set and with the Tomnod examples, we post-process these

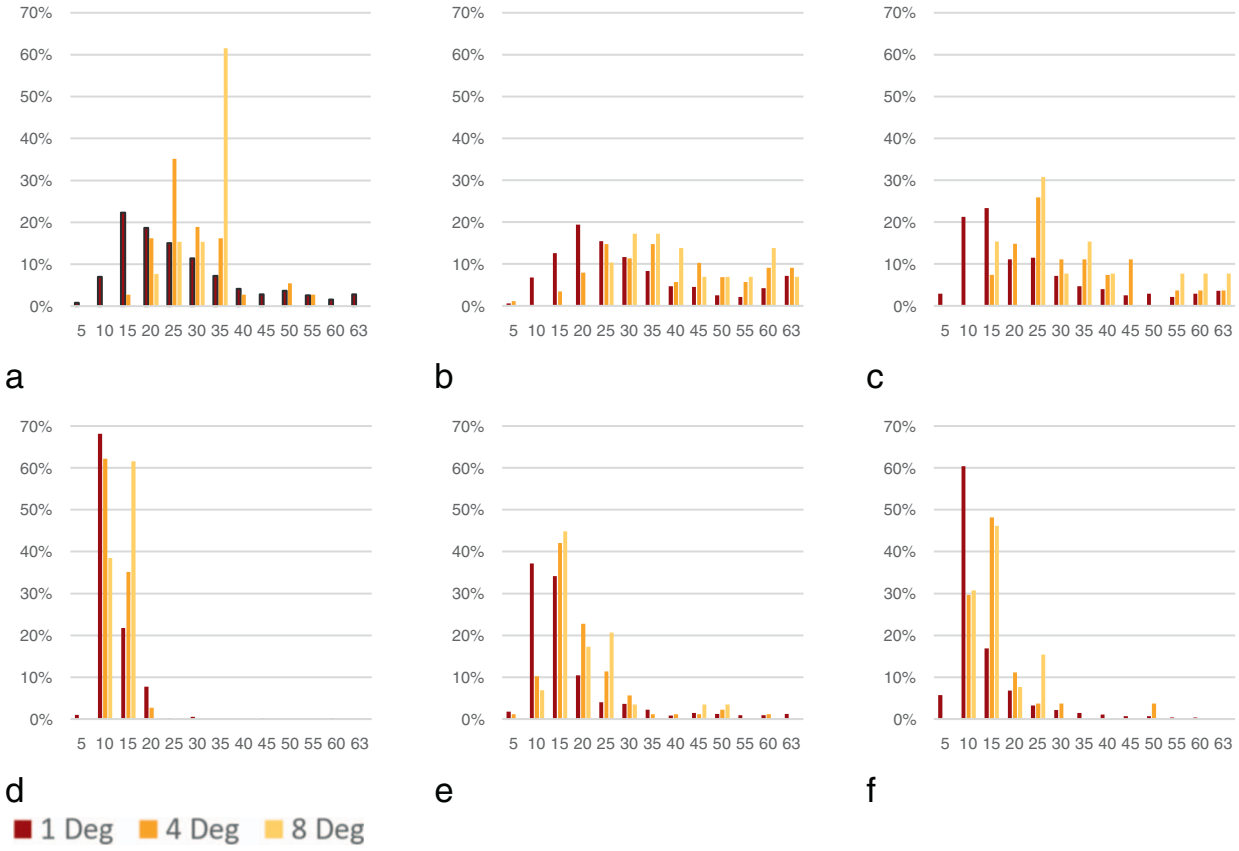

Fig. 9. The distribution of the threshold to identify a pixel as highly-lit ((a)India; (b)US; (c) Mexico), and low-lit ((d) India; (e) US; (f) Mexico) per level of hex-cells (1, 4, and 8°).

maps. We clip them at the extent of each country's border, and remove misclassified built-up pixels from areas with gas flares using the Elvidge et al. (2009) database. The final post-processed classification maps span the full area of the three countries (Fig. 11).

The confusion matrix for the accuracy assessment with the validation set is presented in Table 5. The results show a high balanced accuracy rate of around 75% in India and Mexico and around 81% in the US (Table 5). The accuracy measures for assessment with our validation set and Tomnod examples are very close (Table 5). User's accuracy ranges between 66% and 77% and the overall accuracy measure is above 80% in all countries, (Table 5).

4. Discussion

We present a machine learning approach to map built-up areas at a large geographic scale. Our methodology utilizes nighttime-light data (derived from DMSP-OLS) as a source for training examples of built-up and non-built-up areas. These examples are used for supervised image classification in Landsat 8 imagery. To our knowledge, this is the first study to develop a practical and simple pixel-based image classification of built-up areas that utilizes nighttime light data as a source of training examples and high-resolution daytime imagery as the input source for the classifier. This methodology can be applied to map built-up areas across space. Although many classification products map urban land, they are typically limited in their temporal and/or spatial resolution

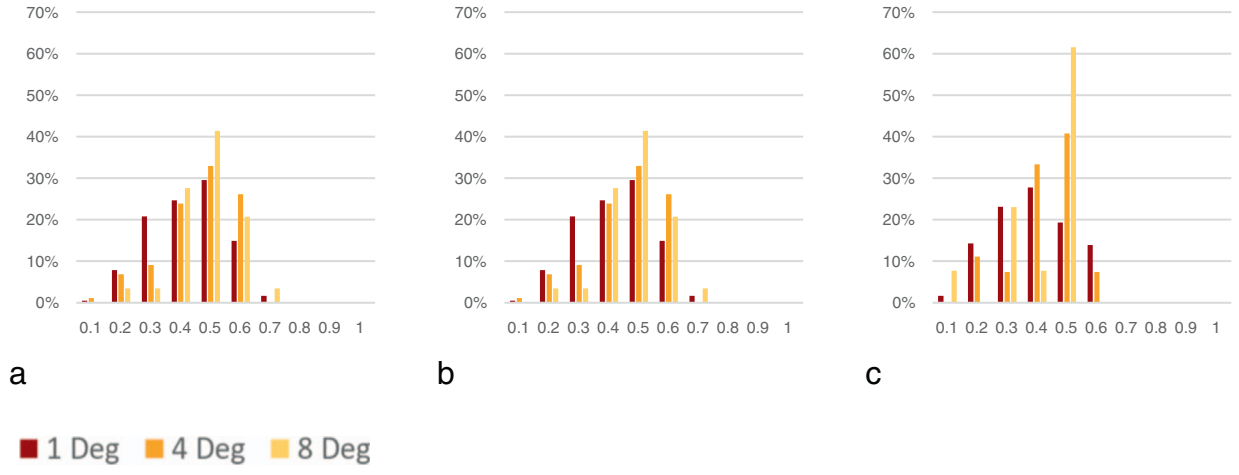

Fig. 10. The distribution of the posterior probability threshold to determine a pixel as “built-up” in India, the US and Mexico.

Table 4
Performance measures as a factor of the hex-cell level for India, the US, and Mexico. Accuracy assessment with test set (classification of all hexagons in each country).

	Overall accuracy	TPR (producer's accuracy)	TNR	Balanced	User's accuracy
India					
8°	84.2%	49.8%	94.0%	71.9%	70.3%
4°	84.7%	53.9%	93.5%	73.7%	70.4%
1°	83.8%	63.2%	89.7%	76.5%	63.9%
US					
8°	87.2%	68.8%	92.7%	80.8%	73.6%
4°	86.8%	72.8%	91.0%	81.9%	70.6%
1°	85.9%	71.1%	90.4%	80.7%	68.7%
Mexico					
8°	80.5%	64.1%	87.7%	75.9%	69.6%
4°	80.9%	64.1%	88.3%	76.2%	70.6%
1°	82.4%	61.4%	91.7%	76.6%	76.7%

which limits their use to track urbanization processes over time.

Mapping built-up areas at a global scale is challenging because of the scarcity of extensive ground-reference data for supervised classification and validation. Crowd-sourced datasets, such as OpenStreetMap (OSM) can also be used to map urban areas (Belgiu and Drăguţ, 2014; Estima and Painho, 2015). OSM is a valuable source for ground-reference data because of its geographic extent and availability, but OSM is subject to concerns about completeness and reliability (Schlesinger, 2015). The use of OSM for supervised image classification remains challenging due to the risk of imbalanced distribution of class labels (including spatial coverage), the presence of errors or missing class assignments ("class-noise"), and inaccurate polygon boundary delineations (Johnson and Iizuka, 2016). Our methodology overcomes the lack of such data by utilizing low-resolution DMSP-OLS data for classification of built-up areas in Landsat imagery. By partitioning countries into smaller regions, we allow the parameters of the classification model to vary across space. The results indicate that the three examined countries differ in the optimal hex-cell scale, or the size of the hex-cell that is used as the unit of analysis (for training and for classification). Classification at a local scale (a hex-cell) allows one to account for heterogeneity in environmental conditions. Although many studies address the effect of the classifiers' hyper-parameters on their performance, here we show that classifiers also have an optimal spatial scale that can be and should be discovered through techniques similar to those we employ here. We leave the choice of the classifier scale for each pixel for future research.

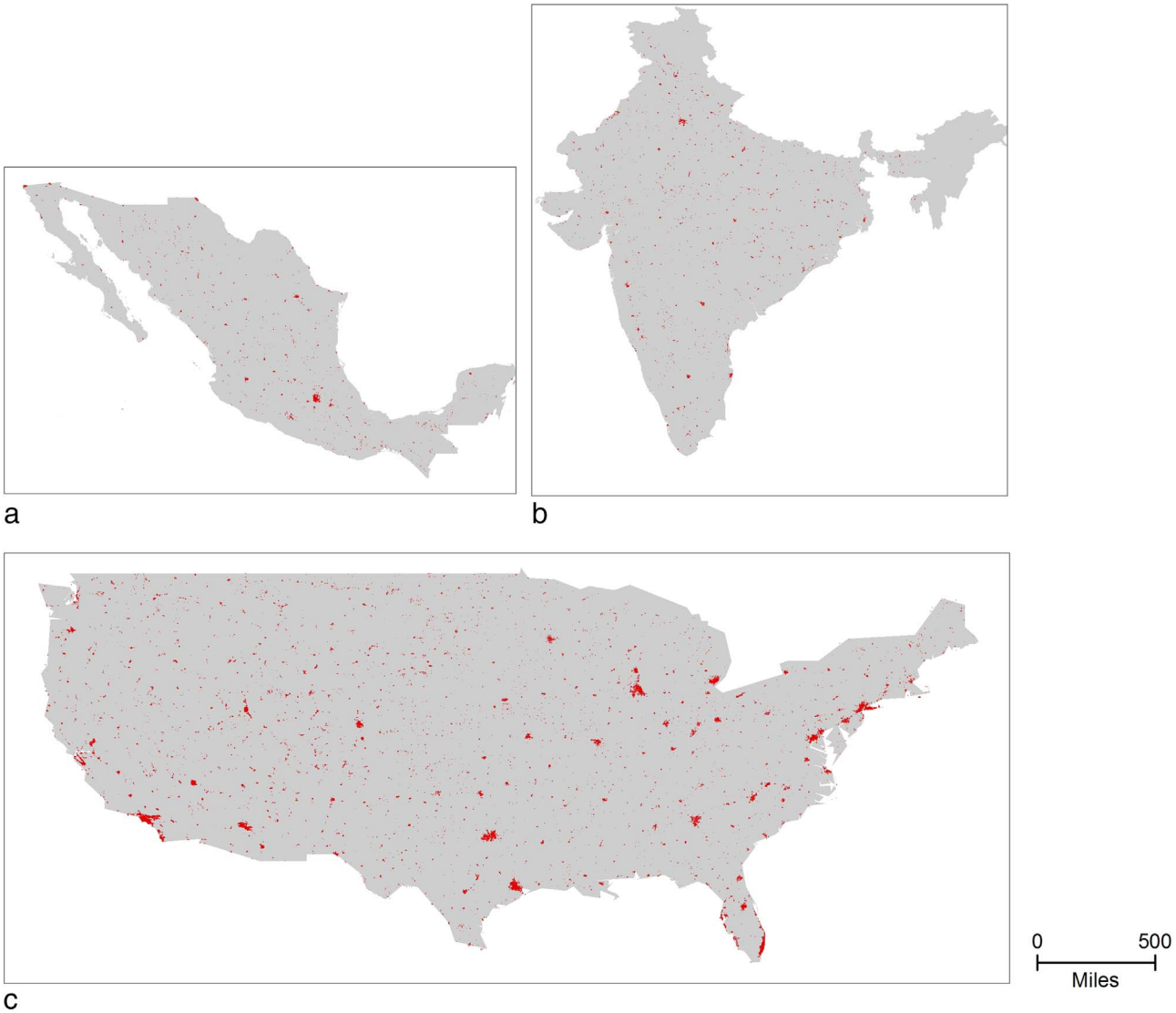

Fig. 11. Classified built-up land cover (in red) across the full countries: (a) Mexico, (b) India and (c) the US. Post-processed classification maps, using 1-degree hex-cells. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Performance measures as a factor of the hex-cell level for India, the US, and Mexico after post processing. Accuracy assessment with validation set and with Tomnod examples.

	Overall accuracy	TPR (producer's accuracy)	TNR	Balanced	User's accuracy
Post processed maps – validation with validation set					
India (1-dgree)	84.0%	59.6%	91.0%	75.3%	65.6%
US (4-degree)	86.2%	70.7%	90.8%	80.7%	69.0%
Mexico (1-degree)	82.0%	58.6%	92.3%	75.4%	77.1%
Post processed maps – validation with Tomnod set					
India (1-dgree)	83.6%	56.5%	91.7%	74.1%	67.2%
US (4-degree)	85.9%	72.1%	92.1%	82.1%	80.3%
Mexico (1-degree)	80.2%	55.0%	91.9%	73.5%	76.0%

Confusion matrix for Post processed maps (validation with validation set)*

		India			US			Mexico		
		Predicted								
		BU	NBU	SUM	BU	NBU	SUM	BU	NBU	SUM
Actual	BU	1521	1031	2552	1734	719	2453	1961	1385	3346
	NBU	799	8110	8909	778	7648	8426	584	6993	7577
	Sum	2320	9141	11,461	2512	8367	10,879	2545	8378	10,923

* Confusion matrices for classification with examples in the validation set. Comparison between predicted and actual built-up (BU) and not-built-up (NBU) examples.

We assess the validity of our approach using an external validation that uses a total of around 84,500 hand-labeled examples. The results demonstrate the robustness of our approach and its applicability in heterogeneous regions. Yet, the degree to which the "localization" of our classifiers affects their performance varies between regions and depends on the characteristics and heterogeneity of the mapped landscape. For example, because of the relatively similar spectral characteristics of built-up areas and bare land, it is often hard to differentiate between these land cover types through remote sensing, especially in regions characterized by heterogenous land cover (e.g. complex pattern of built-up areas, bare land and vegetation) (He et al., 2010). Thus, differences in the complexity of the land cover within and between countries, as well as between different hexagonal scales, may affect the accuracy of the classification as we observe in this study. Applying this method at a global scale will require automatic methods for selecting, in any pixel, the scale of classification that maximizes accuracy. Of course, the accuracy of the classification also depends on the quality of the ground-reference data. Our methodology performs well especially when built-up and not built-up examples can be sufficiently differentiated in highly-lit areas. Because we rely on NDVI and NDWI indices to separate between built-up and not built-up examples in highly-lit DMSP-OLS pixels, our methodology may work best in regions that are characterized by vegetation land cover. Future research could rely on our method to improve accuracy in settings that are characterized by bare land or bare soil, for example, by utilizing the Enhanced Built-Up and Bareness Index (EBBI) (As-Syakur et al., 2012). In addition, due to the lack of on-board calibration and unstable radiometric performance of the DMSP-OLS sensors, the absolute radiance of light cannot directly represent temporal changes in the intensity of the light, and thus, inter-sensor calibration is required to make our approach operational in time.

Our methodology overcomes the need for expensive hand-labeled data for supervised classification as well as many of the limitations associated with DMSP-OLS data. First, by identifying local per-hex-cell thresholds to characterize DMSP-OLS pixels as highly-lit, we account for the blooming effect and identify the boundaries of highly-lit settlements. As illustrated in Fig. 12, due to the blooming effect, lit areas are consistently larger than the built-up land cover they are associated with. Second, by identifying vegetation land cover and bodies of water according to NDVI and NDWI indices, we minimize the effect of the over saturation of DMSP-OLS in city centers, which allows us to capture the fine boundaries of built-up areas with high precision (Fig. 13). We note that although the ground-reference data are not free of measurement error, this methodology allows us to collect millions of labeled examples. This large sample is valuable for training a classifier to map built-up land cover.

We used 60,000 labeled examples (polygons) to assess the accuracy of our classification compared to the definition of urban and built-up land cover according to MCD12Q1 UMD MODIS classification scheme. The results show that the accuracy of our classification exceeds MCD12Q1 UMD MODIS classification scheme, indicated by a higher balanced accuracy of between 2.5%–5.6% (Table 6). As illustrated in Fig. 14, the extent of the built-up land cover that we detect with our methodology is larger than the extent of the land cover that is classified as urban and built-up by the MCD12Q1 classification.

Our classification also exceeds other national high-resolution land-cover and land-use maps. To illustrate, a comparison between our examples and the US National Land Cover Database (NLCD) classification map showed a lower balanced accuracy rate in the NLCD product (72.3%) (because we define a polygon as "built-up" if 50% or greater of its area is built, we compare our classification to the NLCD classes "Developed, Medium Intensity" and "Developed High Intensity" as "built-up"). Fig. 15 presents a comparison between our classification and NLCD's classification of medium and high- intensity in three case studies in the US.

Finally, classification of urban land cover can be performed using Soft Classification (SC) techniques. This technique is the creation of a continuous surface where each pixel is characterized by the fraction of urban land, or impervious surface that is incorporated within it (Huang et al., 2016; MacLachlan et al., 2017; Song et al., 2016; Van de Voorde et al., 2011). Alternatively, one can use Hard Classification (HC) techniques, where the extent of urban land is characterized as discrete per-pixel binary class (presence/absence of urban land). Many of the existing global urban maps (e.g., the Global Rural-Urban Mapping Project (GRUMP), MODIS Urban Land cover and GLC00) represent the surface on Earth according to a discrete thematic classification (Potere et al., 2009). In part, this is because many of the common classifiers, such as Random Forests, predict for each pixel the probability (posterior probability) it belongs to a given class (e.g., a pixel is characterized as urban if the probability it is urban exceeds a given threshold). However, a discrete map can further be translated into percentage of urban land, for example, by aggregating high-resolution binary classification to a

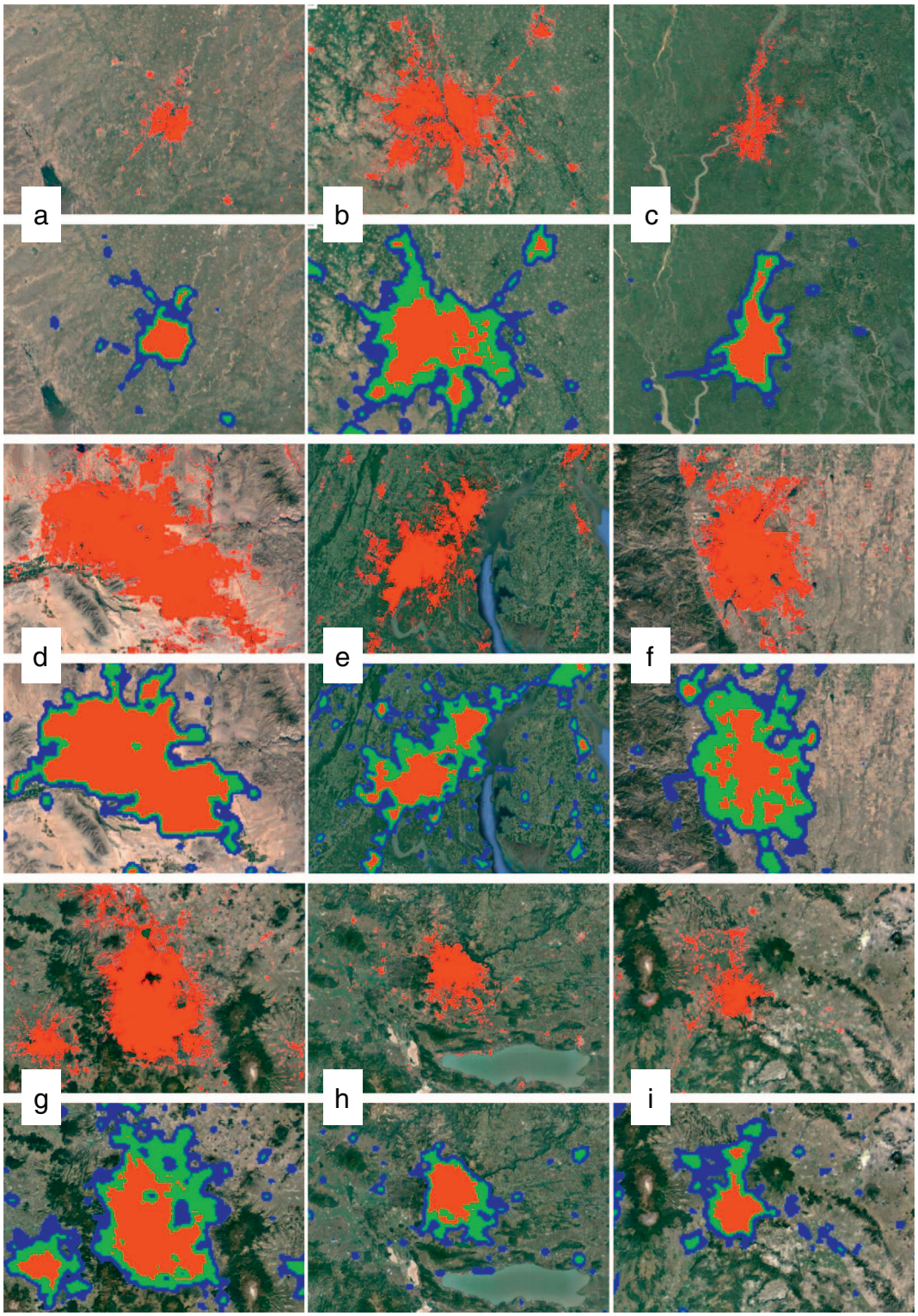

Fig. 12. A comparison between our classification of built-up areas and lit pixels according to DMSP-OLS in (a) Ahmedabad, Gujarat, India; (b) New Delhi, Delhi, India; (c) Kolkata, West Bengal, India; (d) Phoenix, Arizona, US; (e) Washington DC, US; (f) Denver, Colorado, US; (g) Mexico City, Mexico; (h) Guadalajara, Mexico; (i) Puebla, Mexico. The top figure in each city presents our classification; the bottom figure presents the DN values of DMSP-OLS stable lights band: 30–55 (blue), 56–61 (green), 63 (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

continuous coarser resolution surface representing per pixel percentage urban land (Schneider et al., 2009, 2010; Zhou et al., 2015). To illustrate how our classification maps can be translated into a continuous characterization of built-up land cover, we create an artificial grid of 500 m ∗ 500 m cells and calculate the fraction of built-up land cover (according to our classification) in each cell. Fig. 16 presents, as an example, per-pixel fractions of built-up land cover in nine cities in our investigated countries. The figure demonstrates variations in built-up density, both within and between cities. Some cities are characterized by a clear gradient from the center of the city towards the periphery

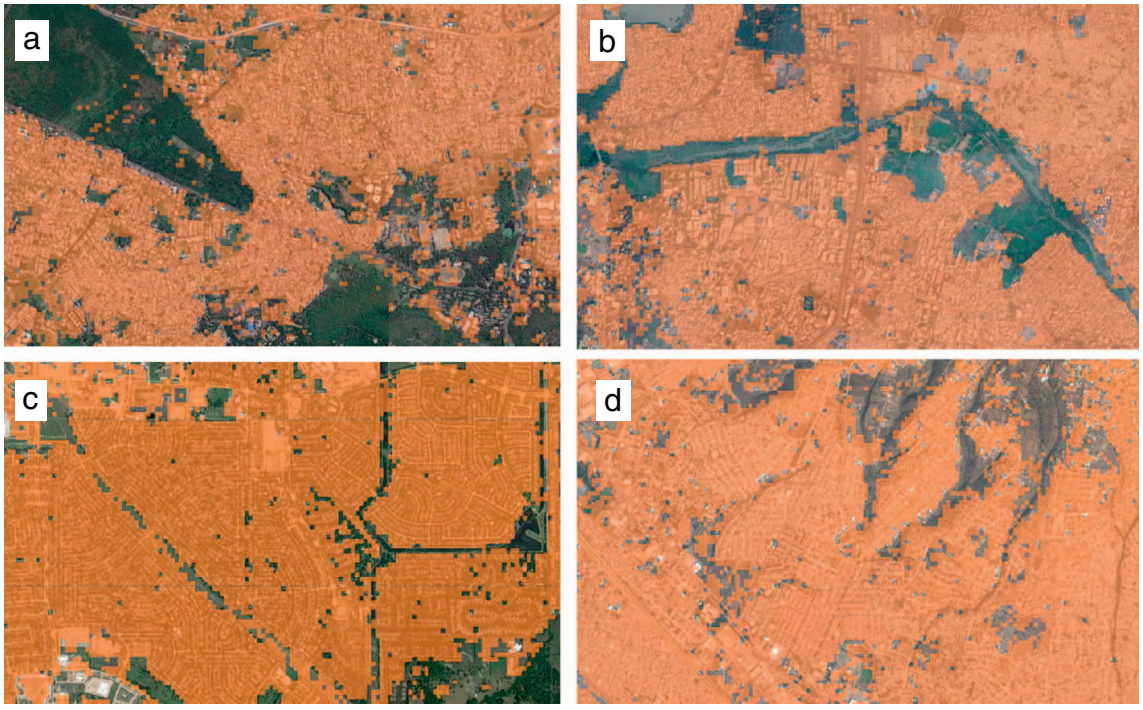

Fig. 13. Classification of built-up areas (in orange) in (a) Nagpur, India; (b) Hyderabad, India; (c) Houston US; (d) Mexico City, Mexico.

Table 6
Accuracy assessment (balanced accuracy) of our methodology for classification, MODIS-MCD12Q1 (MODIS), DMSP-OLS highly-lit areas (defined as pixels with a DN value above the 95th or 90th percentile), Accuracy assessment with 60,000 labeled examples.

	Our BU classification	MODIS UMD
India		
TPR (producer's accuracy)	59.9%	61.7%
TNR	91.2%	84.3%
Balanced accuracy	75.5%	73.0%
US		
TPR (producer's accuracy)	71.8%	64.4%
TNR	90.9%	87.3%
Balanced accuracy	81.4%	75.8%
Mexico		
TPR (producer's accuracy)	62.7%	53.4%
TNR	92.6%	94.1%
Balanced accuracy	77.6%	73.7%

(e.g., Hyderabad, India and Culiacán, Mexico) while other cities (e.g., Atlanta, Georgia, US), are less dense and the built-up density is more evenly distributed throughout the city. Future research could also build on our methodology to estimate sub-pixel fraction of built up land cover, for example, by adopting the FCM and PCM algorithms to derive relative and absolute measures of class membership within a pixel (Foody, 2002) or by applying Spectral Unmixing methods (e.g., Mitraka et al., 2016) that are especially useful for classification in complex urban/peri-urban settings (Poursanidis et al., 2015).

To summarize, we have developed a framework for pixel-based image classification of built-up areas that uses hexagonal tessellation to combine nighttime and daytime light data. The methodology yields high-resolution, high quality depictions of built-up areas across three highly diverse countries.

The study contributes to the existing literature in several aspects. First, previous studies that combine nighttime and daytime data have been limited either in their spatial generalization or in their spatial

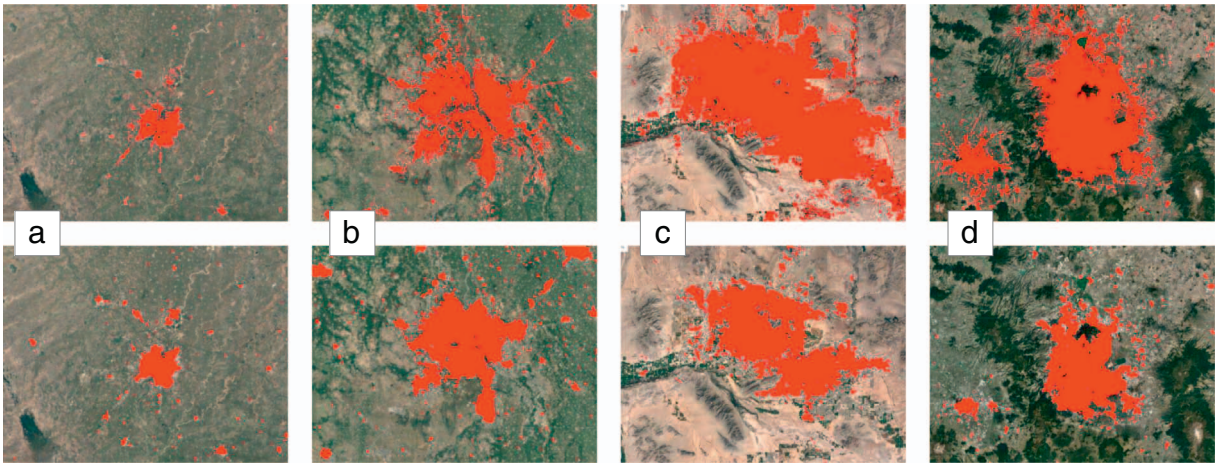

Fig. 14. A comparison between areas classified as built-up using our methodology (top) and areas classified as built up and urban by MCD12Q1 UMD MODIS classification scheme (bottom), in (a) Ahmedabad, Gujarat, India; (b) New Delhi, Delhi, India; (c) Phoenix, Arizona, US; (d) Mexico City, Mexico.

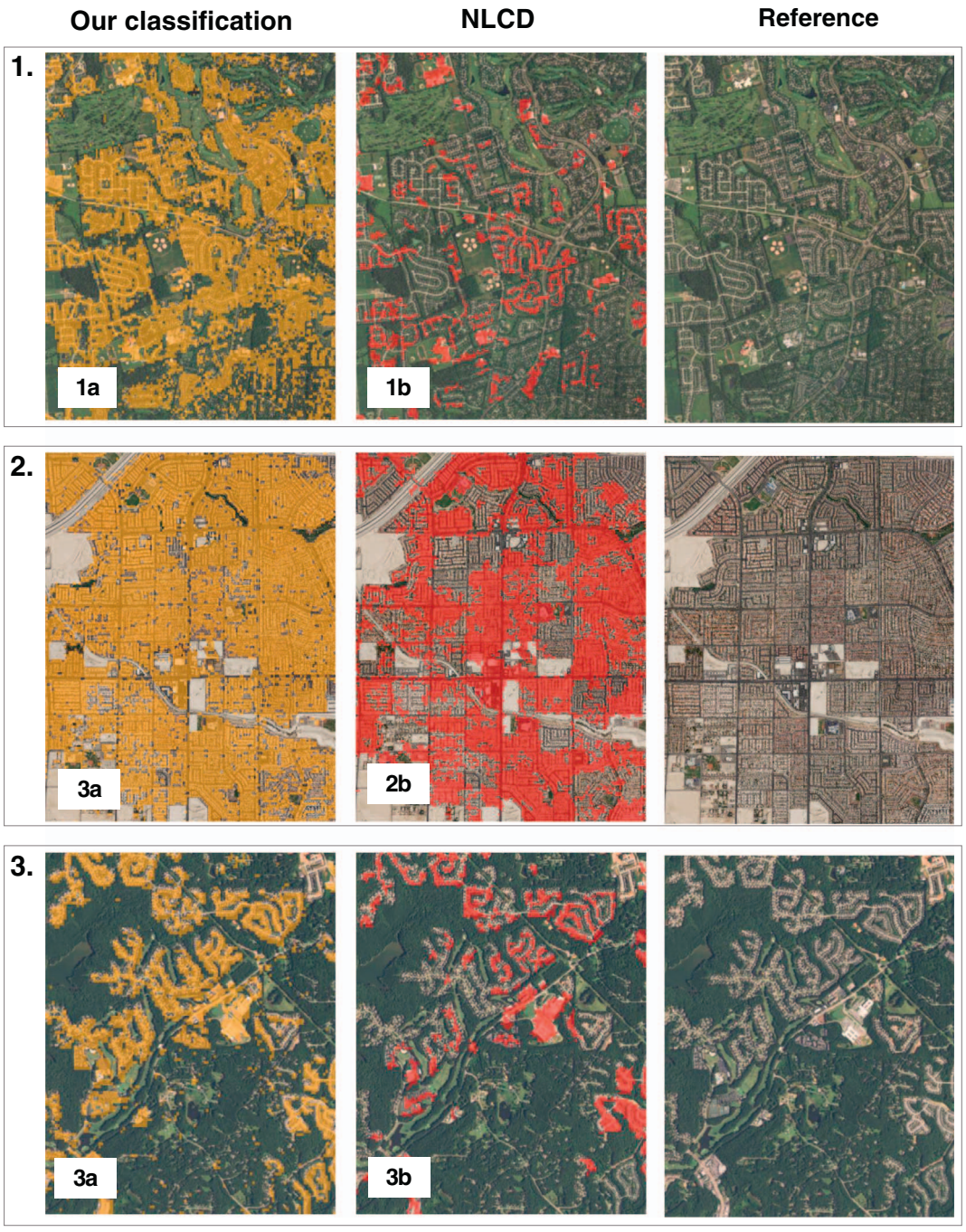

Fig. 15. A comparison of our classification (column a) and NLCD's classification* (column b), in (1) Columbus, Ohio; (2) Las Vegas, Nevada; (3) Atlanta, Georgia. *Classes 23–24: Developed, Medium and Intensity (Impervious surfaces account for 50% to 100% of the total cover).

resolution. To cope with regional variations, we adopt a hexagonal tessellation mapping approach to handle heterogeneous land cover with complex spectral characteristics. Second, despite significant progress in applying machine-learning techniques to process satellite data, there is a paucity of reference datasets to detect urban areas. The approach we present uses DMSP-OLS and Landsat data to collect training examples and Landstat data for the input to the classifier. Third, to determine the binary-class of a pixel (i.e., "built-up" or "not built-up") we use the Otsu algorithm to determine the threshold of the posterior probability for each hex-cell. Finally, our approach requires minimal manual input for training and for classification and does not rely on local adjustable parameters to the classifiers, making it scalable across space and time, which is the subject of ongoing work.

In today's era of big data, a globally consistent and data-driven method of defining and classifying urban areas has extensive applications. Economics, urban planning, climate modeling, water-resource management, hazard-response efforts, and urban-ecosystem assessments all use geographic data on urban areas. Having information on urban extent that is spatially and temporally consistent and defined at high resolution is both essential to a wide range of disciplines and necessary for helping society better understand the drivers of urbanization, therefore promoting sustainable urban development.

Acknowledgements

MG was supported by the National Science Foundation Grants EAR-

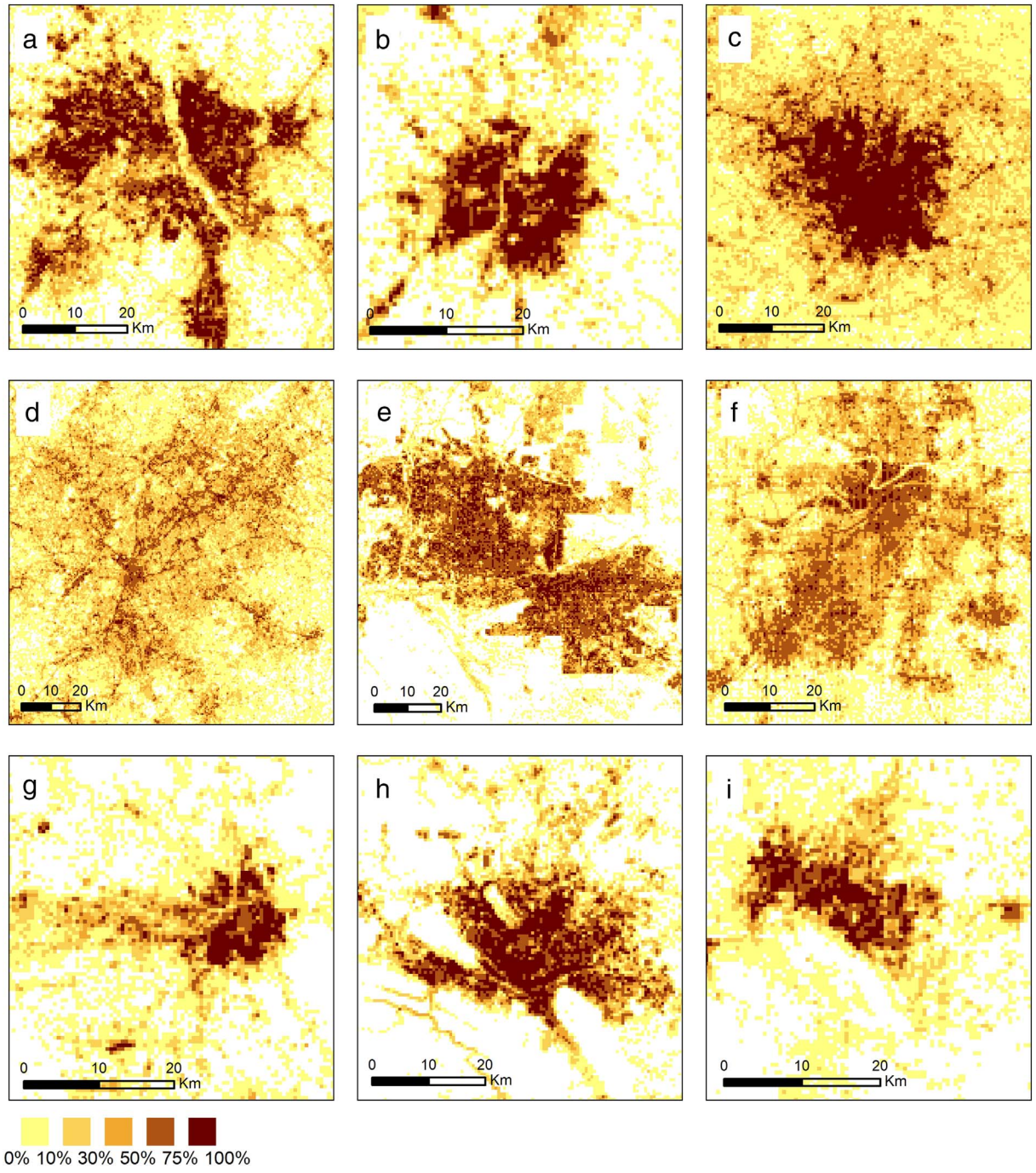

Fig. 16. percentage of built-up land cover in: India (a) New Delhi, (b) Ahmedabad, (c) Hyderabad; US (d) Atlanta, (e) Phoenix, (f) Kansas City; and Mexico (g) Culiacán, (h) Monterrey, (i) Gómez Palacio. The value of each 500 m*500 m pixel represent the percentage of built-up land cover (according to Landsat) that is incorporated in it.

1204774 and DMS-1419593, the U.S. Department of Agriculture NIFA grant 2015-67003-23508. MG and MS were also supported by the National Science Foundation Sustainability Research Network (SRN) Cooperative Agreement 1444758, the Urban Water Innovation Network (UWIN). MS was also supported by Dr. B. L. Turner II's Gilbert F. White Environment and Society Fellowship at Arizona State University. AK, RG and GH were supported by funds from the International Growth Centre (IGC project number 89448). RG and GH were supported by the Center on Global Transformation at UCSD. We thank the Institute of Geography, National Autonomous University of Mexico, Dr. Armando Peralta Higuera, who provided support for two students, Juan Alberto Guerra Moreno and Kimberly Mendez Gomez for validating the algorithm in Mexico.

Appendix A. Indices used as inputs to the classifiers

Table A1presents the bands and the indices used as inputs to the classifiers. We now turn to describe the indices in details.

I. NDVI (Normalized Difference Vegetation Index)

NDVI expresses the relation between red visible light (which is typically absorbed by a plant's chlorophyll) and near-infrared wavelength (which is scattered by the leaf's mesophyll structure) (Pettorelli et al., 2005). It is computed as:

$$(NIR - RED)/(NIR + RED) \quad (A1)$$

where NIR is the near infra-red wavelength and RED is the red wavelength. The values of NDVI range between (−1) and (+1). An average NDVI value was calculated for each pixel.

II. NDBI (Normalized Difference Built-up Index)

NDBI expresses the relation between the medium infra-red and the near infra-red wavelengths (Zha et al., 2003). It is computed as:

$$(MIR - NIR)/(MIR + NIR) \quad (A2)$$

where MIR is the medium infra-red and NIR is the near infra-red wavelength. The index assumes a higher reflectance of built-up areas in the medium infra-red wavelength range than in the near infra-red.

III. NDWI (Normalized Difference Water Index)

NDWI expresses the relation between the green (G) and the NIR (near infra-red) wavelength, with a scaling of −1 to +1 (McFeeters, 1996). It is computed as:

$$(G - NIR)/(G + NIR) \quad (A3)$$

The positive values typically represent open water areas, while negative values represent non-water features, like soil and vegetation.

IV. UI (Urban Index)

UI is the difference between the short infra-red (SWIR) and the near-infrared wavelengths (Kawamura et al., 1996). It is computed as:

$$(SWIR-NIR)/(SWIR+NIR) \quad (A4)$$

Like NDBI, UI assumes high brightness in SWIR in urban areas as opposed to the NIR.

V. EVI (Enhanced Vegetation Index)

EVI is an improved vegetation index with higher sensitivity in high biomass regions where NDVI tends to saturate, reduces atmospheric influences, and removes the canopy background signal (Huete et al., 2002). It is computed as:

$$(2.5*((NIR - R)/(NIR + 6R - 7.5B + 1)) \quad (A5)$$

This is a similar formula to the NDVI, which takes advantage of high reflectance of vegetation in the NIR band as opposed to the R band. The blue band (B) is used (with a coefficient of 7.5) to correct for aerosol influences in the red band (R). There is an integer of 1 added to the denominator to adjustment for nonlinear NIR and R radiant transfer through canopies. 2.5 is applied as a gain to the index.

VI. NDSV (Normalized Difference Spectral Vector)

NDSV (Normalized Difference Spectral Vector) (Trianni et al., 2015) is used to identify consistent human settlement and artificial materials. The index is composed out of all possible normalized indexes combining two bands with the same spatial resolution. Each pixel is characterized by a spectral vector of normalized indexes, which maps the original multispectral pixel into a higher dimensional space (Trianni et al., 2015). The index is calculated with each of Landsat's 30 m resolution bands as follows (Trianni et al., 2015):

$$NDSV_{ij} = \frac{B_i - B_j}{B_i + B_j}$$

where B_i and B_j are two generic bands.

Table A1
Bands in Landsat and indices used.

	Spectral band	Wavelength (micrometers)	Resolution (meters)
Landsat 8			
B1	Band 1 – Ultra blue	0.43–0.45	30
B2	Band 2 - Blue	0.45–0.51	30
B3	Band 3 - Green	0.53–0.59	30
B4	Band 4 – Red	0.64–0.67	30
B5	Band 5 - Near Infrared (NIR)	0.85–0.88	30
B6	Band 6 - SWIR 1	1.57–1.65	30
B7	Band 7 - SWIR 2	2.11–2.29	30
B8	Band 8 - Panchromatic	0.50–0.68	15
B10	Band 10 - Thermal Infrared (TIRS) 1	10.60–11.19	100 (resampled to 30)

B11	Band 11 - Thermal Infrared (TIRS) 2	11.50–12.51	100 (resampled to 30)
NDVI	(B5 − B4) / (B5 + B4)		30
NDWI	(B3 − B5) / (B3 + B5)		30
NDBI	(B6 − B5) / (B6 + B5)		30
EVI	2.5 ∗ ((B5 / B4) / (B5 + 6 ∗ B4 − 7.5 ∗ B2 + 1))		30
UI	(B7 − B5) / (B7 + B5)		30

Appendix B. Performance measures with fifty hex-cells

Table B1
Performance measures with fifty-hex-cells per country.

	NDVI percentile	Highly lit pixels	Inputs to classifier	Low-lit pixels	TPR	TNR	Balanced	User's accuracy
Optimal NDVI threshold								
India	50	99	All*	99	56.16%	93.23%	74.69%	71.89%
	75	99	All*	99	58.24%	92.25%	75.25%	69.86%
	90	99	All*	99	61.27%	91.18%	76.23%	67.33%
	95	99	All*	99	58.75%	91.27%	75.01%	67.49%
	99	99	All*	99	59.04%	90.67%	74.85%	66.13%
US	50	99	All*	99	65.84%	84.51%	75.18%	72.53%
	75	99	All*	99	72.42%	80.71%	76.57%	69.99%
	90	99	All*	99	76.14%	79.85%	77.99%	70.11%
	95	99	All*	99	76.76%	78.79%	77.78%	69.20%
	99	99	All*	99	76.22%	77.06%	76.64%	67.35%
Mexico	50	99	All*	99	69.94%	94.63%	82.28%	86.59%
	75	99	All*	99	72.96%	93.55%	83.26%	84.87%
	90	99	All*	99	72.40%	92.90%	82.65%	83.50%
	95	99	All*	99	72.22%	92.87%	82.55%	83.39%
	99	99	All*	99	72.61%	92.61%	82.61%	82.97%
Highly-lit pixels								
India	90	99	All*	99	58.89%	91.18%	75.04%	67.33%
	90	95	All*	99	72.71%	85.57%	79.14%	60.84%
	90	90	All*	99	78.98%	77.75%	78.36%	52.26%
	90	75	All*	99	87.62%	60.56%	74.09%	40.66%
	90	50	All*	99	91.00%	58.23%	74.61%	40.19%
US	90	99	All*	99	76.14%	79.85%	77.99%	70.11%
	90	95	All*	99	77.15%	78.55%	77.85%	69.07%
	90	90	All*	99	80.64%	76.43%	78.53%	67.99%
	90	75	All*	99	87.45%	64.26%	75.86%	60.31%
	90	50	All*	99	94.72%	37.42%	66.07%	49.81%
Mexico	75	99	All*	99	72.96%	93.55%	83.26%	84.87%
	75	95	All*	99	72.71%	92.66%	82.69%	83.09%
	75	90	All*	99	86.01%	82.58%	84.29%	71.00%
	75	75	All*	99	92.90%	58.65%	75.77%	52.69%
	75	50	All*	99	95.25%	39.42%	67.34%	43.81%
Inputs to classifier								
India	90	95	No Indices	99	71.13%	86.52%	78.83%	61.94%
	90	95	NDVI	99	71.85%	86.28%	79.06%	61.76%
	90	95	NDBI	99	70.34%	86.41%	78.37%	61.49%
	90	95	All*	99	72.71%	85.57%	79.14%	60.84%
	90	95	NDSVI	99	73.87%	86.05%	79.96%	62.03%
US	90	90	No Indices	99	80.25%	74.51%	77.38%	66.16%
	90	90	NDVI	99	78.93%	75.32%	77.13%	66.51%
	90	90	NDBI	99	80.09%	75.61%	77.85%	67.10%
	90	90	All*	99	80.64%	76.43%	78.53%	67.99%
	90	90	NDSVI	99	80.17%	76.48%	78.32%	67.91%
Mexico	90	90	No Indices	99	86.11%	82.44%	84.28%	70.86%
	90	90	NDVI	99	84.81%	82.72%	83.77%	70.88%
	90	90	NDBI	99	85.48%	82.48%	83.98%	70.75%
	90	90	All*	99	86.01%	82.58%	84.29%	71.00%
	90	90	NDSVI	99	85.41%	82.69%	84.05%	70.98%
Low-lit pixels								

India	90	95	NDSV	10	84.16%	69.04%	76.60%	45.61%
	90	95	NDSV	25	82.72%	74.82%	78.77%	50.33%
	90	95	NDSV	50	79.91%	80.70%	80.31%	56.09%
	90	95	NDSV	75	77.54%	84.12%	80.83%	60.10%
US	90	90	All*	10	93.57%	44.49%	69.03%	51.14%
	90	90	All*	25	91.87%	56.18%	74.02%	56.56%
	90	90	All*	50	88.30%	67.48%	77.89%	62.78%
	90	90	All*	75	82.49%	75.08%	78.79%	67.28%
Mexico	90	90	All*	10	93.28%	59.50%	76.39%	53.32%
	90	90	All*	25	92.65%	64.07%	78.36%	56.11%
	90	90	All*	50	91.42%	71.64%	81.53%	61.51%
	90	90	All*	75	88.75%	78.87%	83.81%	67.56%

References

Aguilar, A.G., 2008. Peri-urbanization, illegal settlements and environmental impact in Mexico City. Cities 25, 133–145. http://dx.doi.org/10.1016/j.cities.2008.02.003.

Aguilar, A.G., Guerrero, F.L., 2013. Poverty in peripheral informal settlements in Mexico City: the case of Magdalena Contreras, Federal District. Tijdschr. Voor Econ. En Soc. Geogr. 104, 359–378. http://dx.doi.org/10.1111/tesg.12012.

Angiuli, E., Trianni, G., 2014. Urban mapping in Landsat images based on normalized difference spectral vector. IEEE Geosci. Remote Sens. Lett. 11, 661–665. http://dx.doi.org/10.1109/LGRS.2013.2274327.

As-Syakur, A.R., Adnyana, I., Arthana, I.W., Nuarsa, I.W., 2012. Enhanced built-up and bareness index (EBBI) for mapping built-up and bare land in an urban area. Remote Sens. 4 (10), 2957–2970.

Bagan, H., Yamagata, Y., 2014. Land-cover change analysis in 50 global cities by using a combination of Landsat data and analysis of grid cells. Environ. Res. Lett. 9, 064015. http://dx.doi.org/10.1088/1748-9326/9/6/064015.

Bagan, H., Yamagata, Y., 2015. Analysis of urban growth and estimating population density using satellite images of nighttime lights and land-use and population data. GISci. Remote Sens. 52, 765–780. http://dx.doi.org/10.1080/15481603.2015.1072400.

Ban, Y., Jacob, A., Gamba, P., 2015. Spaceborne SAR data for global urban mapping at 30m resolution using a robust urban extractor. ISPRS J. Photogramm. Remote Sens. 103, 28–37. http://dx.doi.org/10.1016/j.isprsjprs.2014.08.004.

Belgiu, M., Drăguţ, L., 2014. Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. ISPRS J. Photogramm. Remote Sens. 96, 67–75. http://dx.doi.org/10.1016/j.isprsjprs.2014.07.002.

Birch, C.P.D., Oom, S.P., Beecham, J.A., 2007. Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. Ecol. Model. 206, 347–359. http://dx.doi.org/10.1016/j.ecolmodel.2007.03.041.

Boucher, A., Seto, K., 2009. Methods and challenges for using high-temporal resolution data to monitor urban growth. In: Global Mapping of Human Settlement, Remote Sensing Applications Series. CRC Press. http://dx.doi.org/10.1201/9781420083408-c16.

Chen, X.L., Zhao, H.M., Li, P.X., Yin, Z.Y., 2006. Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. Remote Sens. Environ. 104, 133–146. http://dx.doi.org/10.1016/j.rse.2005.11.016.

Chen, C.-F., Son, N.-T., Chang, N.-B., Chen, C.-R., Chang, L.-Y., Valdez, M., Centeno, G., Thompson, C.A., Aceituno, J.L., 2013. Multi-decadal mangrove forest change detection and prediction in Honduras, Central America, with Landsat imagery and a Markov chain model. Remote Sens. 5, 6408–6426. http://dx.doi.org/10.3390/rs5126408.

Chen, Jun, Chen, Jin, Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., Mills, J., 2015. Global land cover mapping at 30 m resolution: a POK-based operational approach. ISPRS J. Photogramm. Remote Sens. 103, 7–27. http://dx.doi.org/10.1016/j.isprsjprs.2014.09.002.

CIESIN, 2005. Gridded Population of the World, Version 3 (GPWv3) Data Collection.

Clinton, N., Gong, P., 2013. MODIS detected surface urban heat islands and sinks: global locations and controls. Remote Sens. Environ. 134, 294–304. http://dx.doi.org/10.1016/j.rse.2013.03.008.

Connolly, P., 2014. Latin Amercian Informal Urbanism: Contexts, Concepts, and Contributions with Specific Reference to Mexico. In: Becerra, A., Hernandez, F. (Eds.), (Re)viewing Latin American Cities. Cambridge Scholars Publishing, Cambridge.

Consejo Nacional de Población, 2012. Sistema Urbano Nacional. Col. Juárez, Del. Cuauhtémoc, Mexico.

Eakin, H., Lerner, A.M., Manuel-Navarrete, D., Hernández Aguilar, B., Martínez-Canedo, A., Tellman, B., Charli-Joseph, L., Fernández Álvarez, R., Bojórquez-Tapia, L., 2016. Adapting to risk and perpetuating poverty: Household's strategies for managing flood risk and water scarcity in Mexico City. Environ. Sci. Pol. 66, 324–333. http://dx.doi.org/10.1016/j.envsci.2016.06.006.

Elvidge, C.D., Ziskin, D., Baugh, K.E., Tuttle, B.T., Ghosh, T., Pack, D.W., Erwin, E.H., Zhizhin, M., 2009. A fifteen year record of global natural gas flaring derived from satellite data. Energies 2, 595–622. http://dx.doi.org/10.3390/en20300595.

Elvidge, C., Hsu, F.-C., Baugh, K., Ghosh, T., 2014. National trends in satellite-observed lighting: 1992–2012. In: Global Urban Monitoring and Assessment through Earth Observation, Remote Sensing Applications Series. CRC Press, pp. 97–120.

Estima, J., Painho, M., 2015. Investigating the Potential of OpenStreetMap for Land Use/Land Cover Production: A Case Study for Continental Portugal. In: Arsanjani, J.J., Zipf, A., Mooney, P., Helbich, M. (Eds.), OpenStreetMap in GIScience, Lecture Notes in Geoinformation and Cartography. Springer International Publishing, pp. 273–293.

Estoque, R.C., Murayama, Y., 2015. Classification and change detection of built-up lands from Landsat-7 ETM + and Landsat-8 OLI/TIRS imageries: a comparative assessment of various spectral indices. Ecol. Indic. 56, 205–217. http://dx.doi.org/10.1016/j.ecolind.2015.03.037.

Flood, N., 2013. Seasonal composite Landsat TM/ETM + images using the Medoid (a multi-dimensional median). Remote Sens. 5, 6481–6500. http://dx.doi.org/10.3390/rs5126481.

Foody, G.M., 2002. Status of land cover classification accuracy assessment. Remote Sens. Environ. 80, 185–201. http://dx.doi.org/10.1016/S0034-4257(01)00295-4.

Gamba, P., Aldrighi, M., Stasolla, M., 2011. Robust extraction of urban area extents in HR and VHR SAR images. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 4, 27–34. http://dx.doi.org/10.1109/JSTARS.2010.2052023.

Gao, B.C., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58 (3), 257–266.

Gaughan, A.E., Stevens, F.R., Linard, C., Jia, P., Tatem, A.J., 2013. High resolution population distribution maps for Southeast Asia in 2010 and 2015. PLoS One 8, e55882. http://dx.doi.org/10.1371/journal.pone.0055882.

Giri, C., Pengra, B., Long, J., Loveland, T.R., 2013. Next generation of global land cover characterization, mapping, and monitoring. Int. J. Appl. Earth Obs. Geoinf. 25, 30–37. http://dx.doi.org/10.1016/j.jag.2013.03.005.

Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random forests for land cover classification. Pattern Recogn. Lett. 27, 294–300. http://dx.doi.org/10.1016/j.patrec.2005.08.011.

Goldblatt, R., You, W., Hanson, G., Khandelwal, K.A., 2016. Detecting the boundaries of urban areas in India: a dataset for pixel-based image classification in Google earth engine. Remote Sens. 8. http://dx.doi.org/10.3390/rs8080634.

Gong, P., Wang, J., Yu, Le, Zhao, Yongchao, Zhao, Yuanyuan, Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, Han, Zhu, P., Zhao, Z., Zhang, Haiying, Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, Liang, Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, Jin, Chen, Jun, 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. http://dx.doi.org/10.1080/01431161.2012.748992.

Guan, H., Li, J., Chapman, M., Deng, F., Ji, Z., Yang, X., 2013. Integration of orthoimagery and lidar data for object-based urban thematic mapping using random forests. Int. J. Remote Sens. 34, 5166–5186. http://dx.doi.org/10.1080/01431161.2013.788261.

He, C., Shi, P., Xie, D., Zhao, Y., 2010. Improving the normalized difference built-up index to map urban built-up areas using a semiautomatic segmentation approach. IEEE Geosci. Remote Sens. Lett. 1, 213–221.

Henderson, M., Yeh, E.T., Gong, P., Elvidge, C., Baugh, K., 2003. Validation of urban boundaries derived from global night-time satellite imagery. Int. J. Remote Sens. 24, 595–609. http://dx.doi.org/10.1080/01431160304982.

Herold, M., 2009. Some recommendations for global efforts in urban monitoring and assessments from remote sensing. In: Global Mapping of Human Settlement, Remote Sensing Applications Series. CRC Press. http://dx.doi.org/10.1201/9781420083408-c2.

Homer, C., Gallant, A., 2001. Partitioning the Conterminous United States into Mapping Zones for Landsat TM Land Cover Mapping. USGS.

Homer, C., Huang, C., Yang, L., Wylie, B.K., Coan, M., 2004. Development of a 2001 National Land-Cover Database for the United States. Photogramm. Eng. Remote. Sens. 7, 829–840.

Hsu, F.-C., Baugh, K.E., Ghosh, T., Zhizhin, M., Elvidge, C.D., 2015. DMSP-OLS radiance calibrated nighttime lights time series with intercalibration. Remote Sens. 7, 1855–1876. http://dx.doi.org/10.3390/rs70201855.

Huang, X., Schneider, A., Friedl, M.A., 2016. Mapping sub-pixel urban expansion in China using MODIS and DMSP/OLS nighttime lights. Remote Sens. Environ. 175, 92–108. http://dx.doi.org/10.1016/j.rse.2015.12.042.

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. http://dx.doi.org/10.1016/S0034-4257(02)00096-2.

Hunsaker, C.T., O'Neill, R.V., Jackson, B.L., Timmins, S.P., Levine, D.A., Norton, D.J., 1994. Sampling to characterize landscape pattern. Landsc. Ecol. 9, 207–226. http://dx.doi.org/10.1007/BF00134748.
Imhoff, M., Lawrence, W.T., Stutzer, D.C., Elvidge, C.D., 1997. A technique for using composite DMSP/OLS "City Lights" satellite data to map urban area. Remote Sens. Environ. 61, 361–370. http://dx.doi.org/10.1016/S0034-4257(97)00046-1.
Jacob, A., Ban, Y., 2015. Sentinel-1A SAR Data for Global Urban Mapping: Preliminary Results. 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Presented at the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)pp. 1179–1182. http://dx.doi.org/10.1109/IGARSS.2015.7325982.
Jean, N., Burke, M., Xie, M., Davis, W.M., Lobell, D.B., Ermon, S., 2016. Combining satellite imagery and machine learning to predict poverty. Science 353, 790–794. http://dx.doi.org/10.1126/science.aaf7894.
Jing, W., Yang, Y., Yue, X., Zhao, X., 2015. Mapping urban areas with integration of DMSP/OLS nighttime light and MODIS data using machine learning techniques. Remote Sens. 7, 12419–12439. http://dx.doi.org/10.3390/rs70912419.
Johnson, B.A., Iizuka, K., 2016. Integrating OpenStreetMap crowdsourced data and Landsat time-series imagery for rapid land use/land cover (LULC) mapping: case study of the Laguna de Bay area of the Philippines. Appl. Geogr. 67, 140–149. http://dx.doi.org/10.1016/j.apgeog.2015.12.006.
Kasimu, A., Tateishi, R., Hoan, N.T., 2009. Global urban characterization using population density, DMSP and MODIS data. In: 2009 Joint Urban Remote Sensing Event. Presented at the 2009 Joint Urban Remote Sensing Event, pp. 1–7. http://dx.doi.org/10.1109/URS.2009.5137493.
Kawamura, M., Jayamana, S., Tsujiko, Y., 1996. Relation between social and environmental conditions in Colombo Sri Lanka and the urban index estimated by satellite remote sensing data. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. 31, 321–326.
Keola, S., Andersson, M., Hall, O., 2015. Monitoring economic development from space: using nighttime light and land cover data to measure economic growth. World Dev. 66, 322–334. http://dx.doi.org/10.1016/j.worlddev.2014.08.017.
Levin, N., Duke, Y., 2012. High spatial resolution night-time light images for demographic and socio-economic studies. Remote Sens. Environ. 119, 1–10. http://dx.doi.org/10.1016/j.rse.2011.12.005.
Li, C., Li, J., Wu, J., 2013. Quantifying the speed, growth modes, and landscape pattern changes of urbanization: a hierarchical patch dynamics approach. Landsc. Ecol. 28, 1875–1888. http://dx.doi.org/10.1007/s10980-013-9933-6.
Liu, H., Zhou, Q., 2004. Accuracy analysis of remote sensing change detection by rule-based rationality evaluation with post-classification comparison. Int. J. Remote Sens. 25, 1037–1050. http://dx.doi.org/10.1080/0143116031000150004.
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. http://dx.doi.org/10.1016/j.landurbplan.2012.02.013.
Liu, Y., Delahunty, T., Zhao, N., Cao, G., 2016. These lit areas are undeveloped: delimiting China's urban extents from thresholded nighttime light imagery. Int. J. Appl. Earth Obs. Geoinf. 50, 39–50. http://dx.doi.org/10.1016/j.jag.2016.02.011.
Lu, D., Tian, H., Zhou, G., Ge, H., 2008. Regional mapping of human settlements in southeastern China with multisensor remotely sensed data. Remote Sens. Environ. 112, 3668–3679. http://dx.doi.org/10.1016/j.rse.2008.05.009.
Ma, L., Wu, J., Li, W., Peng, J., Liu, H., 2014. Evaluating saturation correction methods for DMSP/OLS nighttime light data: a case study from China's cities. Remote Sens. 6, 9853–9872. http://dx.doi.org/10.3390/rs6109853.
MacLachlan, A., Roberts, G., Biggs, E., Boruff, B., 2017. Subpixel land-cover classification for improved urban area estimates using Landsat. Int. J. Remote Sens. 38, 5763–5792. http://dx.doi.org/10.1080/01431161.2017.1346403.
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.
Mertes, C.M., Schneider, A., Sulla-Menashe, D., Tatem, A.J., Tan, B., 2015. Detecting change in urban areas at continental scales with MODIS data. Remote Sens. Environ. 158, 331–347. http://dx.doi.org/10.1016/j.rse.2014.09.023.
Mitraka, Z., Frate, F.D., Carbone, F., 2016. Nonlinear spectral unmixing of Landsat imagery for urban surface cover mapping. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9, 3340–3350. http://dx.doi.org/10.1109/JSTARS.2016.2522181.
Miyazaki, H., Iwao, K., Shibasaki, R., 2011. Development of a new ground truth database for global urban area mapping from a gazetteer. Remote Sens. 3, 1177–1187. http://dx.doi.org/10.3390/rs3061177.
Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., Weng, Q., 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sens. Environ. 115, 1145–1161. http://dx.doi.org/10.1016/j.rse.2010.12.017.
Ng, H.-F., 2006. Automatic thresholding for defect detection. Pattern Recogn. Lett. 27, 1644–1649. http://dx.doi.org/10.1016/j.patrec.2006.03.009.
O'Neill, R.V., Hunsaker, C.T., Timmins, S.P., Jackson, B.L., Jones, K.B., Riitters, K.H., Wickham, J.D., 1996. Scale problems in reporting landscape pattern at the regional scale. Landsc. Ecol. 11, 169–180. http://dx.doi.org/10.1007/BF02447515.
Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. 9, 62–66. http://dx.doi.org/10.1109/TSMC.1979.4310076.
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 (Part B), 199–208. http://dx.doi.org/10.1016/j.jag.2014.09.005.
Pekel, J.-F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. Nature 540, 418–422. http://dx.doi.org/10.1038/nature20584.
Pesaresi, M., Ehrlich, D., Ferri, S., Florczyk, A., Carneiro Freire, S.M., Halkia, S., Julea, A.M., Kemper, T., Soille, P., Syrris, V., 2016. Operating Procedure for the Production of the Global Human Settlement Layer From Landsat Data of the Epochs 1975, 1990, 2000, and 2014. Publications Office of the European Union, Ispra (VA), Italy.
Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.M., Tucker, C.J., Stenseth, N.C., 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends Ecol. Evol. 20 (9), 503–510.
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. http://dx.doi.org/10.1080/01431160903121134.
Poursanidis, D., Chrysoulakis, N., Mitraka, Z., 2015. Landsat 8 vs. Landsat 5: a comparison based on urban and peri-urban land cover mapping. Int. J. Appl. Earth Obs. Geoinf. 35, 259–269. http://dx.doi.org/10.1016/j.jag.2014.09.010.
Richards, T., Gallego, J., Achard, F., 2000. Sampling for forest cover change assessment at the pan-tropical scale. Int. J. Remote Sens. 21, 1473–1490. http://dx.doi.org/10.1080/014311600210272.
Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J. Photogramm. Remote Sens. 67, 93–104. http://dx.doi.org/10.1016/j.isprsjprs.2011.11.002.
Rwanga, S.S., Ndambuki, J.M., 2017. Accuracy assessment of land use/land cover classification using remote sensing and GIS. Int. J. Geosci. 8, 611.
Schlesinger, J., 2015. Using Crowd-Sourced Data to Quantify the Complex Urban Fabric - OpenStreetMap and the Urban–Rural Index. In: Arsanjani, J., Zipf, A., Mooney, P., Helbich, M. (Eds.), OpenStreetMap in GIScience. Springer International Publishing, Switzerland, pp. 295–315.
Schneider, A., Friedl, M.A., Potere, D., 2009. A new map of global urban extent from MODIS satellite data. Environ. Res. Lett. 4, 044003. http://dx.doi.org/10.1088/1748-9326/4/4/044003.
Schneider, A., Friedl, M.A., Potere, D., 2010. Mapping global urban areas using MODIS 500-m data: new methods and datasets based on "urban ecoregions". Remote Sens. Environ. 114, 1733–1746. http://dx.doi.org/10.1016/j.rse.2010.03.003.
Seto, K.C., Fragkias, M., Güneralp, B., Reilly, M.K., 2011. A meta-analysis of global urban land expansion. PLoS One 6, e23777. http://dx.doi.org/10.1371/journal.pone.0023777.
Sirmacek, B., Unsalan, C., 2010. Urban area detection using local feature points and spatial voting. IEEE Geosci. Remote Sens. Lett. 7, 146–150. http://dx.doi.org/10.1109/LGRS.2009.2028744.
Small, C., Elvidge, C.D., 2013. Night on earth: mapping decadal changes of anthropogenic night light in Asia. Int. J. Appl. Earth Obs. Geoinf. 22, 40–52. http://dx.doi.org/10.1016/j.jag.2012.02.009.
Small, C., Pozzi, F., Elvidge, C.D., 2005. Spatial analysis of global urban extent from DMSP-OLS night lights. Remote Sens. Environ. 96, 277–291. http://dx.doi.org/10.1016/j.rse.2005.02.002.
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. http://dx.doi.org/10.1016/j.rse.2015.12.027.
Stevens, F.R., Gaughan, A.E., Linard, C., Tatem, A.J., 2015. Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. PLoS One 10, e0107042. http://dx.doi.org/10.1371/journal.pone.0107042.
Su, Y., Chen, X., Wang, Chongyang, Zhang, H., Liao, J., Ye, Y., Wang, Changjian, 2015. A new method for extracting built-up urban areas using DMSP-OLS nighttime stable lights: a case study in the Pearl River Delta, southern China. GISci. Remote Sens. 52, 218–238. http://dx.doi.org/10.1080/15481603.2015.1007778.
Sutton, P.C., 2003. A scale-adjusted measure of "urban sprawl" using nighttime satellite imagery. Remote Sens. Environ. 86, 353–369. http://dx.doi.org/10.1016/S0034-4257(03)00078-6.
Tatem, A.J., Noor, A.M., von Hagen, C., Gregorio, A.D., Hay, S.I., 2007. High resolution population maps for low income nations: combining land cover and census in East Africa. PLoS One 2, e1298. http://dx.doi.org/10.1371/journal.pone.0001298.
Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., Dech, S., 2012. Monitoring urbanization in mega cities from space. Remote Sens. Environ. 117, 162–176. http://dx.doi.org/10.1016/j.rse.2011.09.015.
Theobald, D.M., 2016. A general-purpose spatial survey design for collaborative science and monitoring of global environmental change: the global grid. Remote Sens. 8, 813. http://dx.doi.org/10.3390/rs8100813.
Trianni, G., Lisini, G., Angiuli, E., Moreno, E.A., Dondi, P., Gaggia, A., Gamba, P., 2015. Scaling up to national/regional urban extent mapping using Landsat data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8, 3710–3719. http://dx.doi.org/10.1109/JSTARS.2015.2398032.
Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8, 127–150. http://dx.doi.org/10.1016/0034-4257(79)90013-0.
Turner, M.G., 1989. Landscape ecology: the effect of pattern on process. Annu. Rev. Ecol. Syst. 20, 171–197.
UN, 2014. World Urbanization Prospects: The 2014 Revision. https://esa.un.org/unpd/wup/CD-ROM/, Accessed date: 17 February 2001.
US Census Bureau, 2012. Increasing Urbanization: Population Distribution by City Size, 1790 to 1890. https://www.census.gov/dataviz/visualizations/005/, Accessed date: 17 February 2001.
US Census Bureau, 2016. Urban and Rural Areas. https://www.census.gov/history/www/programs/geography/urban_and_rural_areas.html, Accessed date: 17 February 2001.
Van de Voorde, T., Jacquet, W., Canters, F., 2011. Mapping form and function in urban areas: an approach based on urban metrics and continuous impervious surface data. Landsc. Urban Plan. 102, 143–155. http://dx.doi.org/10.1016/j.landurbplan.2011.

03.017.
Wan, B., Guo, Q., Fang, F., Su, Y., Wang, R., 2015. Mapping US urban extents from MODIS data using one-class classification method. Remote Sens. 7, 10143–10163. http://dx.doi.org/10.3390/rs70810143.
Wei, Y., Liu, H., Song, W., Yu, B., Xiu, C., 2014. Normalization of time series DMSP-OLS nighttime light images for urban growth analysis with pseudo invariant features. Landsc. Urban Plan. 128, 1–13. http://dx.doi.org/10.1016/j.landurbplan.2014.04.015.
Wood, S., Har-Noy, S., 2013. A Better World with Enhanced Analytics Capabilities. 16. GeoInformatics Emmeloord, pp. 6–8.
Woodcock, C.E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S.N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P.S., Vermote, E.F., Vogelmann, J., Wulder, M.A., Wynne, R., 2008. Free access to Landsat imagery. Science 320, 1011. http://dx.doi.org/10.1126/science.320.5879.1011a.
Xiao, P., Wang, X., Feng, X., Zhang, X., Yang, Y., 2014. Detecting China's urban expansion over the past three decades using nighttime light data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7, 4095–4106. http://dx.doi.org/10.1109/JSTARS.2014.2302855.
Xu, H., 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. Int. J. Remote Sens. 27, 3025–3033. http://dx.doi.org/10.1080/01431160600589179.
Xu, H., 2008. A new index for delineating built-up land features in satellite imagery. Int. J. Remote Sens. 29, 4269–4276. http://dx.doi.org/10.1080/01431160802039957.
Yuan, F., Bauer, M.E., 2007. Comparison of impervious surface area and normalized difference vegetation index as indicators of surface urban heat island effects in Landsat imagery. Remote Sens. Environ. 106, 375–386. http://dx.doi.org/10.1016/j.rse.2006.09.003.
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. http://dx.doi.org/10.1080/01431160304987.
Zhang, Q., Seto, K.C., 2011. Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. Remote Sens. Environ. 115, 2320–2329. http://dx.doi.org/10.1016/j.rse.2011.04.032.
Zhang, Q., Seto, K.C., 2013. Can night-time light data identify typologies of urbanization? A global assessment of successes and failures. Remote Sens. 5, 3476–3494. http://dx.doi.org/10.3390/rs5073476.
Zhang, H., Zhang, Y., Lin, H., 2012. Urban land cover mapping using random forest combined with optical and SAR data. In: 2012 IEEE International Geoscience and Remote Sensing Symposium. Presented at the 2012 IEEE International Geoscience and Remote Sensing Symposium, pp. 6809–6812. http://dx.doi.org/10.1109/IGARSS.2012.6352600.
Zhang, Q., Schaaf, C., Seto, K.C., 2013. The vegetation adjusted NTL urban index: a new approach to reduce saturation and increase variation in nighttime luminosity. Remote Sens. Environ. 129, 32–41. http://dx.doi.org/10.1016/j.rse.2012.10.022.
Zhang, Q., Li, B., Thau, D., Moore, R., 2015. Building a better urban picture: combining day and night remote sensing imagery. Remote Sens. 7, 11887–11913. http://dx.doi.org/10.3390/rs70911887.
Zhao, S., Liu, S., Zhou, D., 2016. Prevalent vegetation growth enhancement in urban environment. Proc. Natl. Acad. Sci. 113, 6313–6318. http://dx.doi.org/10.1073/pnas.1602312113.
Zhou, Y., Smith, S.J., Elvidge, C.D., Zhao, K., Thomson, A., Imhoff, M., 2014. A cluster-based method to map urban area from DMSP/OLS nightlights. Remote Sens. Environ. 147, 173–185. http://dx.doi.org/10.1016/j.rse.2014.03.004.
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. http://dx.doi.org/10.1088/1748-9326/10/5/054011.