Geocarto International

ISSN: 1010-6049 (Print) 1752-0762 (Online) Journal homepage: http://www.tandfonline.com/loi/tgei20

Step-wise Land-class Elimination Approach for extracting mixed-type built-up areas of Kolkata megacity

Ansar Khan, Soumendu Chatterjee, Hashem Akbari, Saad Saleem Bhatti, Apurba Dinda, Chandana Mitra, Haoyuan Hong & Quang Van Doan

To cite this article: Ansar Khan, Soumendu Chatterjee, Hashem Akbari, Saad Saleem Bhatti, Apurba Dinda, Chandana Mitra, Haoyuan Hong & Quang Van Doan (2017): Step-wise Land-class Elimination Approach for extracting mixed-type built-up areas of Kolkata megacity , Geocarto International, DOI: 10.1080/10106049.2017.1408704

To link to this article: https://doi.org/10.1080/10106049.2017.1408704

Accepted author version posted online: 23 Nov 2017.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=tgei20

Download by: [University of Florida] **Date:** 23 November 2017, At: 07:22

Publisher: Taylor & Francis
Journal: *Geocarto International*
DOI: http://doi.org/10.1080/10106049.2017.1408704

Check for updates

Step-wise Land-class Elimination Approach for extracting mixed-type built-up areas of Kolkata megacity

Ansar Khan[a,*], Soumendu Chatterjee[b], Hashem Akbari[c], Saad Saleem Bhatti[d], Apurba Dinda[e], Chandana Mitra[f], Haoyuan Hong[g], Quang Van Doan[h]

Ansar Khan
Assistant Professor
Department of Geography, Lalbaba College, Howrah, India
Email: khanansargeo@gmail.com

Soumendu Chatterjee
Professor and Head
Department of Geography, Presidency University, Kolkata, India
Email: soumendu.geog@presiuniv.ac.in

Hashem Akbari
Professor
Heat Island Group, Building, Civil and Environmental Engineering, Concordia University, Montreal, Canada
Email: hakbari@encs.concordia.ca

Saad Saleem Bhatti
Research Associate
Department of Geography and Planning, School of Environmental Sciences, University of Liverpool, Liverpool, United Kingdom
Email: dearsaad@gmail.com

Apurba Dinda
Research Assistant
Department of Geography and Environment Management, Vidyasagar University, Midnapore, India
Email: apurbadinda0@gmail.com

Chandana Mitra
Associate Professor
Department of Geosciences, Auburn University, Auburn, Alabama
Email: czm0033@auburn.edu

Haoyuan Hong
Research Scientist
Jiangxi Provincial Meteorological Observatory, Jiangxi Meteorological Bureau, Nanchang, China
Email: honghaoyuan@qq.com

ACCEPTED MANUSCRIPT

Downloaded by [University of Florida] at 07:22 23 November 2017

Quang Van Doan
Research Scientist
Center for Computational Sciences, University of Tsukuba, Tsukuba, Japan
Email: doanquangvan@gmail.com

*Corresponding author:
Department of Geography, Lalbaba College, Howrah, India
Email: khanansargeo@gmail.com

Step-wise Land-class Elimination Approach for extracting mixed-type built-up areas of Kolkata megacity

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

Abstract

The extraction of urban built-up areas is an important aspect of urban planning and understanding the complex drivers and biophysical mechanism of urban climate processes. However, built-up area extraction using Landsat data is a challenging task due to spatio-temporal dynamics and spatially intermixed nature of Land Use and Land Cover (LULC) in the cities of the developing countries, particularly in tropics. In the light of advantages and drawbacks of the Normalized Difference Built-up Index (NDBI) and Built-up Area Extraction Method (BAEM), a new and simple method i.e. Step-wise Land-class Elimination Approach (SLEA) is proposed for rapid and accurate mapping of urban built-up areas without depending exclusively on the band specific normalized indices, in order to pursue a more generalized approach. It combines the use of a single band layer, Normalized Difference Vegetation Index (NDVI) image and another binary image obtained through Logit model. Based on the spectral designation of the satellite image in use, a particular band is chosen for identification of water pixels. The Double-window Flexible Pace Search (DFPS) approach is employed for finding the optimum threshold value that segments the selected band image into water and non-water categories- the water pixels are then eliminated from the original image.

The vegetation pixels are similarly identified using the NDVI image and eliminated. The residual pixels left after elimination of water and vegetation categories belong either to the built-up areas or to bare land categories. Logit model is used for separation of the built-up areas from bare lands. The effectiveness of this method was tested through the mapping of built-up areas of the Kolkata Metropolitan Area (KMA)**,** India from Thematic Mapper (TM) images of 2000, 2005 and 2010, and Operational Land Imager (OLI) image of 2015. Results of the proposed SLEA were 95.33% accurate on the whole, while those derived by the NDBI and BAEM approaches returned an overall accuracy of 83.67% and 89.33%, respectively. Comparisons of the results obtained using this method with those obtained from NDBI and BAEM approaches demonstrate that the proposed approach is quite reliable. The SLEA generates new patterns of evidence and hypotheses for built-up areas extraction research, providing an integral link with statistical science and encouraging trans-disciplinary collaborations to build robust knowledge and problem solving capacity in urban areas. It is also brings landscape architecture, urban and regional planning, landscape and ecological engineering, and other practice-oriented fields to bear in processes for identifying problems and analyzing, synthesizing, and evaluating desirable alternatives for urban change. This method produced very accurate results in a more efficient manner compared to the earlier built-up area extraction approaches for the landscape and urban planning.

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

Keywords:

Double-window flexible pace search; Kolkata metropolitan area; Logit model; Spatially intermixed LULC; Step-wise land-class elimination

1. Introduction

Remote sensing techniques for the extraction of urban land use information have drawn immense research attention during the last three decades. Satellite Remote Sensing **(**SRS**)** images with increasingly improved optical quality are now invaluable resources for discrimination of land cover types within cities at appreciably finer levels of accuracy. Automatic and semi-automatic methods of detecting urban features from images have evolved with enhanced details of information captured by more powerful sensors. These imageries, being available historically, allow determination of the dynamic nature of urbanization, characterized by alteration of the natural land areas into unreceptive built-up areas (DeFries, 2008, Mushore et al., 2017). Apart from such modifications in the urban core, urban areas encroach on to the surrounding natural lands such as agricultural fields, forestlands and wetlands as peri-urban dynamic processes (Xu, 2008; Estoque & Murayama, 2015). Remote sensing based (timely and cost-effective) mapping of built-up areas facilitates subsequent urban analysis and modeling of the urban morphology, urban hydrology, urban ecology and urban climate. The significance of expanding built-up areas in intensifying Urban Heat Island (UHI) phenomena cannot be ignored (Stone, 2007; Xu, 2008; Gibril et al., 2017). Apropos recent global warming and climate change concerns, it has been suggested that the warming trends observed at continental weather stations could be partly influenced by rapid urbanization and local UHIs (Jones et al., 2008; Parker, 2010; Khan et al., 2015). As urbanization is an alarming and accelerating phenomenon across the world, especially in developing countries like India, the investigation of UHI effects has accrued momentum (DeFries & Pandey, 2010; Liu & Yang, 2015).

The study of spatial expansion of urban built-up areas always needs accurate information on the size, shape and spatial structure of built-up features (Xu, 2008; Wang et al., 2015). Therefore, a scientific and robust technique is required to quickly retrieve such

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

information, and its success depends on the availability of spatio-temporal data. Fortunately, SRS technology offers considerable pledge to meet these requirements. SRS data are useful for monitoring the growth of urban built-up areas as well as their sprawling effects due to their ability to provide varied spatial and temporal resolutions, consistent and repetitive measurements and synoptic views of Land Use and Land Cover (LULC) (Guindon et al., 2004; Xu, 2008; Bhatta, 2009; Griffiths et al., 2010; Bouzekri et al., 2015).Early researchers have made use of SRS imagery to address urban and suburban problems (Jacquin et al., 2008). Several techniques have been formulated, applied and evaluated to achieve desirable level of accuracy in automated mapping of urban LULC. These techniques can be broadly grouped into three categories:(a) pixel-based classification (Jensen, 1986; Gong et al., 1992; Casals-Carrasco et al., 2000; Zhou and Robson, 2001; Dean & Smith, 2003);(b) object-based classification (Giada et al., 2003; Guindon et al., 2004; Gao et al., 2006; Qian et al., 2007; Cleve et al., 2008); and (c) application of spectral indices such as Normalized Difference Built-up Index (NDBI) (Zha et al., 2003; Zhang et al., 2005). Pixel-based approach is still considered efficient for classificatory purposes when employed on multi-band imageries of medium resolution (Adam et al., 2016). Exploring the advantages of the unique spectral signature of built-up areas and other LULC of Thematic Mapper (TM) images, Zha et al. (2003) formulated NDBI for automatic extraction of built-up areas. The method involves simple arithmetic manipulation of re-coded Normalized Difference Vegetation Index (NDVI) and NDBI binary images which indicated vegetation and built-up areas, respectively for pixel value =1. Reportedly, this process is unable to separate urban built-up areas from bare lands (Zha et al., 2003; Zhai et al., 2015). It has been suggested that the universality of the approach needs to be tested in other geographic areas because spectral response patterns of vegetation and other LULC classes are actually more complicated than were assumed (Zha et al., 2003; Bhatti &Tripathi, 2014; Mushore et al., 2017). The problem is more acute in case of

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

old cities like Kolkata, where mixed urban development has continued throughout its history of urbanization. As a refinement of the original method, He et al. (2010) suggested that continuous images of NDBI and NDVI be used, not the binary images and then a final separation of built-up areas on the basis of an optimal threshold value determined adopting Double-window Flexible Pace Search (DFPS) technique (Chen et al., 2003). This modified approach has been demonstrated to be comparatively better than the customary NDBI approach in terms of overall accuracy (He et al., 2010; Bhatti & Tripathi, 2014; Lu et al., 2015).

NDBI has been widely used for the mapping of urban built-up areas from TM images. Afterwards, Bhatti & Tripathi (2014) tested the applicability of this approach on Landsat-8 Operational Land Imager (OLI) and proposed a new Built-up Area Extraction Method (BAEM) for improving the performance of index-based separation of built-up pixels. This new approach considers Principal Component Analysis (PCA) as a method to ensure optimal utilization of information in image data and accordingly uses the first Principal Component (PC) image of the highly correlated bands in Landsat-8 OLI imagery to replace the spectrally equivalent TM band in the original formula for computing NDBI. Similarly, strategies for calculating NDVI and Normalized Difference Water Index (NDWI) were modified by substitution of TM bands with corresponding OLI bands. Finally, the BAEM image was obtained by simple arithmetic operation on modified NDBI, modified NDVI and modified NDBI images of continuous type (Bhatti & Tripathi, 2014). The threshold value for the extraction of built-up areas from this BAEM image was then determined through the DFPS technique. The BAEM subsequently brought a noticeable improvement in NDBI approach to make it commensurate with the Landsat-8 images and also revealed a higher overall accuracy than modified NDBI method (Ke et al., 2015; Piyoosh & Ghosh, 2017).

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

In spite of all these efforts, the accuracy in extraction of built-up area from satellite images still lies below 80% and currently available methods fail to completely separate built-up areas from bare lands (Pant et al., 2016). Results from a lot of studies revealed that there has not been a single technique to classify the built-up areas but one combines the intrinsic aspect of two or more techniques in order to build a hybrid approach to improve the extraction accuracy of urban LULC.Hence, a satisfactory technique with justifiable approach is necessary. The major challenges encountered while analyzing raster data in wavelength domain arehigh probability of misclassification due to mixed pixel problem (Lo & Choi, 2004) and the juxtaposition of discrete and continuous features in an extremely heterogeneous urban landscape (Aplin, 2006). Moreover, Very High Resolution (VHR) images are prohibitively expensive for urban agencies and researchers to be used in real applications. One way of handling these difficulties is to develop a regression model based on digital number (DN) values of sample pixels from those bands which can significantly distinguish a feature class from the rest, and to apply such models on medium resolution images.

In light of the original NDBI and BAEM approach's recompense and drawback, a new and simple method is proposed here for the rapid and accurate mapping of urban built-up areas without depending exclusively on the band specific normalized indices, in order to pursue a more generalized approach befitting both TM and OLI images. The indifference in the spectral response of built-up areas and bare soils to thermal bands makes the separation of these two land categories difficult, particularly in tropical areas. Encountering this problem for the city of Kolkata, the proposed method excludes the thermal bands in extraction of built-up areas. This new approach, called Step-wise Land-class Elimination Approach (SLEA) combines the use of a single band layer, an NDVI image and another binary image obtained through Logit model. Based on the spectral designation of the satellite image, a

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

particular band is chosen for identification of water pixels. The DFPS approach is employed for finding the optimum threshold value that segments the selected band image into water and non-water categories - the water pixels are then eliminated from the original image in use. The vegetation pixels are similarly identified using the NDVI image and eliminated. The residual pixels left after eliminating water and vegetation categories are assumed to be the built-up and bare land categories, and thus need to be separated. For extracting built-up areas, Logit model is used to select the probable image bands appropriate for recognizing the land characters. It allows, after coefficient optimization, to model the link between a set of explanatory bands and a dichotomous dependent output which, at a given probability, reveals if a pixel belongs to the built-up class or not. The mapping is accomplished through recoding of the final image.This method, importantly, does not involve much of subjective human intervention in the mapping process. The effectiveness of this method was tested through the mapping of built-up areas in the Kolkata metropolis of India from TM images of 2000, 2005 and 2010, and OLI image of 2015. Comparison of the results obtained using this method with those interpreted manually demonstrates that the proposed approach is quite reliable. This method produced very accurate results in a more efficient manner compared to the earlier built-up area extraction approaches.

The past study revealed that most of the studies developed spectral indices for rapid extraction of built-up areas from various types of satellite images based on specific image data (Sameen & Pradhan, 2016). Therefore, it is needed to improve band specific approach for Landsat satellite images, which is not examined in previous study. In addition, inspection further advanced and robust methods are prerequisite to be involved in the improvement of spectral response for tropical environment. Therefore, this study proposes a novel spectral land-class elimination approach for rapid extraction of built-up areas from Landsat images. For this, researchers' use of a single band layer, an NDVI image and another binary image

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

obtained through Logit model. This approach offers paramount solution for numerous problems including the selection of relevant bands and determining the best combination of spectral bands for the elimination of final land-class in tropical city.

2. Study Area

The present study of built-up area extraction was conducted on Kolkata Metropolitan Area (KMA), delimited by the latitudes of 22.01°N and 23.08°N, and the longitudes of 88.03°E and 88.45°E, developed mainly along the banks of the river Hugli about 150 km to the north of the Bay of Bengal (BoB), right over the Gangetic delta plains (Fig. 1). This is one of the leading urbanized areas in India, as well as of the urban world characterized by swift urbanization and associated spatio-temporal variation in biophysical processes. KMA, the largest urban agglomeration in eastern India, extending over 1851.41km^{2}comprises 4 municipal corporations, 39 municipalities, and 1 cantonment, and parts of 24 panchayat samiti (rural local governments at the intermediate level in administrative structure of India)(Census of India, 2011). Kolkata is the most important commercial and industrialized centre of east and northeast Indian hinterland, and it holds essential manufacturing and transportation infrastructure (Nath et al., 2015).The population of core Kolkata was 1.5 million in 1901, 11 million in 1991 and an exceptional 14.2 million in 2011 (Census of India, 2011). In some parts of the KMA, the population density exceeds10,000 persons/km^2. Owing to the huge demand of land for housing the tremendously growing population, accommodating the ever increasing vehicular traffic and facilitating the booming trade and commerce, the city has intensified the use of land and, at the same time, has also intruded into the adjacent natural back swamp and marshy lands to the east. Planned development, particularly in the Salt Lake City and New Town-Rajarhat areas, and spontaneous haphazard growth in inner city areas characterize the urbanization process in Kolkata. Both filling up of

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

widespread natural areas in the former instance, and densification of landuse in the later, are adding to the intensity of the heat island effect in Kolkata at the city scale as well as at the neighbourhood level (Stone, 2007; Bhatta et al., 2009; Khan, & Chatterjee, 2016). More than 80% of land in the city core is built-up with high-rise residential buildings, overcrowded commercial areas, hospitals and different other institutions (Nandy, 2007). Aesthetically, the city exhibits wide variations from leafy residential areas like Alipore and Rajarhat to 'grey' areas of Binoy-Badal-Dinesh (BBD) Bag (erstwhile Dalhousie Square). Kolkata's business district shows unmistakable signs of haphazard developments mirroring the reality in other big cities in the country.

<Fig. 1>

3. Methods and materials

3.1 Pre-processing of imagery

Images from two different sensors of Landsat satellite series were used for delineation of built-up areas in Kolkata metropolis i.e. TM of Landsat-5 and OLI and Thermal Infrared Sensor (TIRS) of Landsat-8 (Table 1). All multi-temporal images were pre-processed in ERDAS 2014 and subsequent analysis was carried out in ArcGIS 10.1 software. Only the cloud-free images were selected for this study, hence no atmospheric correction was required (Deng & Wu, 2013; Bhatti & Tripathi, 2014). For the purpose of comparing three different approaches (NDBI, BAEM and SLEA), the entire set of bands of all the selected images required necessary pre-processing, though the proposed SLEA does not include the thermal bands. Initially, a rectangular area encompassing the KMA was clipped from the original images in order to avoid null pixels at the processing stage. Careful observation of the standard False Colour Composite (FCC) band combination helped to distinguish diverse types of LULC (Zha et al., 2003). Field survey and unsupervised classification of the images with variable number of classes revealed that the area comprises four principal land-class

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

types namely water bodies, vegetated areas (including other green coverage), bare lands and built-up areas. TM band 4 and 5 were most pertinent for discerning the urban built-up areas from other LULC types (Zha et al., 2003; He et al., 2010; Bhatti & Tripathi, 2014). In case of OLI bands, the spectral ranges differ from those of TM, so, it was necessary to examine the spectral signatures of built-up areas in different bands of OLI bands before developing any suitable method for built-up areas extraction (Bhatti & Tripathi, 2014). For this, the spectral signatures associated with built-up, vegetation and water areas in different bands of TM and OLI were examined. The plots for the Digital Number (DN) values of 20 sample pixels each from every land class typein: (a) optical bands 1-5 and 7 of Landsat-5 TMand; (b) optical bands 2–7 of Landsat-8 OLI are shown in Fig. 2. It reveals that vegetation show higher spectral reflectance than other LULC classes in the TM Band 4 and OLI band 5. Therefore, subtraction of band 3 from band 4 (for TM images) and subtraction of band 4 from band 5 (for OLI images) produce positive values for vegetation pixels (Zha et al., 2003; Bhatti & Tripathi, 2014).

<Table 1> Satellite data used in the study.

<Fig. 2>

The multi-temporal sub-set (clipped) images have been geometrically rectified using 200 Ground Control Points (GCPs). The Root Mean Square Error (RMSE) at all GCPs ranged from 0.25-0.86 pixels, which is acceptable and quite suitable for further processing (Gibson, 2000; Askne et al., 2003). Afterwards, the original images were projected into Universal Transverse Mercator (UTM) system and resampled to the same spatial resolution using nearest neighbour technique (Zha et al., 2003). In case of Landsat 8 OLI/TIRS, digital DNs of optical bands 2-8 were converted to Top of Atmosphere (ToA) reflectance and DNs of thermal bands 10 and 11 were first converted to ToA reflectance and subsequently processed to obtain at-satellite brightness temperature in °C (USGS, 2013; Bhatti & Tripathi,

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

2014). The working formulae for translating the DN values into ToA reflectance and at-satellite brightness are available in the Landsat-8 data user handbook (USGS, 2013). The optical bands 2-7 (except the panchromatic band 8) were then stacked into a multi-band image file and similarly, the thermal bands 10 and 11 were stacked into another image file. The spectrally rectified optical stack (bands 2-7, spatial resolution 30m) was then combined with high-resolution panchromatic layer (band 8, spatial resolution 15m) by using the High Pass Filter (HPF) resolution merge function to obtain an output (spatial resolution 15m) with greater detail (Gangkofner et al., 2008: Bhatti and Tripathi, 2014). The values of the parameters, used in this image sharpening process, were set accordingly. The Fig. 3 compares the original image with spatially enhanced image for a sample area of the Salt Lake City in Kolkata. The thermal bands 10 and 11 were resampled to a spatial resolution of 15m, in order to make them compatible with the spectral bands.

<Fig. 3>

In case of Landsat-5 TM images for 2000, 2005 and 2010, the above resolution merging was not possible as the sensor does not deliver any panchromatic band of higher resolution. Therefore, after re-projection and re-sampling, the bands 1-5 and 7 (exceptthe thermal band 6) were directly stacked into a multi-band image file for further processing.

3.2 Built up area extraction

The proposed method for extraction of built-up areas using TM and OLI imagery considers elimination, in phases, of the LULC categories, and retains the built-up areas finally (Fig. 4). Starting with all the component LULC classes, water and vegetation pixels were detected and excluded, and the resultant image was then used as the input to a Logit model for separation of built-up pixels from the bare land pixels.

3.2.1 Step 1: Elimination of water class

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

The most significant spectral character of water is energy absorption at the Shortwave Infrared (SWIR) wavelengths. A single band density slices classification of Landsat-5 images using the TM band 5 (1.55-1.75μm) produces 96.9% accuracy in mapping of water boundaries. The optical bands are not fit for the method, but while using TM band 5, the performance of density slicing in detecting water pixels is comparable with multispectral maximum likelihood algorithm (Frazier et al., 2000). Band 6 of OLI image provides similar data in the SWIR1 band (1.57 – 16.5μm), though the spectral range is a bit narrower than the TM band 5. Therefore, band 5 of Landsat-5 images and band 6 of Landsat-8 images were selected since they exhibit comparable spectral character suitable for the purpose of detecting water pixels. The greyscale image (band 5 for TM or band 6 for OLI) is segmented into water and non-water classes by thresholding method where the spatially continuous pixels are partitioned into two disjoint sets in reference to a fixed constant pixel value, called *threshold value*. The pixels with values less than the threshold value indicate water and such pixels were re-coded to 0, while pixels with values greater than the threshold value signify non-water and they were re-coded to 1. Thus, thresholding converts the grey-level band images into a binary image by turning all water pixels to black and others to white. A DFPS search approach was used for fixing the threshold value.

<Fig. 4>

3.2.1.1 Double-window flexible pace search

The DFPS is a systematic approach to determine the optimum threshold value between change and no-change pixels while examining LULC transitions (Chen et al., 2003; He et al., 2010; Bhatti & Tripathi, 2014). This technique is an improved version of Change Vector Analysis (CVA) which determines the threshold of change first and then determines change direction (from one category to another) by application of minimum-distance classification algorithm on direction cosines of the change vectors (Chen et al., 2003). DFPS

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

approach has been widely used to determine optimum threshold value for recognizing vegetation, water and bare land in a thematic image while extracting built-up areas (He et al., 2010; Bhatti & Tripathi, 2014). The technique is based on choosing an optimum threshold from representative training samples, which should contain all possible feature classes to represent in the study area optimally (Chen et al., 2003). It relies on determination of threshold value at the highest level of accuracy such that the value eliminates all the pixels under a targeted class from all other classes present in the sample area (Chen et al., 2003; He et al., 2010; Bhatti & Tripathi, 2014). The technique involves some steps to apply in SLEA method (Fig 5).

Fig. 5

3.2.1.2 Selection of typical sample area

Selection of representative training sample areas in processed image is decisive for actual extraction of water areas. In this study, the sample water bodies have been selected through rapid visual survey of study area and visual interpretation of the processed image, subject to the condition that sample areas include only water pixels in inner window and surrounded by non-water pixels in the outer window (Fig 6).

<Fig. 6>

3.2.1.3 Determination of search range and pace

In this stage, the threshold range has been setup using simple arithmetic difference between minimum pixel value *(a)* and the maximum pixel value *(b)* of the magnitude of change in processed image by first search (Chen et al., 2003; He et al., 2010; Bhatti & Tripathi, 2014). The pace of first search, P_1 could be computed by Equation 1 (Chen et al., 2003).

$$P_1 = \frac{(b-a)}{m} \qquad (1)$$

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

where, *m* is the positive integer that reveals the number of optimum threshold in a iteration search and it can be manually determined. The optimum thresholds to detect magnitude change pixels from training sample in a iteration search process are given within the range of (a,b) as $(b-P_1, b-2P_1, ..., b-mP_1)$. It is to be remembered that the value of *m* has no bearing on the search process and ultimate results, but it is a time dependant parameter. A large value of *m* increases to the number of optimum threshold during first search, but it gradually decreases as the search process is iterated (Bhatti & Tripathi, 2014).

3.2.1.4 Calculation of test parameter

The procedure for calculating test parameter has been slightly modified from that suggested by Chen et al. (2003). A success rate of target pixel detection is defined to evaluate the performance of each potential threshold during one search process for identifying change/no-change pixels. The success rate (L_k) is calculated for a potential threshold of k using Equation 2.

$$L_k = \left(\frac{A_{k_1}}{A_1} \times \frac{A_{k_2}}{A_2} \right) \times 100 \qquad (2)$$

where, A_1 is the total number of pixels under water within the inner window frame, A_{k_1} is the number of pixels detected correctly as water inside the inner window, A_2 is the total number of pixels in the non-water rectangular strip lying between the boundaries of outer and inner window, and A_{k_2} is the number of pixels correctly detected as non-water in A_2 .For all *m* thresholds in one particular search process, the maximum and minimum values of L_k were calculated and assigned to $L_{\max}$ and $L_{\min}$, respectively. These two values were examined whether or not they satisfied the exit condition given in the final step of the DFPS process. If the exit condition was not satisfied, a new search began with a new search range ($k_{\max} - P_1$,

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

$k_{max} + P_1$) and a new search pace that was calculated using Equation (2), where k_{max} was the threshold value corresponding to L_{max} in that particular search iteration.

3.2.1.5 Condition to exit the iteration

The condition to exit from the iteration search process was performed until differences between L_{max} and L_{min} satisfied the stipulation given in Equation 3.

$$L_{max} - L_{min} \leq \delta \quad (3)$$

where δ denotes an acceptable error in the search process, and its value is 1 in case of an inter-type image. The desired threshold value to separate water areas from other areas in the continuous image was equivalent to L_{max} in the final iteration of DFPS search technique. Thus, the said value has been employed in continuous layers of band 5 (for TM images of 2000, 2005 and 2010) and band 6 (for OLI image of 2015) to separate water bodies; waterpixels were assigned a value of 0 and the rest were given a value of 1.

The output binary image was multiplied to their corresponding multi-band image so that the water pixels in each layer of the stacked images receive zero value while the other pixels retain their original values. Consequently, the water areas were eliminated and the output images were ready for subsequent analyses.

3.2.2 Step 2: Elimination of vegetation class

The elimination of vegetation pixels follows the same technique as employed for the delineation of water (see Step 1 above); the only difference is that it was applied to the NDVI images instead of any single band. Continuous NDVI images were obtained through raster calculation using Equation 4 that involves Red (R) and Near Infrared (NIR) bands.

$$NDVI = \frac{NIR - R}{NIR + R} \quad (4)$$

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

The red band corresponds to the band 3 of TM and band 4 of OLI, while the infrared band corresponds to the band 4 of TM and band 5 of OLI images. Accordingly, the formula was applied for preparing the NDVI image.

The DFPS method, with all its steps, was applied to the NDVI images to detect the pixels that cover the green areas including woodlands, crop fields and vegetated wetlands. The threshold based classification segmented the vegetation pixels from the pixels under other land classes and this classified image was re-coded to obtain a binary image with values 0 for vegetation and 1 for others. The image was then multiplied to the output image from step 1 to eliminate the vegetation pixels. The multi-band images layers now comprises the pixels under only two categories- built-up and bare land.

3.2.3Step 3: Separating bare lands from built-up areas

In an effort to segregate built-up pixels from the bare land pixels, one Logit model each was formulated for the images of 2000, 2005, 2010 and 2015 on the basis of the DN values of the pixels sampled from the corresponding images obtained from Step 2 above.The proposed method includes several components: (1) training and test data preparation;(2) formulation of Logit model; (3) accuracy assessment of the model using test data set; and (4) selection of the final model by repeating the steps from 2 to 3 until the best level of accuracy is achieved.

3.2.3.1 Sampling pixels and preparation of data sets

In order to sample pixels, we carried out unsupervised classification of the input image into two categories for a primary detection of built-up and bare land pixels, and both the categories were sampled randomly. The sampled points, originally saved in shape file format, were converted into a Google Earth (GE) Keyhole Markup Language (KML) file for verification of the land class that the sample points represent. In the process, we took the advantage of the availability of historical high resolution Landsat GeoCover imageries for the

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

study area on GE. Moreover, horizontal positional accuracy of GE high-resolution imagery is sufficient for assessing moderate-resolution satellite images for majority of the world cities (Potere, 2008). Sample points in KML file were drawn on GE images and any mismatch in land class category of the sample points, led to withdrawal of the points from the sample pool. A final round of verification was done with ground-based survey using Global Positioning System (GPS) receivers and parcel level detail landuse map of municipal wards available for the study period. The size of the sample was 576, same for all the images processed. Sampling was followed by building of database that includes the DN values of the sampled pixels in each band of the respective layer-stacks (independent variables), and the land class of the pixels in a binary format, 1 for built-up and 0 for bare land (dependent variable). The data was split into two parts in the proportion of 60 to 40 for preparing the training and test data sets, respectively, which were used in formulating the Logit model.

3.2.3.2 Logit model

Logit model, also known as logistic regression (Dyke and Patterson, 1952; Cox, 1958; Cox, 1970; Sankoff and Labov, 1979), is a promising approach for modeling binary responses $Y = y_i \in \{1, 0\}$ as the linear combination of one or more continuous or categorical predictor variables $(x_1, x_2, x_3 \ldots\ldots x_n)$, plus some normalized random error, ε (and optionally the intercept β_0), expressed as (Equation 5,6 and 7):

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots\ldots\ldots + \beta_n x_n + \varepsilon \quad (5)$$

$$E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots\ldots\ldots + \beta_n x_n \quad (6)$$

$$E(Y) = \mathbf{x}'\boldsymbol{\beta} \quad (7)$$

where, $E(Y)$ is the expected value of Y, vector $\mathbf{x}' = (1, \mathrm{x}_1, x_2, \ldots\ldots\ldots, x_n)^T$ and vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \ldots\ldots\ldots, \beta_n)^T$ .In order to accommodate categorical nature of Y, the model considers $p = P(Y = 1 \mid \mathbf{x}')$ i.e. probability of the success of an observation that assumes a

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

logistic distribution bounded between 0 and 1 ($p = e^{\mathbf{x}'\boldsymbol{\beta}}/1+e^{\mathbf{x}'\boldsymbol{\beta}} = 1/1+e^{-\mathbf{x}'\boldsymbol{\beta}}$) and uses natural logarithm of odds called *Logit* $\left[logit(p) = \ln(p/1-p)\right]$ as the output variable that ranges unboundedly from -∞ to +∞. Thus, the Logit function transforms the probability space into a log-odds space where the logistic relation becomes linear as (Equation 8).

$$\ln \frac{p}{1-p} = Y' = \mathbf{x}'\boldsymbol{\beta} \qquad (8)$$

The ordinary Logit model aims at estimating the parameters $(\beta_0, \beta_1, \beta_2, ..., \beta_n)$ from a training data with observed responses y_1 through y_m, by way of maximizing a likelihood function defined as (Equation 9):

$$L(p; y_1, y_2, ..., y_m) = f(y_1, y_2, ..., y_m; p) = \prod_{i=1}^{m} P(Y = y_i) = p^k (1-p)^{m-k} \qquad (9)$$

where, $k = \sum_{i=1}^{m} y_i$ and $0 \le p \le 1$.

Statistical conditions required for application of Logit model are fulfilled in the present context of multiband image transformation to obtain a binary image depicting probability of a pixel to be under built-up class. Logistic regression is independent of many of the vital assumptions made in the ordinary least square approach. It does not need that the dependent and independent variables to be in a linear relationship, because log-odds transformation of the dependent variable makes the system capable of handling all types of relationship in a linear framework. It is not mandatory for the independent variables to be multivariate normal and homoscedastic. Land class of a sample pixel was entered into the model in a binary form because the logistic regression needs its dependent variable to be dichotomous. The issue of over fitting or under fitting of the model was tackled by retaining only the significant independent variables in the model. The sampling strategy complies with the recommendation of 20-30 samples per independent variable.

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

For the present purpose of separating built-up areas from bare lands, Logit model is implemented on the multiband image containing those two land classes only (output from Step 2). This entails digital transformation of input image into a single band output image by combining selected bands through linear operation.

3.2.3.3 Assessment of model fit

An iterative modeling strategy was adopted where predictive bands were progressively removed (Campling et al., 2002), based on significance test for the co-efficients in the model using Wald's method. The Wald's test considers square of the ratio of regression coefficients to its standard error and thus assesses the statistical significance of each co-efficient in the model. The test suggested the variables to be retained in the next round of repetition. On every reiteration, fitting of the model was further confirmed by assessing its predictive capacity through application of the model on the test data, at a chosen level of probability. A model is finalized when it includes only the statistically significant predictive bands and has attained a satisfactory level of accuracy (> 95%) in predicting the output.

Implementation of the final model on the appropriate image produces a single band raster layer in greyscale where the pixel value represents the probability of a pixel to be the built-up area. Finally, segmentation of the output image by use of DFPS technique gives the desired map of built-up area.

4. Results and discussion

The mapping of mixed and swift urban growth and their spatial pattern in relation to LULC is an indispensablerequirementfor urban planning. One of the foremost tasks for this purpose is the actual mapping of urban built up area. The issue is more crucial for the cities of the developing countries like India where haphazard unplanned use of land has led to mixed type of development where the continuity of a land class is often interrupted by

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

appearance of another class. This situation makes the extraction of built-up area challenging due to adjacency and similar spectral response of built-up area and bare land. The problem is still unresolved in spite of several attempts made to address the issue.

4.1 Previous approaches: NDBI and BAEM

The NDBI Zha et al. (2003) has been useful for automatic mapping urban built-up areas using TM data. It was prepared on the basis of the unique spectral signatures and responses of built up land and other LULC. The authors have argued that the built up areas and open bare land provide a sharp increment in their reflectance from band 4 to band 5 and this pace of increment greatly exceeds that of any other LULC (Zha et al., 2003). The normalized differentiation of thesetwo bands is given by Zha et al., 2003 as (Equation 10):

$$NDBI = \frac{(TM5 - TM4)}{(TM5 + TM4)} \begin{pmatrix} close\ to\ 0 = vegetation \\ 0\ to\ -1 = waterbodies \\ 0\ to\ 1 = built\ up \end{pmatrix} \quad (10)$$

The index was developed through simple arithmetic manipulation of two bands from TM satellite imagery. In the output image, the values range between -1 and 1 where the built up areas and bare land pixels have positive values which allow the built up areas to be mapped automatically.

Recently, the applicability of NDBI to the Landsat-8 OLI imagery was examined and tested by Bhatti & Tripathi (2014), and they proposed a new technique for urban built up area extraction by utilizing PCA of similar spectral bands and thermal bands of OLI data. The new method called BAEM was the revised version of a modified NDBI approach (He et al., 2010), and was employed on Landsat-8 OLI imagery. The possibility of built up areas is given by Bhatti & Tripathi (2014) as:

$$BAEM = (NDBI - NDVI - MNDWI) \begin{pmatrix} high\ DN = built\ up \\ low\ DN = other\ land \end{pmatrix} \quad (11)$$

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

The BAEM approach is an amalgamation and integration of Land Surface Temperature (LST), NDVI, and Modified Normalized Difference Water Index (MNDWI) given by Xu (2005). The BAEM approach was able to extract the built up areas from OLI images more satisfactorily as compared to the modified NDBI approach. Instead of using binary NDBI, NDVI and MNDWI images directly, continuous images of these indices were employed and the outputs were re-coded to map the built-up areas based on optimal threshold value determined through the DFPS method.

The aforementioned methods of NDBI and BAEM were applied and tested in this study, and it was found that both the methods could not completely separate built-up areas from bare land in case of Kolkata, one of the leading metropolitan areas in eastern India. These two techniques of built-up area extraction are more suitable for temperate climate where a distinct contrast exits in spectral reflectance from built up lands and open bare lands. In the case of tropical environment, however, these methods overestimates the built-up areas because no distinct contrast exits in spectral reflectance from built up lands and open bare lands; both the land categories are classified as built up in some parts of the study area, particularly in margins of the city.Results show that many of the built-up areas could not be identified by the index-based methods due to less difference between spectral reflectance between built-up areas and other land-classes in the bands used in building the indices. NDBI and BAEM techniques returned exceptionally good results for continuously built-up areas but failed to detect built-up pixels from areas of mixed-type LULC. Thus, for the tropical cities with haphazard growth, these methods do not provide the assessment of urban built up areas with desirable accuracy because no spatial thresholding method has been suggested to separate the misclassified pixels. This is a major problem and challenge for urban planning and policy implementation purposes.

4.2 Proposed SLEA approach

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

The approaches developed so far for built-up area extraction suffers from generic limitations either from technical or from methodological or from applicative point of view. The present study is an attempt to eliminate those discrepancies and proposes a simple technique that involves application of Logit model and DFPS approach to precisely map urban built-up areas by successive elimination of other land classes. The suggested method is more straightforward and eliminates users' control over the output. When applied in case of Kolkata metropolis, it turned out to be more efficient than the other methods available.

The impervious surfaces have effects on the land surface temperature which is directly related to their thermal properties and therefore, the impervious urban surface covers exhibit direct relation with the intensity of reflectance in the thermal bands (Oke, 1973; Weng, 2001; Yuan & Bauer, 2007). The uncovered land parcels within the built-up areas have spectral reflectance in thermal bands which is comparable to that of the impervious built-up areas. Techniques of built-up area extraction using thermal bands are, therefore, prone to overstimulation. For this, thermal bands were excluded from the analysis. The method, therefore, only considers the optical bands of TM and OLI sensors. After elimination of the water bodies and the green covers, Logit model was employed to eliminate bare lands. The Logit model was finalized on the basis of best level of accuracy obtained through several trials with different combinations of the bands. The model retains only those bands for which the co-efficient in the linear combination are statistically significant. Thus the built-up areas extraction approach was simple and follows a strategy of eliminating other land classes to finally obtain the built-up areas.

The spectral features and ranges of TM bands differ from those of OLI, accordingly, it was essential to examine the signatures of spectral ranges of LULC class in different bands of TM and OLI data before applyingthis sample based approach of eliminating the unintended land classes and subsequent extraction of built-up areas. For this reason, the spectral response

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

of water bodies, vegetation areas and built-up areas in different band of TM and OLI were examined after necessary pre-processing. It was found that the band 4 (0.76-0.90μm) and band 5 (1.55-1.75μm) of TM data are closely analogous to band 5 (0.85-0.88μm) and band 6 (1.57-1.65μm) of OLI data, respectively, in respect of spectral character. The spectral responses of the major land classes differ significantly in some bands. For example, water responds quite differently than built-up and vegetation class in band 5 and 6 of OLI while such variation is distinct in band 4 and 5 of TM (Fig. 7). Importantly, Band 4 and band 5 of TM are very strongly correlated at a high level of statistical significance and a similarly strong correlation exists between the band 6 and 7 of OLI (Bhatti & Tripathi, 2014).

<Fig. 7>

In SLEA, each of the land classes to be eliminated are identified and separated out of the original raster image sequentially. The continuous image was first converted to unsigned integer, and then its histogram was plotted and examined to set the initial range of search. The search process for identifying water pixels in the Landsat 8 OLI image of 2015 and the results obtained thereby are shown in the Table 2, as an example. This is a 16 bit data with DN values ranging between 0 and 65535. For the first cycle, 13 individual searches were desired and accordingly threshold values were set at an interval of 5000 within the said range of DN values. The first search turned out to be successful (with success rate > 0) for the threshold values 15535, 10535 and 20535 suggesting that the second round of searches to start with threshold value 15535 and be continued till 5535. The pace or interval for the second search was considered to be 1000. In subsequent searches, both the search range and pace gradually diminished to achieve a precise threshold value for which the success rate attains the maxima. For the image under consideration, the search process identified 8355to be the optimal DN value for segmentation of the image into water and non-water classes with

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

a success rate of 96.35%. The Table 3 summarises the results of DFPS run for all other similar pursuits of image segmentation involved in the study.

<Table 2> An example of the DFPS for the determination of optimal threshold value of water pixels of the Landsat-8 OLI image of 2015

<Table 3>Summary results of the DFPS for determination optimum threshold values in eliminating the water and vegetation pixels and final extraction of built-up areas.

After eliminating the water and vegetation pixels (values recoded as zero), the multi-band images with layers of selected bands were sampled for preparing pixel-based database to be employed in formulating Logit model that uses the image bands as the independent variables. The best fit Logit model transforms the corresponding multi-band image into a binary image which when segmented through DFPS delineates the built-up areas (Fig. 8). In formulating the Logit model, an iterative approach was adopted that initiated with all the bands in the image progressively discarded insignificant bands from the model. Optimality of the model was judged by the level of accuracy measured with reference to the test data. Results show that for the TM image of 2000, the best model involves band 5 and band 7 with an accuracy of 98.21%, signifying that the model is highly accurate in distinguishing built-up area from bare lands (Table 4). The results for the images of other years are summarized in Table 3. The pixels with values abovethe optimum threshold were classified as built-up areas (Fig. 9).

<Table 4> Co-efficient values of Logit model for built up area extraction.

<Fig. 8>

<Fig. 9>

4.3 Built-up areas and level and agreement of accuracy

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

In order to assess the classification performance, the results of the proposed SLEA were verified at 350 locations by rapid visual survey with corresponding image pixels using a stratified random sampling technique (Congalton, 1991). At first, a specific area was split into two strata, built-up areas and non-built-up areas, and then samples were taken from each one stratum on the basis of proportion in the output image (Bhatti & Tripathi, 2014). In addition, both NDBI and BAEM approaches were also used to generate a map of built-up areas, which was tested using the same accuracy assessment method in order to compare the performance of the three approaches (Fig. 10). Producer and user accuracy of three approaches were compared to each other for assessing the classification efficiency. Selected sampled areas were visually observed to comprehend the discrepancies in the outputs of three approaches. The Fig. 10 demonstrates the misclassification of built-up area in NDBI and BAEM approach very clearly. The red 'A' rectangles indicate sample areas with large number of bare land parcels situated within the developed areas which could not be detected by NDBI and BAEM methods and those bare lands have been classed under the built-up areas. The disconnected built-up areas interspersed with homestead have been underestimated in both NDBI and BAEM approaches. Such areas were sampled, indicated by red 'B' rectangles in the Fig. 10.

<Fig. 10>

The confusion matrix has been derived for NDBI, BAEM and SLEA to assess the accuracy of the three approaches (Table 5). Results of the proposed SLEA were 95.33% accurate on the whole, while those derived by the NDBI and BAEM approaches returned an overall accuracy of 83.67% and 89.33%, respectively. These comparative accuracy discrepancies occur mainly due to incorrect detection of built-up and non-built-up areas in different approaches. The commission error greater than the omission error in NDBI and BAEM approaches has resulted in outputs containing some pixels wrongly detected as built-up areas. However, the

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

least omission error signifies that the approach is capable to capture the built-up areas well (Table 6). Nevertheless, least values of omission and commission errors, represented by the kappa index, suggested that the SLEA has the highest agreement level or overall accuracy than the NDBI and BAEM approaches.

<Table 5>Confusion matrix for NDBI, BAEM and SLEA

<Table 6> Summary of the accuracy assessment of the images derived by NDBI, BAEM and SLEA

For the purpose of comparison, the input TM and OLI images and final outputs of three built-up areas extraction approaches have been displayed in Fig. 10 showing the same sample locations. It was found that the NDBI and BAEM approaches have over classified (inclusion of bare lands) the built-up areas as compared to the SLEA. The NDBI and BAEM classifications include higher commission error in the outputs than that obtained through the SLEA method. The use of thermal bands in BAEM approach helped in separating less dense built-up areas from other LULC types (Bhatti & Tripathi, 2014). However, in case of Kolkata, inclusion of thermal bands and seasonality in BAEM approach misclassified a considerable number of bare lands pixels as built-up areas, and thus reduced the overall accuracy of BAEM approach. On the other hand, NDBI approach is also incapable to segregate built-up areas from open bare lands because built-up areas and bare lands have comparable spectral signatures in TM bands (Zha et al., 2003). Zha et al., (2003) have suggested that the drawback may be tackled through the integration of spatial knowledge, that the bare lands are located next to water. But obviously, this is totally a time dependant protocol. The established methods in use could not completely separate the bare lands from built-up areas because of the resemblances in spectral response of these two LULC classes which may be due to monotonous nature of surface, seasonality and climate type of the study area (Zha et al., 2003; Herold et al. 2004; Lu &Weng, 2004; Lwin & Murayama, 2013; Bhatti

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

& Tripathi, 2014).Thus, in case of Kolkata metropolis, the rationale for the high commission error may be endorsed to spectral mixing of built-up pixels with bare lands. It is interesting to note that SLEA was able to recognize water, vegetation and bare lands from continuously built-up areas, and to eliminate them for separating out the built-up areas. Moreover, the integration of optical band-based Logit model for segregating built up areas from bare lands increased the performance of the proposed land class elimination technique.

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

5. Conclusion

Both NDBI and BAEM approaches were capable to map continuously built-up areas but they could not perform satisfactorily when applied to highly intermixed type built-up areas of tropical environment as in the case of Kolkata metropolis. This study proves that the accuracy and objectivity of NDBI and BAEM approaches decrease towards urban fringe areas where the built-up areas are dotted with bare lands. Both the methods wrongly detected most of the open bare lands as built-up areas. NDBI and BAEM approaches are based on indices which rely on some fixed bands and if any two types of objects have similar spectral responses in these bands, and then the indices fail to distinguish between the objects in question. In the present context, the built-up areas and the bare lands respond comparably to energy in thermal and some visible bands, particularly during the summer. These shortcomings were addressed in this study to develop a method that was simultaneously applicable to both TM and OLI sensor data in order to assess urban growth. The new approach, SLEA, extracts and maps the built-up areas by eliminating water bodies, vegetation and bare lands. This method is capable of distinguishing built-up areas from bare land parcels by using Logit model which flexibly involves only those bands of an image to which the land classes respond distinctly. The Logit model is developed on the basis of sample pixels collected from the image in use, which makes the method free from any theoretical assumption regarding spectral response of

the objects. The SLEA enables to map built-up areas (excluding barren lands) objectively with higher consistency and accuracy even when applied to images of tropical urban areas like KMA. In comparison to NDBI and BAEM approaches, SLEA maps the built-up areas at higher degree of spatial and visual accuracy because SLEA can easily exclude the bare lands from the built-up areas. The confusion matrix has been derived for NDBI, BAEM and SLEA to assess the accuracy of the three approaches. Results of the proposed SLEA were 95.33% accurate on the whole, while those derived by the NDBI and BAEM approaches returned an overall accuracy of 83.67% and 89.33%, respectively. SLEA does not entail a multifaceted mathematical computation protocols or any spatial filtering technique with user defined window at the cost of spatial accuracy of the output. Its methodological simplicity substantially expedites the built-up area extraction accomplished by simple arithmetic manipulation and recoding of intermediate images.

Acknowledgements

We are grateful to the Earth Resources Observation and Science Centre (EROS) and United States Geological Survey (USGS) for offering Landsat data free of charge. We also thank the editor and three anonymous reviewers for their constructive comments, which helped us to improve the manuscript.

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

References

Adam, H. E., Csaplovics, E., & Elhaja, M. E. 2016."A Comparison of Pixel-Based and Object-Based Approaches for Land Use Land Cover Classification in Semi-Arid Areas, Sudan". In *IOP Conference Series: Earth and Environmental Science* (Vol. 37, No. 1, p. 012061). IOP Publishing.

Aplin, P. 2006. "On Scales and Dynamics in Observing the Environment." *International Journal of Remote Sensing*, *27*(11), 2123-2140.

Askne, J., Santoro, M., Smith, G., &Fransson, J. E. 2003."Multitemporal Repeat-Pass SAR Interferometry of Boreal Forests."*Geoscience and Remote Sensing, IEEE Transactions on*, *41*(7), 1540-1550.

Bhatta, B. 2009. "Analysis of Urban Growth Pattern Using Remote Sensing and GIS: A Case Study of Kolkata, India." *International Journal of Remote Sensing*, *30*(18), 4733-4746.

Bhatti S. S., and Tripathi N. K. 2014. "Built-up Area Extraction Using Landsat 8 OLI Imagery." *GIScience & Remote Sensing* 51 (4): 445–467. doi:10.1080/15481603.2014.939539.

Bouzekri, S., Lasbet, A. A., & Lachehab, A. (2015). A new spectral index for extraction of built-up area using landsat-8 data. *Journal of the Indian Society of Remote Sensing*, *43*(4), 867-873.

Campling, P., Gobin, A., &Feyen, J. 2002. "Logistic Modeling to Spatially Predict the Probability of Soil Drainage Classes." *Soil Science Society of America Journal*, *66*(4), 1390-1401.

Chen, J., P. Gong, C. He, R. Pu, and P. Shi. 2003. "Land-Use/Land-Cover Change Detection Using Improved Change-Vector Analysis." *Photogrammetric Engineering & Remote Sensing* 69 (4): 369–379. doi:10.14358/PERS.69.4.369.

Cleve, C., M. Kelly, F. R. Kearns, and M. Moritz. 2008. "Classification of the Wildland–Urban Interface: A Comparison of Pixel- and Object-Based Classifications Using High-Resolution Aerial Photography." *Computers, Environment and Urban Systems* 32 (4): 317–326. doi:10.1016/j.compenvurbsys.2007.10.001.

Congalton, R. G. 1991. "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data." *Remote Sensing of Environment* 37 (1): 35–46. doi:10.1016/0034-4257(91) 90048-B.

DeFries, R. 2008. "Terrestrial Vegetation in the Coupled Human-Earth System: Contributions of Remote Sensing." *Annual Review of Environment and Resources*, 33, 369-390.

DeFries, R., and Pandey, D. 2010. "Urbanization, the Energy Ladder and Forest Transitions in India's Emerging Economy." *Land Use Policy*, *27*(2), 130-138.

Deng, C., and C. Wu. 2013. "A Spatially Adaptive Spectral Mixture Analysis for Mapping Subpixel Urban Impervious Surface Distribution." *Remote Sensing of Environment* 133: 62–70. doi:10.1016/j.rse.2013.02.005.

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. *Ecological Indicators*, *56*, 205-217.

Frazier, P. S., & Page, K. J. 2000)."Water Body Detection and Delineation with Landsat TM Data." *Photogrammetric Engineering and Remote Sensing*, *66*(12), 1461-1468.

Gangkofner, U. G., P. S. Pradhan, and D. W. Holcomb. 2008. "Optimizing the High-Pass Filter Addition Technique for Image Fusion." *Photogrammetric Engineering & Remote Sensing* 74 (9): 1107–1118. doi:10.14358/PERS.74.9.1107.

Gibril, M. B. A., Bakar, S. A., Yao, K., Idrees, M. O., & Pradhan, B. (2017). Fusion of RADARSAT-2 and multispectral optical remote sensing data for LULC extraction in a tropical agricultural area. *Geocarto International*, *32*(7), 735-748.

Gibson, P. J., Power, C. H., Goldin, S. E., & Rudahl, K. T. 2000."Introductory Remote Sensing: Digital Image Processing and Applications" (Vol. 11). London, UK: Routledge.

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

Griffiths, P., P. Hostert, O. Gruebner, and V. D. L. Sebastian. 2010. “Mapping Megacity Growth with Multi-Sensor Data.” *Remote Sensing of Environment* 114 (2): 426–439. doi:10.1016/j.rse.2009.09.012.

Guindon, B., Y. Zhang, and C. Dillabaugh. 2004. “Landsat Urban Mapping Based on a Combined Spectral–Spatial Methodology.” *Remote Sensing of Environment* 92 (2): 218–232. doi:10.1016/ j.rse.2004.06.015.

He, C., P. Shi, D. Xie, and Y. Zhao. 2010. “Improving the Normalized Difference Built-Up Index to Map Urban Built-Up Areas Using a Semiautomatic Segmentation Approach.” *Remote Sensing Letters* 1 (4): 213–221. doi:10.1080/01431161.2010.481681.

Herold, M., D. A. Roberts, M. E. Gardner, and P. E. Dennison. 2004. “Spectrometry for Urban Area Remote Sensing–Development and Analysis of a Spectral Library from 350 to 2400 Nm.” *Remote Sensing of Environment* 91 (3–4): 304–319. doi:10.1016/j.rse.2004.02.013.

Jones, P. D., Lister, D. H., & Li, Q. 2008. “Urbanization Effects in Large-Scale Temperature Records, with an Emphasis on China.” *Journal of Geophysical Research: Atmospheres*, 113(D16).

Ke, Y., Im, J., Lee, J., Gong, H., &Ryu, Y. 2015. “Characteristics of Landsat 8 OLI-Derived NDVI by Comparison with Multiple Satellite Sensors and In-Situ Observations.” *Remote Sensing of Environment*, *164*, 298-313.

Khan, A., &Chatterjee, S. 2016. “Numerical Simulation of Urban Heat Island Intensity Under Urban–Suburban Surface And Reference Site In Kolkata, India.” *Modeling Earth Systems and Environment*, *2*(2), 1-11.

Khan, A., Chatterjee, S., &Bisai, D. 2015.“On the long-term variability of temperature trends and changes in surface air temperature in Kolkata Weather Observatory, West Bengal, India.” *Meteorology Hydrology and Water Management. Research and Operational Applications*, *3*(2), 9-16.

Liu, T., & Yang, X. (2015). Monitoring land changes in an urban area using satellite imagery, GIS and landscape metrics. *Applied Geography*, *56*, 42-54.

Lo, C. P., & Choi, J. 2004. “A Hybrid Approach to Urban Land Use/Cover Mapping Using Landsat 7 Enhanced Thematic Mapper Plus (ETM+) Images.” *International Journal of Remote Sensing*, *25*(14), 2687-2700.

Lu, D., and Q. Weng. 2004. “Spectral Mixture Analysis of the Urban Landscape in Indianapolis with Landsat ETM+ Imagery.” *Photogrammetric Engineering & Remote Sensing* 70 (9): 1053–1062. doi:10.14358/PERS.70.9.1053.

Lu, N., Hernandez, A. J., & Ramsey, R. D. (2015). Land cover dynamics monitoring with Landsat data in Kunming, China: a cost-effective sampling and modelling scheme using Google Earth imagery and random forests. *Geocarto International*, *30*(2), 186-201.

Lwin, K. K., and Y. Murayama. 2013. “Evaluation of Land Cover Classification Based on Multispectral versus Pansharpened Landsat ETM+ Imagery.” *GIScience & Remote Sensing* 50 (4): 458–472.

Mushore, T. D., Mutanga, O., Odindi, J., & Dube, T. (2017). Assessing the potential of integrated Landsat 8 thermal bands, with the traditional reflective bands and derived vegetation indices in classifying urban landscapes. *Geocarto International*, *32*(8), 886-899.

Mushore, T. D., Mutanga, O., Odindi, J., & Dube, T. (2017). Assessing the potential of integrated Landsat 8 thermal bands, with the traditional reflective bands and derived vegetation indices in classifying urban landscapes. *Geocarto International*, *32*(8), 886-899.

Nandy, D. R. 2007. “Need for Seismic Microzonation of Kolkata Megacity.” In *Proceedings of workshop on microzonation, Indian Institute of Science, Bangalore, India* (Vol. 2627).

Nath, S. K., M. D. Adhikari, N. Devaraj, and S. K. Maiti. 2015. "Seismic Vulnerability and Risk Assessment of Kolkata City, India." *Natural Hazards and Earth System Sciences* 15 (6) :1103-1121.

Oke, T. R. 1973. “City Size and the Urban Heat Island.” *Atmospheric Environment* (1967) 7 (8): 769–779. doi:10.1016/0004-6981(73)90140-6.

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

Pant, M., Deep, K., Bansal, J. C., Nagar, A., & Das, K. N. (Eds.). 2016. "*Proceedings of Fifth International Conference on Soft Computing for Problem Solving:" SocProS 2015* (Vol. 1). Springer.

Parker, D. E. 2010. "Urban heat island effects on estimates of observed climate change." *Wiley Interdisciplinary Reviews: Climate Change*, 1(1), 123-133.

Piyoosh, A. K., & Ghosh, S. K. (2017). Development of a modified bare soil and urban index for Landsat 8 satellite data. *Geocarto International*, 1-20.

Potere, D. 2008. "Horizontal Positional Accuracy of Google Earth's High-Resolution Imagery Archive." *Sensors*, *8*(12), 7973-7981.

Powell, R. L., D. A. Roberts, P. E. Dennison, and L. L. Hess. 2007. "Sub-Pixel Mapping of Urban Land Cover Using Multiple Endmember Spectral Mixture Analysis: Manaus, Brazil." *Remote Sensing of Environment* 106 (2): 253–267. doi:10.1016/j.rse.2006.09.005.

Qian, J., Zhou, Q., &Hou, Q. 2007. "{Comparison of Pixel-Based And Object-Oriented Classification Methods For Extracting Built-up Areas in Aridzone." In *ISPRS Workshop on Updating Geo-spatial Databases with Imagery & The 5th ISPRS Workshop on DMGISs* (pp. 163-171).

Registrar General I 2011. "Census of India 2011: Provisional Population Totals-India Data Sheet. Office of the Registrar General Census Commissioner, India."Indian Census Bureau.

Sameen, M. I., & Pradhan, B. (2016, June). A novel built-up spectral index developed by using multi-objective particle-swarm-optimization technique. *In IOP Conference Series: Earth and Environmental Science* (Vol. 37, No. 1, p. 012006). IOP Publishing.

Stone, B. 2007. "Urban and Rural Temperature Trends in Proximity to Large US Cities: 1951–2000." *International Journal of Climatology*, *27*(13), 1801-1807.

USGS. 2013. "Using the USGS Landsat 8 Product." Accessed November 7. https://landsat.usgs.gov/Landsat8_Using_Product.php.

Wang, Z., Gang, C., Li, X., Chen, Y., & Li, J. (2015). Application of a normalized difference impervious index (NDII) to extract urban impervious surface features based on Landsat TM images. *International Journal of Remote Sensing*, *36*(4), 1055-1069.

Weng, Q. 2001. "A Remote Sensing-GIS Evaluation of Urban Expansion and Its Impact on Surface Temperature in the Zhujiang Delta, China." *International Journal of Remote Sensing* 22 (10): 1999–2014.

Xu, H. 2008. "A New Index for Delineating Built-Up Land Features in Satellite Imagery." *International Journal of Remote Sensing* 29 (14): 4269–4276. doi:10.1080/01431160802039957.

Yuan, F., and M. E. Bauer. 2007. "Comparison of Impervious Surface Area and Normalized Difference Vegetation Index as Indicators of Surface Urban Heat Island Effects in Landsat Imagery." *Remote Sensing of Environment* 106 (3): 375–386. doi:10.1016/j.rse.2006.09.003.

Zha, Y., J. Gao, and S. Ni. 2003. "Use of Normalized Difference Built-Up Index in Automatically Mapping Urban Areas from TM Imagery." *International Journal of Remote Sensing* 24 (3): 583–594. doi:10.1080/01431160304987.

Zhai, K., Wu, X., Qin, Y., & Du, P. (2015). Comparison of surface water extraction performances of different classic water indices using OLI and TM imageries in different situations. *Geo-spatial Information Science*, *18*(1), 32-42.

Zhang, Q., G. Pavlic, W. Chen, R. Fraser, S. Leblanc, and J. Cihlar. 2005. "A Semi-Automatic Segmentation Procedure for Feature Extraction in Remotely Sensed Imagery." *Computers & Geosciences* 31 (3): 289–296. doi:10.1016/j.cageo.2004.10.003.

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

Fig. 1 The KMA in FCC of satellite image (2005).

Fig. 2 Spectral signatures of built-up area, bare land, water body and vegetation area in (a) optical bands 1-5 and 7 of Landsat-5 TM and (b) optical bands 2-7 of Landsat-8 OLI.

Fig. 3 Landsat-8 OLI in true colour composite (a) before resolution merging (30m spatial resolution) and (b) after resolution merging (15m spatial resolution).This illustrates the visual difference between non-enhanced and spatially enhanced image, at the location of Salt Lake City, Kolkata. The increase in sharpness and level of detail in the enhanced image obtained by merging original low resolution band images with high resolution image (15m spatial resolution) using a bilinear 4-nearest neighbour algorithm.

Fig. 4 Graphical flow chart for step wise elimination and extraction of water bodies, vegetation area, bare land and built up area from Landsat image.

Fig. 5 Diagrammatic representation of DFPS process for determination of threshold DN value that segments the NDVI image into vegetated and other areas. Gradual narrowing of search range and shortening of pace range increase success rate of the search. Iterative searching finally determines the threshold value to be 0.25 with the highest success rate of 96.26 for this example of thresholding NDVI image derived from Landsat-5 TM of 2000.

Fig. 6 (a-a′) The OLI image (bands 3, 4, 5 as red-green-blue) of 2015, (b) output of vegetation area segmentation from NDVI image and (b′) output of water bodies segmentation from band 6. Red rectangles in each panel show the sample areas where the DFPS approach applied to determine threshold values.

Fig. 7 Scatter plots showing the correlation between (a) band 4 (x-axis) and band 5 (y-axis) (b) band 4 (x-axis) and band 5 (y-axis) and (c) band 5 (x-axis) and band 6 (y-axis) for 2005; 2010 and 2015 respectively for (a-b-TM data) and (c-OLI data). The Pearson's correlation value (R) of (a) 0.951, 0.942 and 0.961; (b) 0.962, 0.951 and 0.976 and (c) 0.957, 0.964 and 0.941 for water bodies, vegetation areas and built-up areas respectively.

Fig. 8 Fitting of Logit models for separating built-up areas from bare lands in the images (water bodies and vegetation areas eliminated): 2000, 2005, 2010 and 2015. Bold numerics denote image bands involved in the best fit model with highest accuracy.

Fig. 9 Map of built-up areas in KMA derived from satellite images through application of SLEA for (a) 2000 (b) 2005 (c) 2010 and (d) 2015.

Fig. 10 Comparison of the results from the NDBI, BAEM and SLEA at sample locations, (a-a′) Landsat-5 TM and Lansat-8 OLI of 2010 and 2015, respectively; (b-b′) output from NDBI and BAEM respectively; (c-c′) output from SLEA for TM and OLI, respectively. Red A rectangles in each panel show the places where the NDBI and BAEM approaches incorrectly extracted other land covers as built-up areas and red B rectangles show the places where the NDBI and BAEM approaches underestimated the built-up areas while SLEA extracted built-up areas properly.

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

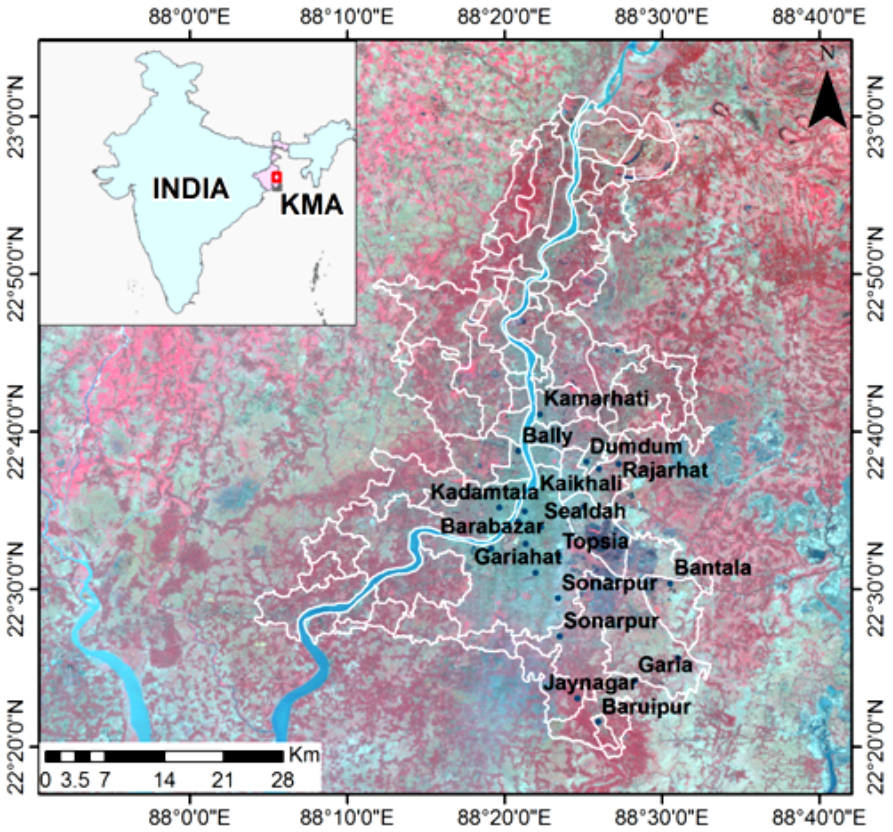

Fig. 1 The KMA in FCC of satellite image (2005).

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

Downloaded by [University of Florida] at 07:22 23 November 2017

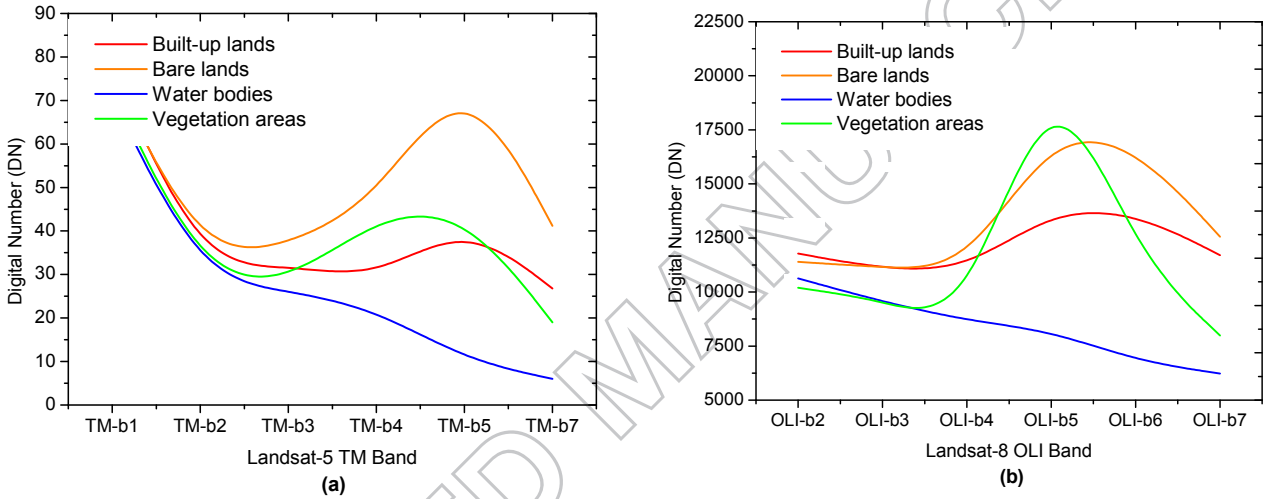

Fig. 2 Spectral signatures of built-up area, bare land, water body and vegetation area in (a) optical bands 1-5 and 7 of Landsat-5 TM and (b) optical bands 2-7 of Landsat-8 OLI.

ACCEPTED MANUSCRIPT

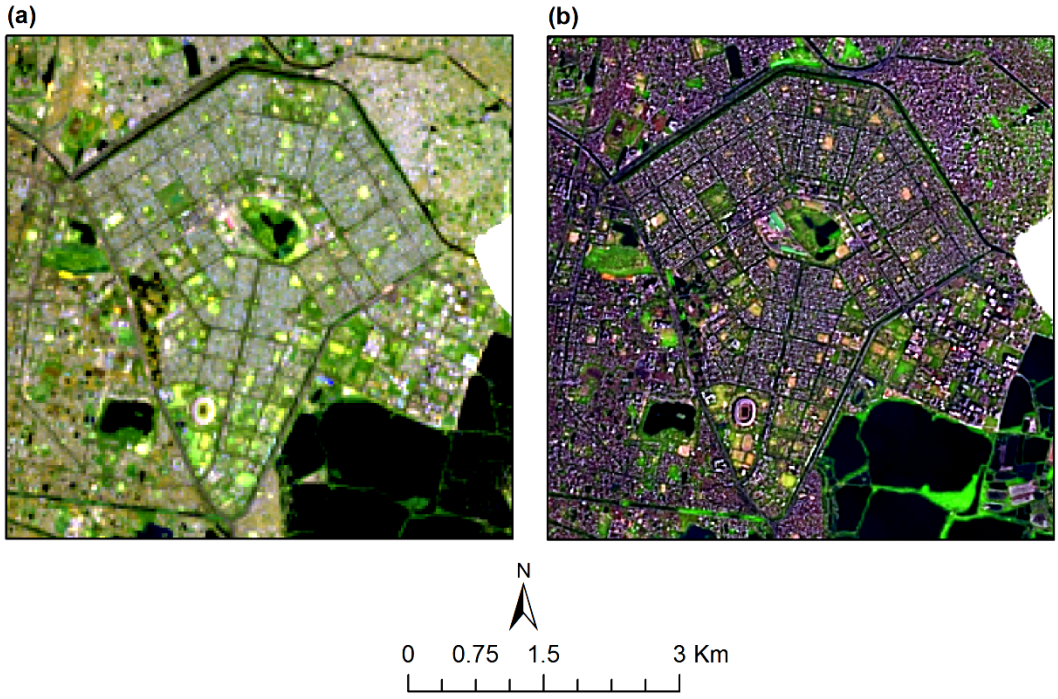

Fig. 3 Landsat-8 OLI in true colour composite (a) before resolution merging (30m spatial resolution) and (b) after resolution merging (15m spatial resolution).This illustrates the visual difference between non-enhanced and spatially enhanced image, at the location of Salt Lake City, Kolkata. The increase in sharpness and level of detail in the enhanced image obtained by merging original low resolution band images with high resolution image (15m spatial resolution) using a bilinear 4-nearest neighbour algorithm.

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED

Downloaded by [University of Florida] at 07:22 23 November 2017

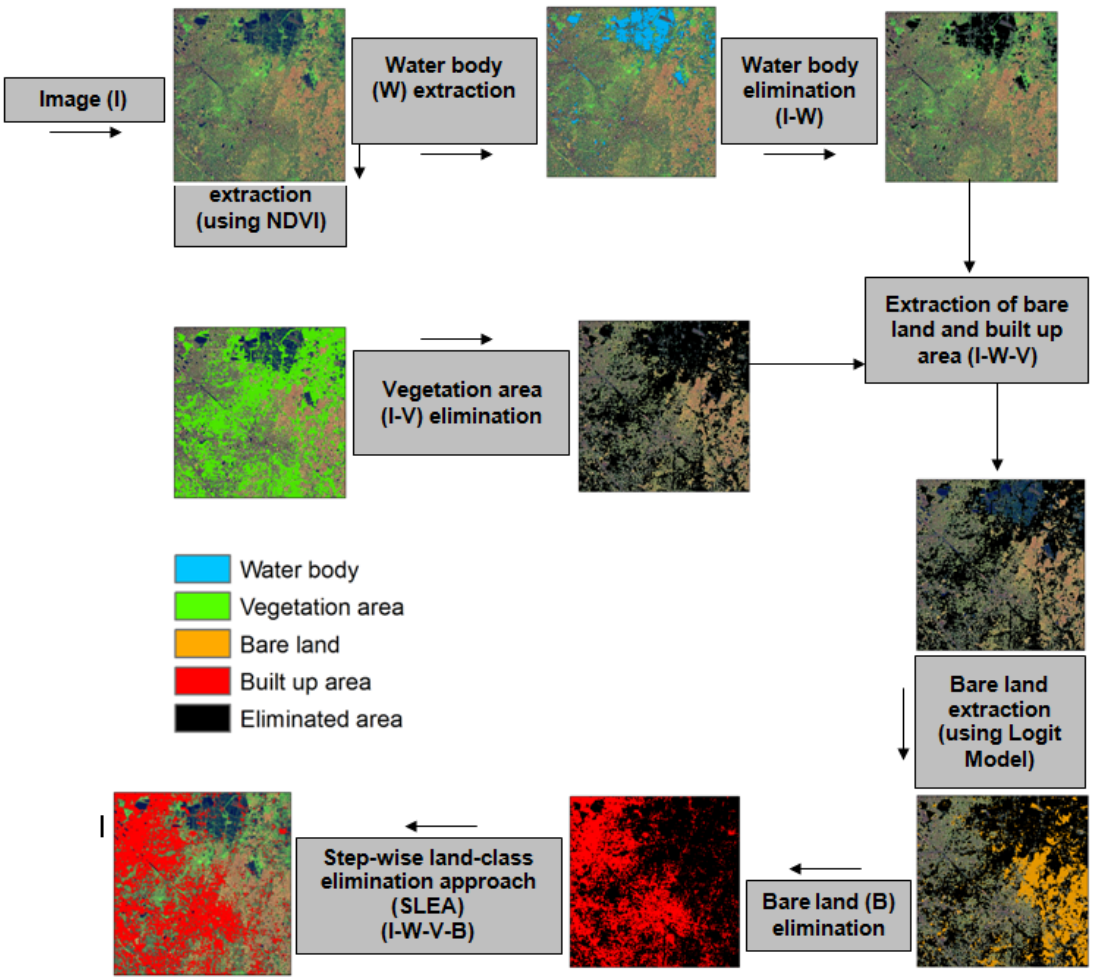

Fig. 4 Graphical flow chart for step wise elimination and extraction of water bodies, vegetation area, bare land and built up area from Landsat image.

Downloaded by [University of Florida] at 07:22 23 November 2017

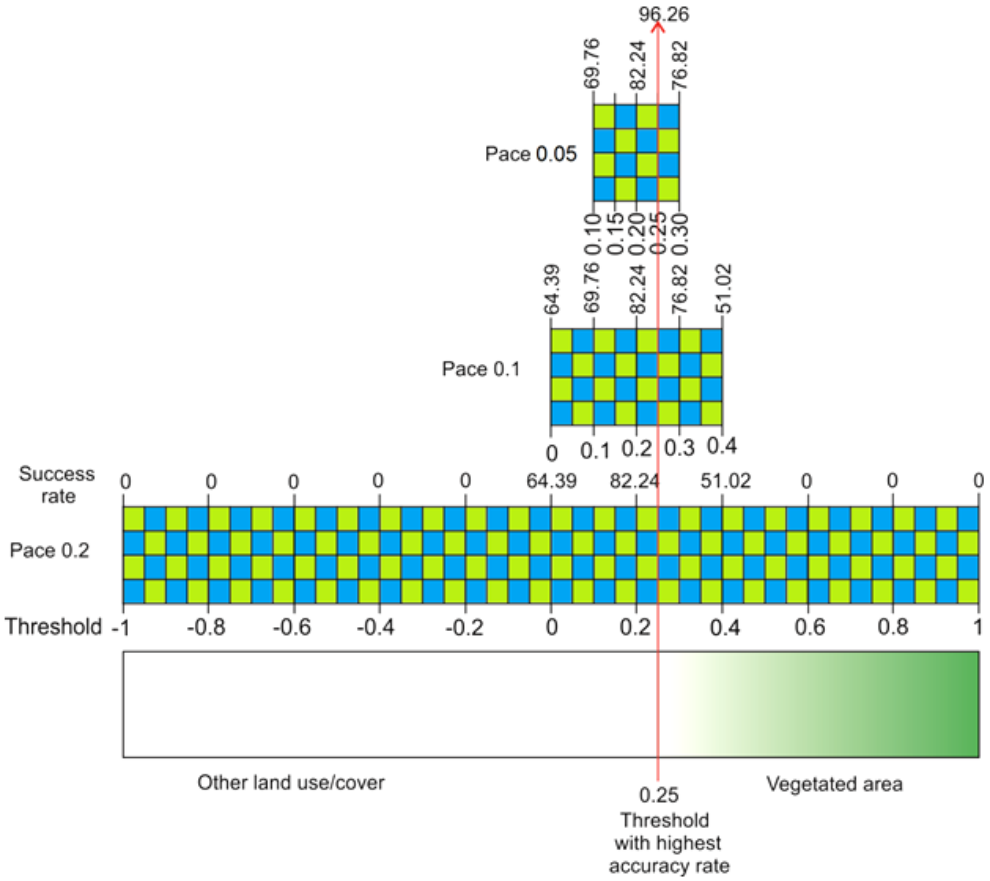

Fig. 5 Diagrammatic representation of DFPS process for determination of threshold DN value that segments the NDVI image into vegetated and other areas. Gradual narrowing of search range and shortening of pace range increase success rate of the search. Iterative searching finally determines the threshold value to be 0.25 with the highest success rate of 96.26 for this example of thresholding NDVI image derived from Landsat-5 TM of 2000.

Downloaded by [University of Florida] at 07:22 23 November 2017

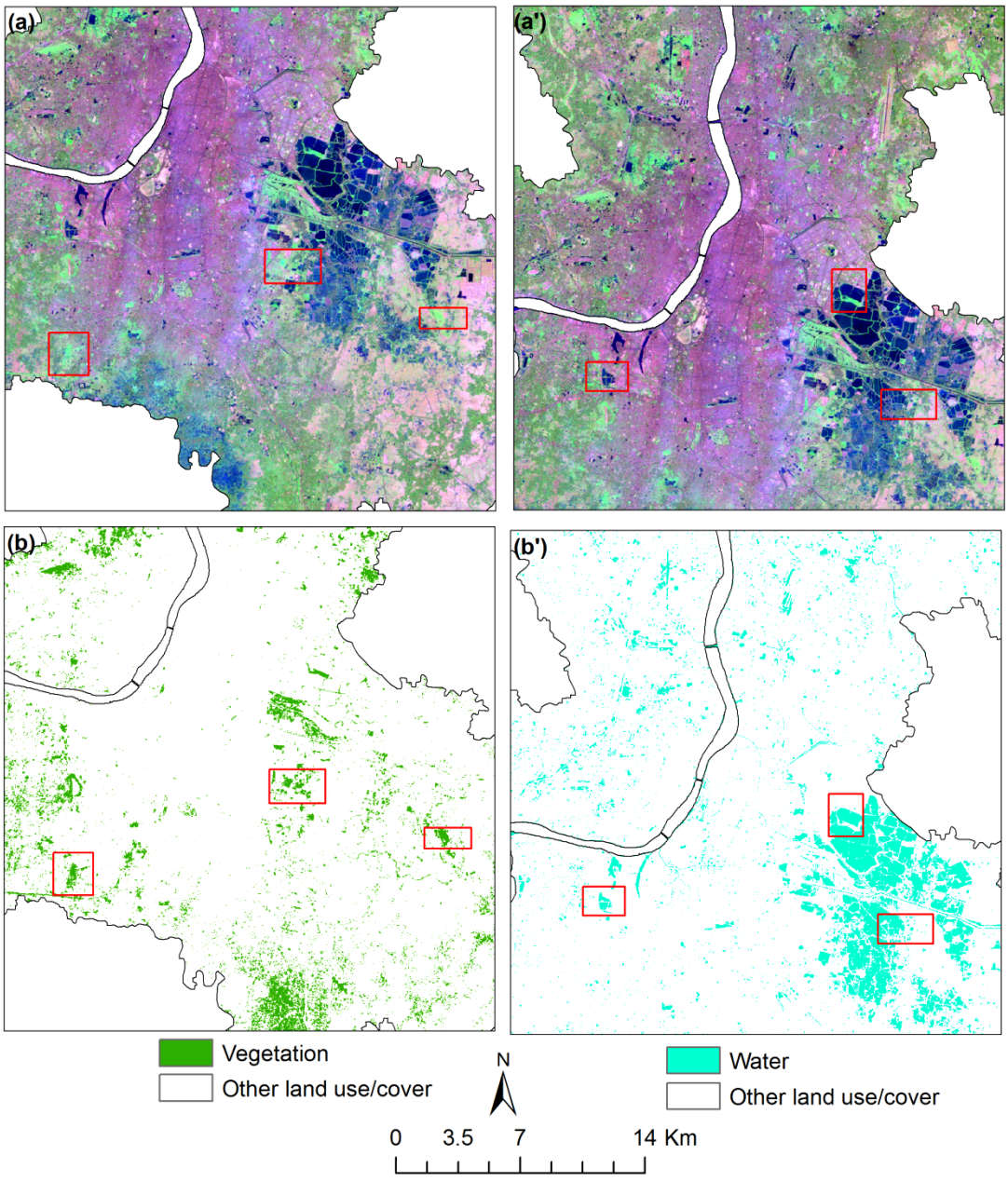

Fig. 6 (a-a´) The OLI image (bands 3, 4, 5 as red-green-blue) of 2015, (b) output of vegetation area segmentation from NDVI image and (b´) output of water bodies segmentation from band 6. Red rectangles in each panel show the sample areas where the DFPS approach applied to determine threshold values.

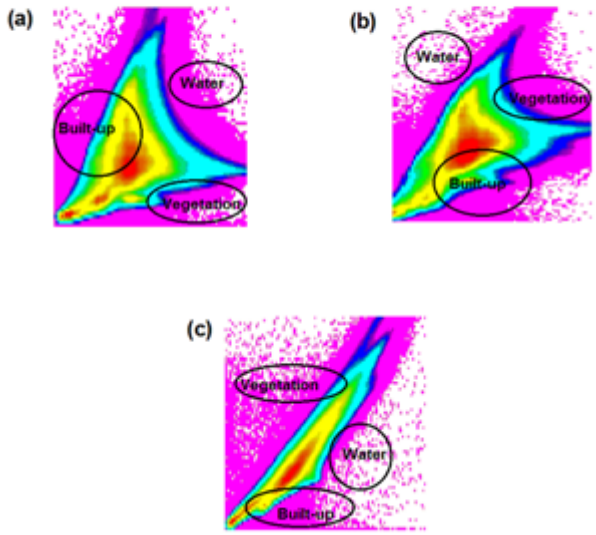

Fig. 7. Scatter plots showing the correlation between (a) band 4 (x-axis) and band 5 (y-axis) (b) band 4 (x-axis) and band 5 (y-axis) and (c) band 5 (x-axis) and band 6 (y-axis) for 2005; 2010 and 2015 respectively for (a-b-TM data) and (c-OLI data). The Pearson's correlation value (R) of (a) 0.951, 0.942 and 0.961; (b) 0.962, 0.951 and 0.976 and (c) 0.957, 0.964 and 0.941 for water bodies, vegetation areas and built-up areas respectively.

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

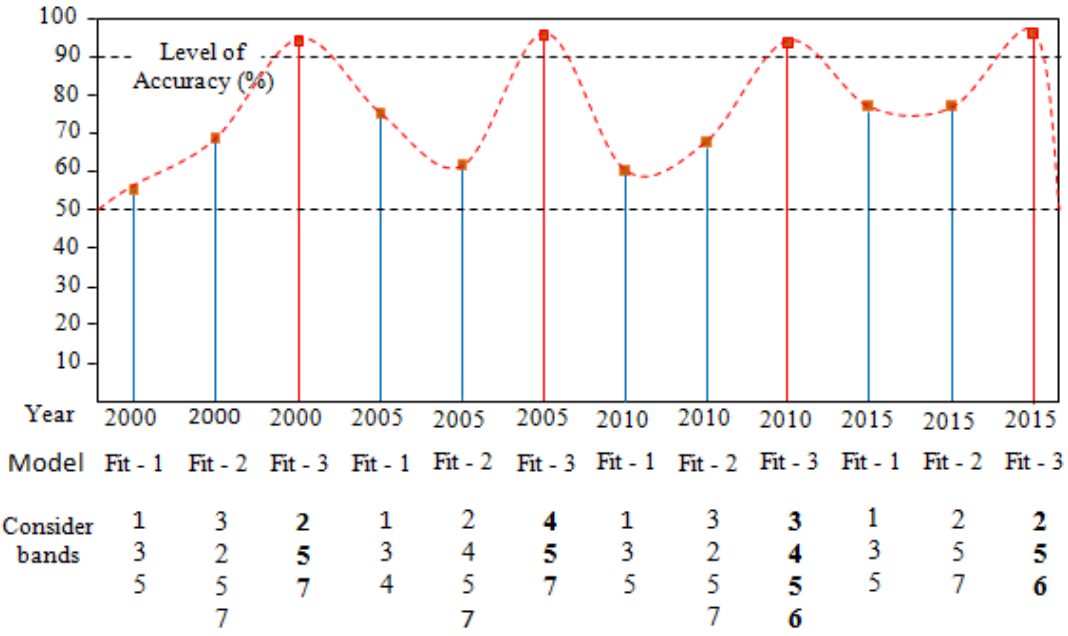

Fig. 8 Fitting of Logit models for separating built-up areas from bare lands in the images (water bodies and vegetation areas eliminated): 2000, 2005, 2010 and 2015. Bold numerics denote image bands involved in the best fit model with highest accuracy.

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED

Downloaded by [University of Florida] at 07:22 23 November 2017

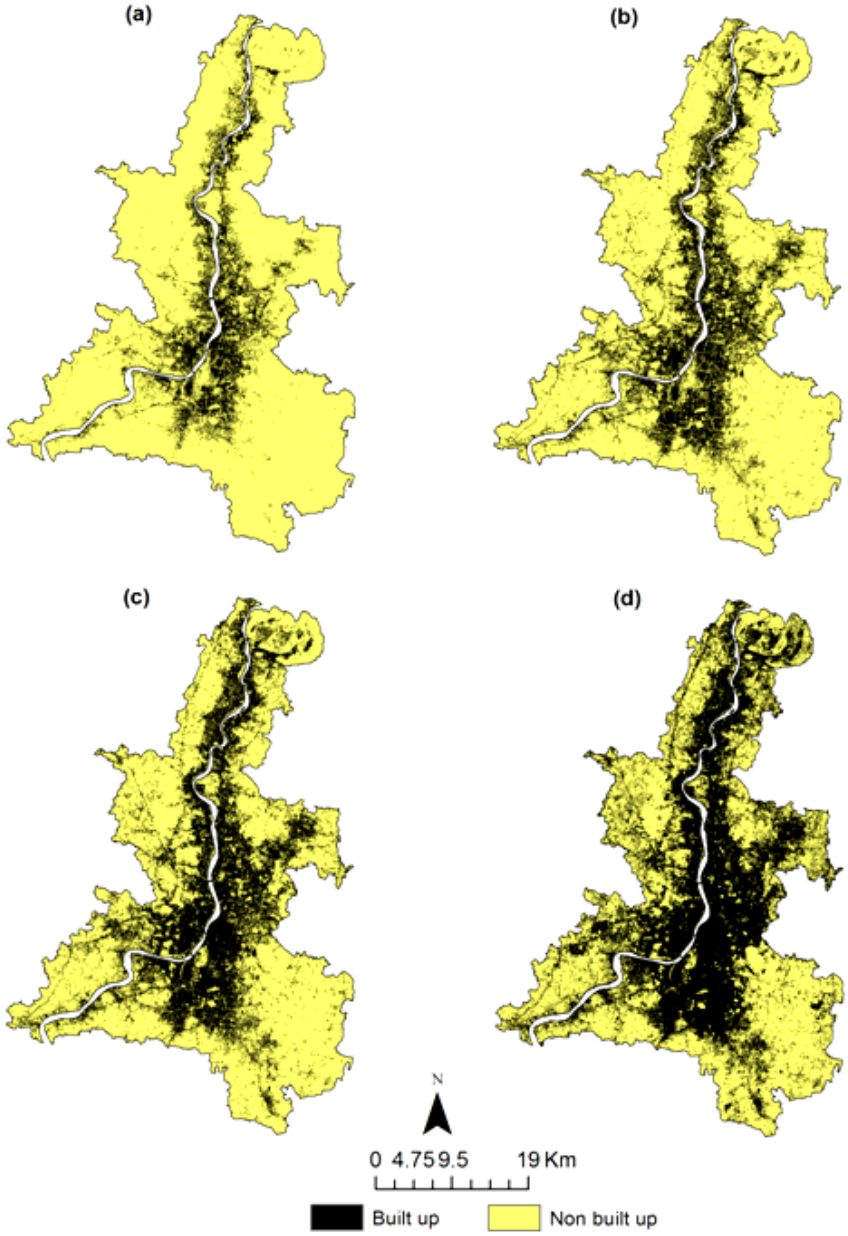

Fig. 9 Map of built-up areas in KMA derived from satellite images through application of SLEA for (a) 2000 (b) 2005 (c) 2010 and (d) 2015.

Downloaded by [University of Florida] at 07:22 23 November 2017

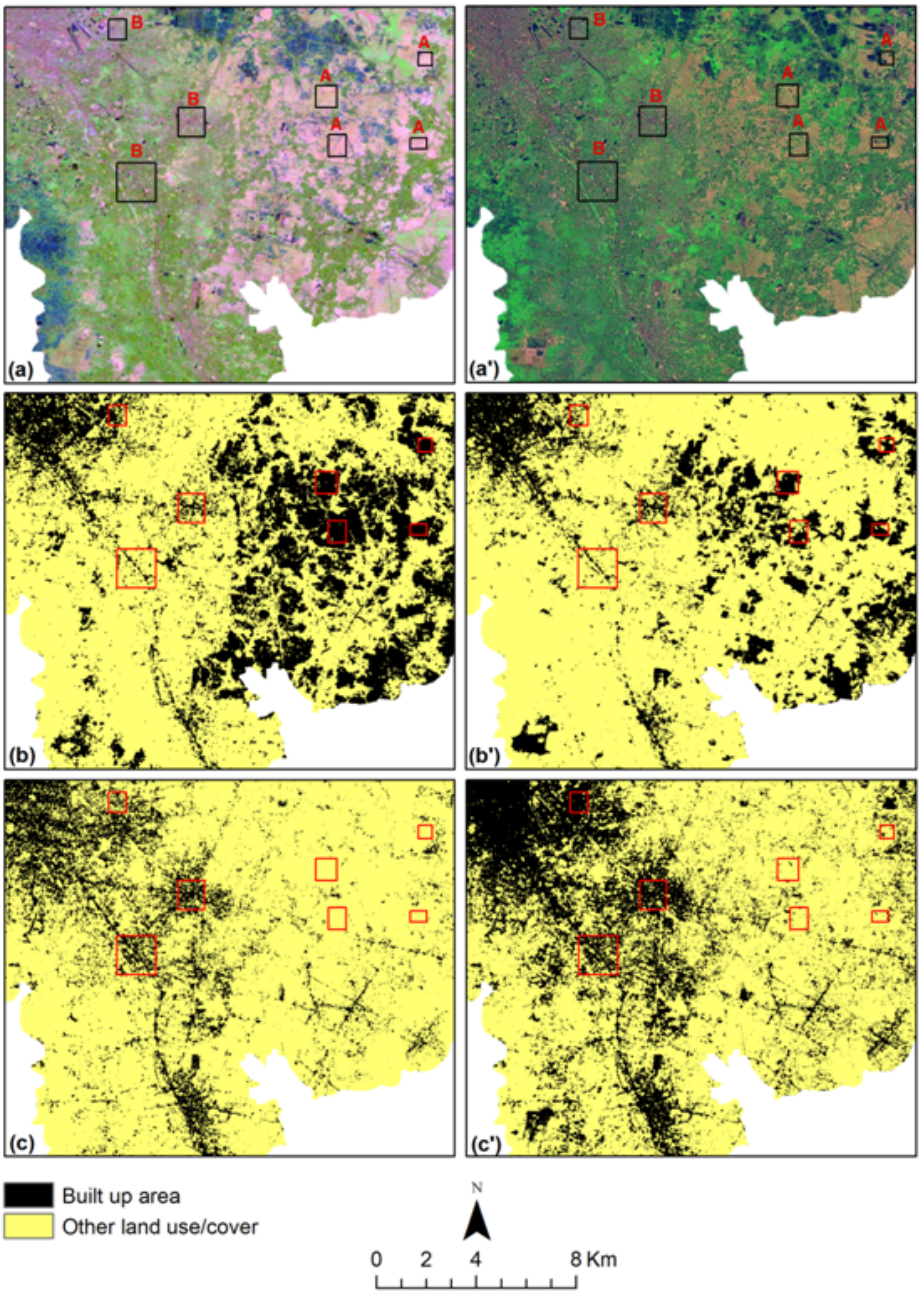

Fig. 10 Comparison of the results from the NDBI, BAEM and SLEA at sample locations, (a-a´) Landsat-5 TM and Lansat-8 OLI of 2010 and 2015, respectively; (b-b´) output from NDBI and BAEM respectively; (c-c´) output from SLEA for TM and OLI, respectively. Red A rectangles in each panel show the places where the NDBI and BAEM approaches incorrectly extracted other land covers as built-up areas and red **B** rectangles show the places where the NDBI and BAEM approaches underestimated the built-up areas while SLEA extracted built-up areas properly.

Downloaded by [University of Florida] at 07:22 23 November 2017

Table 1 Satellite data used in the study.

Satellite	Date of acquisition	Path and row	Referencing system
Landsat-5 TM	12 January 2000	138 & 44 and 138 & 45	UTM and WGS84
Landsat-5 TM	10 January 2005	138 & 44 and 138 & 45	UTM and WGS84
Landsat-5 TM	20 May 2010	138 &44 and 138 & 45	UTM and WGS84
Landsat-8OLI and	15 May 2015	138 & 44	UTM and WGS84

Table 2 An example of the DFPS for the determination of optimal threshold value of water pixels of the Landsat-8 OLI image of 2015

Range: 65535- 535 Pace 5000		Range: 15535 - 5535 Pace 1000		Range: 8535 - 7535 Pace 200		Range: 8535 - 8135 Pace 50		Range: 8435 - 8335 Pace 10	
Threshold	Success Rate	Threshold	Success Rate	Threshold	Success Rate	Threshold	Success Rate	Threshold	Success Rate
65535	0	15535	3.59	8535	58.79	8535	58.79	8435	82.35
60535	0	14535	5.38	8335	**78.48**	8485	72.65	8425	84.37
55535	0	13535	11.28	8135	76.57	8435	82.35	8415	85.58
50535	0	12535	16.27	7935	71.24	8385	**84.27**	8405	87.54
45535	0	11535	22.59	7735	60.45	8335	78.48	8395	89.69
40535	0	10535	26.25	7535	62.62	8285	79.56	8385	91.57
35535	0	9535	32.14			8235	79.56	8375	94.59
30535	0	8535	58.79			8185	78.92	8365	95.37
25535	0	7535	**62.62**			8135	76.57	**8355**	**96.35**
20535	3.59	6535	37.54					8345	92..38
15535	26.25	5535	29.24					8335	88.48
10535	**29.24**								
5535	0								
535	0								

ACCEPTED MANUSCRIPT

Table 3 Summary results of the DFPS for determination optimum threshold values in eliminating the water and vegetation pixels and final extraction of built-up areas.

Year	Target land class	First search			Second search			Third search			Fourth search			Fifth search			Optimum threshold
		Range	Pace	Maximum success rate	Range	Pace	Maximum success Rate	Range	Pace	Maximum success Rate	Range	Pace	Maximum success rate	Range	Pace	Maximum success rate	
2000	NDVI	1 - -1	0.20	72.36	0 – 0.40	0.10	82.93	0.10 – 0.30	0.05	96.26							0.25
	Water	255 - 5	50	16.25	105 - 5	20	57.28	45 - 25	5	82.42	45 - 35	1	95.21				37
	Built-up	1 - 0	0.10	95.83	0.05-0.35	0.05	98.21										0.45
2005	NDVI	1 - -1	0.20	69.52	0 - 0.40	0.10	86.65	0.10 – 0.30	0.05	95.83							0.25
	Water	255 - 5	50	10.24	105 - 5	20	48.67	45 - 25	5	87.54	40 - 35	1	94.52				37
	Built-up	1 - 0	0.10	94.68	0.50 - 0.35	0.05	97.83										0.45
2010	NDVI	1 - -1	0.20	72.36	0 – 0.40	0.10	88.35	0.10 – 0.30	0.05	97.53							0.25
	Water	255 - 5	50	22.34	105 - 5	20	54.54	45 – 25	5	82.57	40 - 35	1	96.24				37
	Built-up	1 - 0	0.10	96.28	0.05 - 0.35	0.05	97.85										0.45
2015	NDVI	1 - -1	0.20	46.25	0 – 0.4	0.10	46.25	0.10 – 0.30	0.05	68.24	0.20 – 0.25	0.01	97.56				0.23
	Water	60689 - 689	5000	29.24	1568 9 - 5689	100 0	62.62	8689 - 7689	200	78.48	8489 - 8289	25	84.27	8489 - 8439	5	96.35	8449
	Built-up	1 - 0	0.10	76.31	0.30 - 0	0.05	86.29	0.10 - 0	0.01	96.38							0.07

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

Table 4 Co-efficient values of Logit model for built up area extraction.

Time	Whole intercept value (α)	Selected band						
		band 1	band 2	band 3	band 4	band 5	band 6	band 7
		Intercept value (α)						
2000	10.80				-0.00	-0.35		0.34
	13.05		-0.08			-0.36		0.35
2005	8.73				-0.02	-0.25		0.21
	9.55				-0.16			-0.10
2010	39.06	0.41		-1.07		-0.48		0.92
	25.18			-0.76	-0.02	-0.43		
2015	29.06					-0.52	-0.27	
	27.21					-0.49	0.80	

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

Table 5 Confusion matrix for NDBI, BAEM and SLEA

Classified data	Reference data			
NDBI	Non-built-up area	Built-up area	Classified total	User's accuracy (%)
Non-built-up area	55	5	60	91.67
Built-up area	44	196	240	81.67
Reference total	99	201	**300**	
Producer's accuracy (%)	55.56	97.51		
Overall accuracy NDBI (%)	83.67			
BAEM	Non-built-up area	Built-up area	Classified total	User's accuracy (%)
Non-built-up area	101	3	104	97.11
Built-up area	29	167	196	85.20
Reference total	130	170	**300**	
Producer's accuracy (%)	77.69	98.24		
Overall accuracy BAEM (%)	89.33			
SLEA	Non-built-up area	Built-up area	Classified total	User's accuracy (%)
Non-built-up area	106	2	108	98.14
Built-up area	12	180	192	93.75
Reference total	118	182	**300**	
Producer's accuracy (%)	89.83	98.90		
Overall accuracy SLEA (%)	95.33			

Table 6 Summary of the accuracy assessment of the images derived by NDBI, BAEM and SLEA

Method	Overall accuracy (%)	Omission error (%)	Commission error (%)	k coefficient
NDBI	83.67	2.49	18.33	0.51
BAEM	89.33	1.76	14.80	0.62
SLEA	95.33	1.10	6.25	0.66

Downloaded by [University of Florida] at 07:22 23 November 2017

ACCEPTED MANUSCRIPT

View publication stats