Geocarto International

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

A new spectral index for the extraction of built-up land features from Landsat 8 satellite imagery

Rafik Bouhennache, Toufik Bouden, Abdmalik Taleb-Ahmed & Abbas Chaddad

To cite this article: Rafik Bouhennache, Toufik Bouden, Abdmalik Taleb-Ahmed & Abbas Chaddad (2018): A new spectral index for the extraction of built-up land features from Landsat 8 satellite imagery, Geocarto International, DOI: 10.1080/10106049.2018.1497094

To link to this article: https://doi.org/10.1080/10106049.2018.1497094

Accepted author version posted online: 09 Jul 2018.

A new spectral index for the extraction of built-up land features from Landsat 8 satellite imagery

Rafik Bouhennache[1], Toufik Bouden[2], Abdmalik TALEB-AHMED[3], Abbas Chaddad[4]

[1]*Science and technology department, Science and technology institute, university center of Mila, Algeria.*

[2]*Non Destructif Testing Laboratory (NDT Lab), Automatic Department, Sciences and Technology Faculty, Mohammed Seddik Ben Yahia University of Jijel, BP°98 Ouled Aissa, 18000, Jijel, Algeria.*

[3]*IEMN DOAE Le Mont Houy, university of Valenciennes, France.*

[4]*Dept. of Computer Science and Engineering, Blekinge Institute of Technology, 371 79 Karlskrona, Sweden*

Corresponding author: rafik.bouhennache@gmail.com

1. Introduction

Urban ecosystems are deeply influenced by many factors such as rapidity of expansion, loss of forests, loss of vegetation and global disasters. Following the Land Use/Land Cover (LU/LC) changes is a good strategy to manage the urbanization and to avoid the undesired situations and catastrophes like Tsunami. Remote sensing satellite images have been contributing to updating and actualizing maps, which are highly desired by the policy makers. Moreover, the LU/LC indices became a good tool to interpret, manage, follow and control land features such as Normalized Difference Vegetation Index (NDVI) which has been largely used to follow vegetation and Normalized Difference Built-up Index (NDBI) which has also been widely utilized to identify and map the built-up areas from medium spatial and spectral resolution satellite images (Stathakis et al. 2012; Zha et al. 2003). There are three categories of built-up extraction methods, the spectral and spatial indices, the combination of the spectral data and texture information i.e. classification, and the combination of sensors or multi-sensors (Zhang et al. 2014). However, the developed indices are outperformed other methods by their simplicity and rapidity of calculation, reduction of processing time and the high applicability. Since the creation of NDVI, researchers have made a huge effort to produce a similar accurate built-up index. The results show that the built-up lands are well separated from vegetation but they are poorly isolated from bare soil and water (Piyoosh & Ghosh 2018) because the calculated indices misclassify a quantity of barren and water regions as built-up areas due to heterogeneity of complex urban areas. Moreover, Kassawmar et al. (2018) have implemented a method to reduce this heterogeneity and improve the classification of LU/LC features. Kawamura et al. (1996) have proposed the Urban Index (UI) using TM7 and TM4 bands from Landsat Thematic Mapper (TM) sensor. Zha et al. (2003) have introduced the known built-up index, NDBI. To make NDBI more efficient at automatically mapping built-up terrain, they have subtracted the recoded NDVI image from

the recoded NDBI image. Earlier in 2008, Xu (2008) formulated a new expression of built-up index using the NDBI, the Soil Adjusted Vegetation Index (SAVI), the Modified Normalized Difference Water Index (MNDWI) and the introduction of the so-called Index-based Built-up Index (IBI). Deng and Wu (2012) proposed the Biophysical Composition Index (BCI). The BCI was shown to be the most effective index of the evaluated indices for separating impervious surfaces and bare soil. Bhatti and Tripathi (2014) have proposed the Built-up Area Extraction Method ($BAEM_{OLI}$) applied to the Operational Land Imager, and the Thermal Infrared Sensor (OLI-TIRS) Landsat 8 images and based on the subtraction of the sum of ($NDVI_{OLI}$) and ($MNDWI_{OLI}$) from the $NDBI_{OLI}$. The $NDBI_{OLI}$ has a specific expression using the Principal Component Analysis (PCA). Bouzekri et al. (2015) have used the green (G), red (R) and (SWIR1) bands of OLI Landsat 8 sensor and have suggested the Built-up Area Extraction Index (BAEI). Sinha et al. (2016) have proposed the New Built-Up Index (NBUI), which applies most of the wavelengths of Landsat images to represent the major urban land use classes. Piyoosh and Ghosh (2018) have proposed the so-called Normalized Ratio Urban Index (NRUI) and they have mentioned that using panchromatic (PAN) band (Band 8) of Landsat 8 data leads to an overall improvement in discriminating between built-up, barren (bare soil) and vegetation. However, generating a satisfactory built-up index image from remotely sensed data like Landsat 8 is not a straightforward task. There are many factors that may reduce the accuracy of classification, such as the nature of the study area, the spatial and spectral resolution of satellite remotely acquired data. As a matter of fact the built-up land feature is smaller than the spatial resolution of sensors, besides the multiple equations and transformations which may amplify errors. It is a hard and a complex process started from the registration and preprocessing operation until the generation of the accuracy map. The purpose of this study is to develop a new simple accuracy built-up land features extraction index (BLFEI) which robustly differentiates the built-up areas from the surrounding barren,

vegetation and water surfaces, and takes advantage of the quick processing time and absence of human intervention or interaction. This new index is applied to Algiers image taken from Landsat 8 OLI sensor, but remains available for Enhanced Thematic Mapper Plus (ETM+) sensor. The proposed index is compared to some recently developed indices where it showed higher performance in terms of both separability as well as accuracy.

2. Related works and the existing indices

In remote sensing an index is a mathematical spectral transformation formula of two or more bands that have the ability to highlight the desired land feature and to graphically indicate in a uniform toned color the pixels which have the similarity of spectral value in a small range. The creation of an index is based on the unique pattern of each land cover and the spectral response of signature features.

Researchers in the field of remote sensing have waited until 2003 when Zha et al. (2003) announced their index (NDBI). This index is similar to NDVI for its segmentation of built-up areas but achieves lower accuracy.

The NDBI index uses the difference and the ratio of Middle InfraRed Band (MIR) or (B5) and Near InfraRed band (NIR) or (B4) to highlight the built-up areas and it is given by the following equation:

$$\mathrm{NDBI} = \frac{(B_5 - B_4)}{(B_5 + B_4)} \quad (1)$$

This index takes advantage of its simplicity and its computation speed.

An alternative way to extract more precisely built-up areas and to eliminate the noise of vegetation and water is the IBI index proposed by (Xu 2008), which is applied to ETM+ sensor and is given by the following equation:

$$\mathrm{IBI} = \frac{(\mathrm{NDBI} - ((\mathrm{SAVI} + \mathrm{MNDWI})/2)}{(\mathrm{NDBI} + ((\mathrm{SAVI} + \mathrm{MNDWI})/2)} \quad (2)$$

Where SAVI and MNDWI are respectively expressed as:

$$SAVI = \frac{(B_4 - B_3)(1+l)}{(B_4 + B_3 + l)} \quad (3)$$

$$MNDWI = \frac{(B_2 - B_5)}{(B_2 + B_5)} \quad (4)$$

B2, B3 are respectively the green and red bands of the ETM+ sensor and l is a factor implemented to minimize the vegetation index sensitivity to soil background reflectance variation. If l is zero, SAVI becomes the same as NDVI. For intermediate vegetation cover ranges, l is typically set around 0.5 as reported in the work of (Zhang et al. 2009).

Let us now delve into the most recent indices reported in the literature. In the passed few years, researchers have made a considerable effort to establish a good index that reflects the reality of built-up regions (Piyoosh & Ghosh 2018; Sinha et al. 2016; Li et al. 2015; Bouzekri et al. 2015; Estoque & Murayama 2015; Bhatti & Tripathi 2014; Stathakis et al. 2012; Deng & Wu 2012). Some of them are based on the Tasseled Cap Transformation (TCT) and PCA. The first is BCI, an index developed with Landsat ETM+, IKONOS and MODIS satellites and similar to IBI. The BCI uses the three of first TCT components and it is computed as:

$$BCI = \frac{(\frac{(H+V)}{2} - L)}{(\frac{(H+V)}{2} + L)} \quad (5)$$

Where H, V and L are the brightness (TC1), the wetness (TC3) and the greenness (TC2) components of the tasseled Cap transformation, respectively.

Estoque and Murayama (2015) have proposed the Visible green-based built-up index (VgNIR-BI) which is a simple and accurate index, and is applied for Landsat 7 as well as Landsat 8. The expression of this index is given by:

$$VgNIR - BI = \frac{(\rho_{Green} - \rho_{NIR})}{(\rho_{Green} + \rho_{NIR})} \quad (6)$$

Where ρ_{Green}, ρ_{NIR} are respectively the reflectance of the bands (B2), (B4) for the ETM + sensor and the reflectance of the bands (OLI3), (OLI5) for the OLI sensor.

The BAEI index is derived from band ratios and applied to Landsat 8, its formula is:

$$BAEI = \frac{OLI_4 + 0.3}{OLI_3 + OLI_6} \quad (7)$$

Another approach is NBUI, an index applied to Landsat 5 and based on the subtraction of the SAVI and MNDWI from the Enhanced Built-up and Bareness Index (EBBI) (As-Syakur et al. 2012) and it is computed as:

$$NBUI = \frac{B_5 - B_4}{10 \times \sqrt{B_5 + B_6}} - \left(\frac{(B_4 - B_3)(1+l)}{(B_4 + B_3 + l)} + \frac{(B_2 - B_5)}{(B_2 + B_5)}\right) \quad (8)$$

To extract built-up area from Landsat 8 imagery through NBUI index, the thermal band (B6) is replaced with (OLI10) thermal band. The formula of NBUI and many other indices illustrate the importance for the utility of SWIR1 band in the creation of indices.

The indices not mentioned in this section are listed in Table 1 by citing the authors as well as the satellites, the sensors used, the Overall Accuracy (OA) and the kappa coefficient (k) if are given by authors.

3. Study area, data sets and preprocessing

3.1 Study area and data used

The OLI sensor Landsat 8 satellite image (L8-19635) of level 1 acquired on 1/5/2015 corresponds to the path of 196 and the row 35 georeferenced to UTM WGS 84 zone 31 was subset by a shapefile of Algiers located in the middle North Africa as shown in Figure 1. A high resolution image of Google earth captured on the same day and at the same location, besides, a Landsat 8 image of another scene footprint corresponding to 196 path and 34 row (L8-19634) for Level 1 acquired on the same day (1/5/2015) but eight seconds earlier from

than the first image and georeferenced to the same reference (UTM WGS 84 zone 31), are both used for results validation.

3.2 Data preprocessing

To enhance the satellite images and increase the classification accuracy, a preprocessing process is required. This process is applied to image L8-19635 as well as to L8-19634. To improve the spatial resolution and achieve a resolution of 14.25 m. First a resampling method (change of resolution) using the nearest neighbor algorithm was performed to obtain a resolution of 28.5 m (Tucker et al. 2004; Ehlers & Welch 1987), subsequently followed by a pansharpening where the OLI Bands 2-7, 10 used have a resolution of 28.5 m, and thus were merged with Band 8, that has the high resolution of 14.25 m using Gram-Schmidt pansharpening (Xu et al. 2014; Ehlers et al.2010). Pansharpening is an image fusion technique in which high resolution panchromatic data is combined with lower resolution multispectral data to obtain a colorized high-resolution dataset. An atmospheric correction has been applied to remove the influence of atmospheric scattering (Zhou et al. 2014). The Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FlAASH) and the ATmospheric CORection (ATCOR) modules are used. The Landsat 8 OLI sensor is very sensitive, the digital data is rescaled to 16-bit DN and ranges from 0 to 65536 as shown in Table 2. To extract built-up area from Landsat 8 imagery through the cited indices and the proposed index, these images have been converted to reflectance rather than radiance. A radiometric calibration is available in ENVI software (Environment for Visualizing Images) to calculate the Top-Of-Atmosphere (TOA) radiance used as input, in the FLAASH module. Fuyi et al. (2013) concluded that for ground reflectance the most accurate results are obtained with ATCOR. However, Morteza et al. (2015) announced that FLAASH atmospheric correction outperformed ATCOR in the majority of cases. In our case, the FLAASH and ATCOR reflectance's values are almost the same as shown in Figure 2. Figure 2 also exhibits that even the separation of built-up from the

G and R bands is good, the separation from the SWIR1 and SWIR2 even bands is better. Displayed color composites were formed using various band combinations to differentiate diverse types of land cover such as 765, 543, and 654 (false color) composites. By using the 654 composite, the built-up areas appear in varying shades of magenta, vegetation is green bright, barren (soil) is mauve and water is very dark (Figure4), and by using the 543 composite, vegetation appears in shades of red, built-up areas are cyan blue, barren land vary from dark to light browns, and water appears very dark as shown in Figure 1.

4 Evaluation of some existing indices

In this section, we briefly evaluate the built-up indices cited in section 2 (without NDBI, SAVI and MNDWI) using the histograms overlap method and the spectral discrimination index (SDI) technique (Piyoosh & Ghosh 2018; Sun et al. 2016; Deng & Wu 2012). To show the images of color coded indices instead of grayscale, a classification using Support Vector Machine (SVM) was carried out. SVM is well known in the field of classification for remote sensing and leads to better results (Feyisa et al. 2016; Hazini & Hashim 2015). The Region Of Interest (ROI) are used as a spectral signature of land use and land cover categories namely built-up, barren, vegetation and water. To perform the SVM classification, the ROI are carefully selected and constructed until they exhibit satisfactory separability and they are shown in Figure 8. The classified indices are shown in Figure 3. The detailed evaluation and discussion also using Otsu's method and accuracy assessment will be discussed in the results and discussion section.

Assessment is based on visual interpretation and analysis of overestimation, underestimation or separability of each class. It is mentioned that all indices can segment the built-up regions and there is no index that we can single out which can efficiently extract the built-up lands. The accuracy of the extracted built-up areas through the indices varies from one index to another. To analyze the overlap between classes, we use the SDI based on the

means and Standard Deviations (SD) of the classes and is defined by (Kaufman & Remer 1994) as the difference of the mean values of two classes divided by the sum of their standard deviations. To avoid negative values, the difference in means is replaced by the absolute difference. If $SDI < 1$ classes overlap and the ability to discriminate the classes is poor, whereas if $1 \leq SDI < 3$ Histogram means are well (good) separated and that regions are relatively easy to discriminate. However, if $SDI \geq 3$ an excellent discrimination of land features is reached and there is no overlap that occurs. Table 3 shows the SDI values for the four features of landscape, namely, the built-up, bare soil, vegetation and water and Figure 4 shows the overlapping histograms between land covers. With SDI values of 3.05 and 4 between built-up-vegetation and built-up-water respectively, IBI index has better separated the vegetation and water from the built-up areas; however, it suffer from discriminating the built-up regions from the barren land which is reflected in Figure 4 by isolation of histograms between built-up-vegetation, built-up-water and deep intersection of histograms between built-up and barren, and it is reflected in Figure 3 for SVM classification by underestimation of the barren class, but it is closer to reality as shown in Figure 3. The new tested indices from 2012 until 2016 (i.e. BCI, VgNIR-BI, BAEI and NBUI) can extract and map the built-up more precisely, but the disadvantage is their enormous value, creating a problem of coding grayscale images. The BCI index has the ability to better segment built-up lands and separate them from other classes, but it confuses a small amount of barren land with built-up regions, in addition there is low isolation between barren and vegetation. BAEI index has perfectly separate water from barren moreover, it confused barren with vegetation and a small quantity of barren with built-up. VgNIR-BI index is still a good accurate index, but it also overestimates urbanized areas in arid lands as shown in Figure 4, the disadvantage being that it only exploits two bands from the entire electromagnetic spectrum and remains to be tested for a complex urban system. The NBUI index has the highest SDI value between barren and

vegetation (2.93) and remains the best index for extracting both barren and built-up. The NBUI method has largely merged urbanized zones with barren land; however, it has largely separated the vegetative area from barren class.

5. Methodology

5.1. Creation of the spectral index BLFEI

To develop the new index, the LU/LC profiles shown in Figure 2 are analyzed to determine the unique pattern of the land features. The concept is to determine the strongest and the weakest values of the built-up reflectance. It is obvious that for OLI7 and OLI6 bands, the built-up areas are well spectrally distinguishable from the others LU/LC. The most useful bands from which some cover lands can be potentially differentiated and separated are OLI_3 (G), OLI_4 (R), OLI_6 (SWIR1) and OLI_7 (SWIR2) bands. In the electromagnetic spectrum the built-up areas have high values reflectance for SWIR1 (1.60 µm), SWIR2 (2.20 µm) and low values for green (0.56 µm) and red (0.65 µm), the vegetation and barren have high values in the OLI_5 (0.86 µm). Figure 2 shows that barren land and asphalt like roads have an almost equal spectral response for the spectrum ranging from NIR to SWIR1; these responses have roughly intersected for the NIR band. This is why the NIR band does not appear in the BLFEI formula given by the following equation:

$$\mathrm{BLFEI} = \frac{\frac{(\mathrm{OLI3}+\mathrm{OLI4}+\mathrm{OLI7})}{3} - \mathrm{OLI6}}{\frac{(\mathrm{OLI3}+\mathrm{OLI4}+\mathrm{OLI7})}{3} + \mathrm{OLI6}} \qquad (9)$$

By using equation (9) the water areas will have the highest values and appear in the white tones. The built-up areas' features will have the medium values and appear in bright grey. The vegetation will have the lowest values and appears in black and dark grey as shown in Figure 5. The barren will appear in grey tones because its values are greater than the vegetation and less than the values of the built-up areas. The robustness of BLFEI index is due to the fact that it exploits almost the whole spectrum of visible (Vis) and infra red

shortwave (SWIR). It is based on the R and G bands from the visible spectrum and on the SWIR1, SWIR2 bands from the infra red spectrum. Figure 5 shows the grayscale image of the BLFEI index and its SVM classification. It is obvious that the new index provides striking spatial details like roads and airports better than almost all built-up indices in comparison with SVM classification of the others indices as shown Figure 3.

5.2. Optimal thresholding and Otsu's algorithm for separating built-up from non built-up regions

Most of the previous studies have used the manual thresholding to generate built-up and non built-up binary images (Garg 2016; Patel & Mukherjee 2015; Xu 2007; Zha et al. 2003). An automated thresholding algorithm is swiftly needed to classify the built-up index image into built-up and non built-up regions (Varshney & Rajesh 2014). Otsu's method for finding the threshold of an image is based on minimizing the interclass variance of two classes representing the background and foreground of the image (Otsu 1979). It selects an optimal threshold by maximizing between-class variance in an image characterized by a bimodal histogram. Liu et al. (2010) have conducted an experimental research of different methods of threshold segmentation and have concluded that Otsu's method outperforms others techniques by speed of processing and its stable effect and have reported also that Otsu's thresholding algorithm gives better results even when more than the two peaks are present in the histogram. Li et al. (2013) have used the Otsu's threshold to automatically calculate the threshold value for separating the water feature from other lands. If we assume that a gray level histogram corresponds to an image obviously has two modes (sides), one for built-up areas (class *C1* ranging from [$\alpha, \ldots, t$]) and the other for non built-up regions (vegetation and barren) (Class *C2* ranging from [$t, \ldots, \beta$]), where t is the threshold value. We note Mb, Mn the means of built-up and non built-up classes respectively and Pb, Pn the probabilities that a

pixel belonged to class *C1* or class *C2* respectively. Therefore, the between class variance is given by variance σ and it is expressed as:

$$\sigma = \sqrt{P_b.(M - M_b)^2 + P_n.(M - M_n)^2} \quad (10)$$

Where *M* is the mean of entire image $M = P_b .M_b + P_n .M_n$ and $P_b + P_n = 1$

The Otsu optimal threshold T_h is given by

$$T_h = ArgMax_{\alpha \le t \le \beta}(\sigma^2) \quad (11)$$

The optimal threshold value's algorithm is implemented in the Environment of ENVI 5.3 and with a script of IDL 8.5 (Integrated Development Language). To avoid the enhancement of water with built-up when applying the new index, an operation of masking out the water still needed before applying the index as reported by (Waqar et al. 2012). After the water regions are masked, the histogram presents two modes allowing to find automatically the optimal threshold used for binary coding (built-up and non built-up) and built-up land extraction as shown in Figure 7. Any pixel of image is notified belonged to built-up class if its value is greater than the optimal threshold value (Th). Table 4 shows the means, the SD and the optimal threshold values and Figure 9 shows the extracted built-up from the waterless surfaces (barren and vegetation, the water is shown but is masked out).

5.3. Comparison with the existing indices

To demonstrate the ability of BLFEI index to extract the built-up land features, this study compares the new index with the evaluated indices described in the section 3.3. (BAEI, BCI, IBI, NBUI and VgNIR-BI). The comparison is based on the evaluation using histogram overlap and spectral discrimination index, the results from Otsu's optimal thresholding or how each index responds to Otsu's method and accuracy assessment as well as the comparison of built-up area surfaces extracted using the SVM classification and the Otsu's method. SVM output classification images of these spectral indices are shown in Figure 3. Analysis, quantitative and qualitative comparison will be discussed in the results and discussion section.

5.4 Built-up areas extracted and accuracy assessment

To evaluate the accuracy of the cited indices in the section 3.3 and the new index for their ability to segment built-up lands, 80 pixels belonging to built-up area and 120 pixels belonging to non built-up region (vegetation and barren) were randomly collected using stratified random sampling method according to Tumb rule (Xu 2007). These 200 random points were generated in ArcMap 10.4.1 and converted to KML in order to use in Google Earth. The 200 points are shown in Figure 8 and are generate in each index image to sort them according to the class they belong to (built-up or non built-up). After those pixels are examined one at a time in the 654 composite of the multispectral image (L8-19634) with the resolution of 14.25 m, they are verified using the Google Earth of high image resolution. Google Earth provides high resolution images with the date were taken, without metadata and without the time-stamp, so we used L8-19634 image which is eight seconds earlier than L8-19635 image to ensure that no change will occur in the studied area. Table 5 summarizes the error matrix (confusion matrix) including the Overall accuracy (OA), the kappa coefficient (k) as well as the commission error and omission error and Table 6 exhibits the scale used to describes the strength of agreement for kappa statistic (Rwanga & Ndambuki 2017). Figure 10 shows the OA and k values and exhibits that OA values range from high (95%) to low (81%). The highest value is attributed to the BLFEI index and the low value is for NBUI. In addition, the value of k varies from 0.60 to 0.90 and is considered substantial (0.61-0.80) for all indices except for BLFEI which is considered almost perfect (0.81-1.00).

6. Results and discussion

6.1 Results from SDI separability and overlapping histograms

Table 3, Figure 4 and Figure 6 exhibit that BLFEI, the new index discriminates the four land cover categories more precisely and its separability is higher than indices used for comparison. Figure 2 shows the spectral response of the terrestrial features after being

corrected atmospherically with FLAASH and ATCOR and pansharpened with PAN band leading to a spatial and spectral enhancement. Figure 2 also exhibits the similarity of the spectral response between barren and asphalt, particularly around the NIR spectrum, so that by using the BLFEI formula, the built-up area was highlighted in foreground. Figure 4 and Figure 6 show the separability between land features or the overlapping phenomena and illustrate that for the new approach, the classes are well discriminated but for the rest of indices they are overlapped and the barren land is spectrally confused with the built-up regions at different proportions, which also is reflects by intersection of the land cover histograms. SDI values and the intersection length of histograms determine the proportion of confusion between land covers. SDI values for BLFEI are ranged between 1.75 and 7.44, and are rated as good and excellent by considering SDI's decision. SDI of barren - vegetation, and barren - water is perfect (excellent) because its value is respectively 3.72 and 4.81. Moreover, BLFEI and the studied indices have excellent separability between built-up and vegetation on the one hand and between vegetation and water on the other hand, as shown in Table 3. In addition, BLFEI index presents a perfect isolation between land cover histograms except between built-up and barren where a negligible overlap has occurred. This overlapping behavior is undesired for a precise extraction of lands. Table 3 shows that the highest value of SDI between built-up land and barren class is 2.10 and it assigned to BLFEI index, subsequently followed by VgNIR-BI with a value of 1.69 which is 62% lower than our proposed index and makes BLFEI the new index improves separability by 25%. After that comes the SDI of BAEI index with a value of 1.25 which is 84% inferior, thereafter comes the SDI of BCI with a value of 1.08 which is 97% less than SDI value of BLFEI index. This fact is also supported by Figure 6 which shows the means of the four urban components features and demonstrates that almost all built-up indices suffer from separating the barren lands and it is only the BLFEI index which is greatly discriminate between them followed by the VgNIR-

BI index. With an SDI value of 3.72 which is the greatest, BLFEI is the best to make great discrimination between built-up and vegetation. The largest value of SDI between barren and vegetation is 2.95, which makes NBUI the effective index for discrimination between barren and vegetation. The poor separability for the rest of the indices is a disadvantage because of, essentially the lack of insufficient information that improves the separation such as the texture features and due also to the existence of the shadow in the studied area mainly for the building and roads (Labib & Harris 2018).

6.2 Results from the extraction of built-up areas using Otsu's optimal threshold

The water reflectance values in G and R bands are higher than those in SWIR1 band as shown in Figure 2; therefore by processing BLFEI index, the water returns the highest values (positive values as shown in Figure 5) and will be enhanced like the built-up lands. This problem mostly occurs with BLFEI, BAEI and VgNIR-BI indices and the necessity to mask out the water before executing the formula of the said indices and processing the optimal thresholding is essential, but there is an alternative solution of water masking by using the double manual threshold. Table 4 and Figure 10 show that the optimal Otsu thresholding values are lower than the mean values; these values are displayed on the histogram of each index as shown in Figure 7. The classes and their overlaps in case they exist are presented on the histograms by different colors, as shown in Figure 7. The Otsu's method works very well with the BLFEI index where the Otsu's optimal threshloding value is detected in the middle of a two predominant symmetrical sides (peaks) of the histogram corresponding to built-up area, on the one hand, and to barren and vegetation areas on the other hand. Followed by the BCI index which also works well with the Otsu's method according to its histogram which presents two peaks, one for the built-up class and the other for the two classes mentioned above (note that the water zones are masked out). Then, the BAEI and IBI indices moderately responded to the Otsu's method for optimal thresholding. Based on the optimal thresholding

values, the output binary indices are coded as built-up and non built-up regions. Table 4 and Figure 7 show the Otsu's optimal thresholding values and Figure 9 shows the binary coding image where the built-up lands are finally extracted from the waterless lands. To visually evaluate the new proposed index and the indices of investigation, four locations are chosen objectively: first the bay of Algiers, the second is the park of Alhamma (a region in Algiers), the third is Algiers airport, and some roads as shown in Figure 9. Bay of Algiers is perfectly extracted by the built-up indices. The park which is located in the middle of a potentially urbanized city is also discriminated by the new method (BLFEI) as well as by BCI, IBI, NBUI indices, but it is extracted with a moderate precision for the VgNIR-BI index and it is confusedly segmented with built-up and barren lands according to the BAEI index. For the airport, BLFEI outperforms for the extraction of the airport's headquarters, however the indices developed like BCI, IBI and NBUI segment more precisely the airport runways than the BLFEI index. With the size of the images shown in Figure 9 (421×305), some roads are only extracted by the BLFEI approach. As mentioned in the introduction section, the resolution of moderate satellite images, such as Landsat 8 limits the segmentation of urban areas even in our case, the Landsat 8 image has undergone preprocessing and pansharpening and has reached a resolution of 14.25 m. With the Otsu's threshold method, BLFEI index gives better results compared to the indices studied, but they remain also available for the extraction of built-up land though the performances vary from one index to another. The indices are tested in different areas with different sensors, in addition to specific limitations such as image acquisition conditions, preprocessing, topographic and climatic characteristics are imposed. Therefore, performance and evaluation vary from one place to another and also from a sensor to another. Finally, we aim to know which evaluation method is the most accurate.

Some indices such as BLFEI, BAEI and VgNIR-BI showed better results over others by using overlapping histogram and SDI; nevertheless, the Otsu method is more specifically suited to the BLFEI, BCI and IBI approaches, since each side (peak) of their histograms has a better symmetry with respect to the other. Moreover, as mentioned in the introduction Kassawmar et al. (2018) implemented a new method to increase the accuracy of classification based on increasing number of classes. In the first method based on the SDI separability and overlapping histograms, the classes are four, but for the second method (Otsu's method) the classes are only two, which means that the first method is more precise.

6.3 Results from accuracy assessment

With an OA of 95% and a coefficient k of 0.90 as shown in Table 5 and Figure 10, BLFEI, the new index is able to segment more efficiently and precisely the built-up areas than previously developed indices.

The errors of commission and omission of the built-up class for BLFEI are respectively 6% and 5% and are the lowest compared to the other studied indexes. Nevertheless, this slight error of omission means that the new approach effectively captures the built-up areas. BLFEI, the new method has reduced the commission and omission errors by (6% - 4%) and has only misclassified 9 pixels. These nine points have been wrongly classified, and are in majority, localized in the overlapping zone between the built-up and the barren areas. Our proposed index is followed by VgNIR-BI index with an OA equal to 90% and a k coefficient equal to 0.78 which has misclassified a twenty pixels, followed by BCI with an OA equal to 87% and a k coefficient equal to 0.72 which it has misclassified twenty six pixels. BLFEI the new index have an overall accuracy (5-14) % and a coefficient k (12-26) % higher than that of the state-of-the-art indices. For the quantitative comparison, Figure 10 shows in Hectares the surfaces occupied by built-up land for each index using Otsu's method and SVM classification. The surface of waterless studied zone is 338543 × 28.5 ×

28.5 = 27498.15 hectares. As shown in Figure 10, the extracted built-up surface by BLFEI index is not affected by the method used, unlike with other indices where the built-up region surfaces vary from the Otsu's method to the SVM classification. The percentages of built-up area extracted by BLFEI approach are 40.36% and 40.80% using Otsu's method and SVM classification, respectively, which leads to a relative error of 1% only. Thus the obtained results demonstrate that BLFEI index can be used as an alternative way for the extraction of built-up lands and for discriminating between land covers.

Conclusion

Extracting built-up areas from satellite images with medium spectral and spatial resolution like Landsat 8 is not an easy task. The recent developed built-up indices, such as BAEI, BCI, IBI, NBUI and VgNIR-BI, are able to extract built-up areas but they often merge them at different proportions with land cover categories such as barren and water. The purpose of this study is to develop and test a new built-up index and to compare it with the newly existing indices in order to examine its capacity in delimiting built-up areas and in distinguishing the components of land cover. By using BLFEI, the new index, built-up areas have been identified automatically, which minimizes the time required for the image classification process. The analyses of overlapping histograms, spectral discrimination index (SDI), Otsu's optimal thresholding and accuracy assessment make it relevant that our proposed new index outperforms existing indices. Firstly, by its robustness in discriminating built-up land from the barren class, secondly by its histogram of land without water which presented two predominant symmetrical sides leading to the best Otsu's optimal thresholding for the extraction of built-up areas, and thirdly by the accuracy assessment where its overall accuracy and its kappa coefficient are higher than those of the indices used for comparison. However, the values of the BLFEI index are the highest for water zones, therefore, a masking

operation is necessary for optimal thresholding, which is not an advantageous point for the new proposed index. Another point is that BLFEI index is applied only to one city and it is essential to test it for other regions with different land cover characteristics.

Acknowledgments

The authors are thankful to the editor and the anonymous reviewers for their valuable comments and insightful ideas, authors also acknowledge USGS and Google Earth, for providing the satellite data and images which we used in this study.

References

As-Syakur AR, Adnyana I, Arthana IW, Nuarsa IW. 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.

Bhatti SS, Tripathi NK. 2014. Built-up area extraction using Landsat 8 OLI imagery. GISc Remote Sens. 51(4): 445-467.

Bouzekri S, Lasbet AA, Lachehab A. 2015. A New Spectral Index for Extraction of Built-Up Area Using Landsat-8 Data. Indian Journal of society and Remote Sensing. 43(4):867–873.

Deng C, Wu C. 2012. BCI: A biophysical composition index for remote sensing of urban environments. Remote Sens Environ. 127:247-259.

Ehlers M, Klonus S, Åstrand PJ, Rosso P. 2010. Multi-sensor image fusion for pansharpening in remote sensing. International Journal of Image and Data Fusion. 1(1): 25-45. DOI: 10.1080/19479830903561985.

Ehlers M, Welch R. 1987. Stereocorrelation of Landsat TM Images. Photogrammetric Engineering & Remote Sensing. 53(9): 1231-1237.

Estoque CR, 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.

Feyisa GL, Meilby H, Jenerette GD, Pauliet S. 2016. Locally optimized separability enhancement indices for urban land cover mapping: Exploring thermal environmental

consequences of rapid urbanization in Addis Ababa, Ethiopia. Remote Sensing of Environment. 175: 14-31.

Fuyi T, Mohammed SK, Abdullah K, Lim HS, Ishola KS. 2013. A comparison of Atmospheric Correction Techniques for Environmental Applications. Proceeding of the IEEE International Conference on Space Science and Communication (IconSpace), July 1-3, Melaka, Malaysia.

Garg A, Pal D, Singh H, Pandey DC. 2016. A Comparative study of NDBI, NDISI and NDII for extraction of Urban Impervious Surface of Dehradun [Uttarakhand, India] using Landsat 8 Imagery. Proceeding of the IEEE. International Conference on Emerging Trends in Communication Technologies (ETCT). Nov 18-19. Dehradun, India. DOI: 10.1109/ETCT.2016.7882963.

Hazini S, Hashim M. 2015. Comparative analysis of product-level fusion, support vector machine, and artificial neural network approaches for land cover mapping. Arabian Journal of Geosciences. 8(11): 9763–9773.

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. Remote Sensing Letters. 1(4): 213–221.

Labib SM, Harris A. 2018. The potentials of Sentinel-2 and LandSat-8 data in green infrastructure extraction, using object based image analysis (OBIA) method. European Journal of Remote Sensing. 51(1): 231-240.

Li E, Du P, Samat A , Xia J, Che M. 2015. An automatic approach for urban land-cover classification from Landsat-8 OLI data. International Journal of Remote Sensing. 36(24): 5983–6007.

Li W, Du Z, Ling F, Zhou D, Wang H, Gui Y, Sun B, Zhang X. 2013. A comparison of land surface water mapping using the normalized difference water index from TM, ETM+ and ALI. Remote Sens. 5(11): 5530–5549.

Liu B, Chen X, Ma C, Zhang D, Zhou X, He Y. 2010. Research on threshold segmentation in tracking technology of moving objects. Proceeding of the IEEE. 2nd International Conference on Industrial Mechatronics and Automation (ICIMA). May 30-31, Wuhan, China. DOI:10.1109/ICINDMA.2010.5538048.

Morteza S, Nia S, Wang D Z , Bohlman S A , Gader P , Sarah J, Graves, Petrovic M. 2015. Impact of atmospheric correction and image filtering on hyperspectral classification of tree species using support vector machine. Journal of Applied Remote Sensing. 9(1): 095990. DOI:10.1117/1.JRS.9.095990.

Kaufman YJ, Remer LA. 1994. Detection of forests using mid-IR reflectance: an application for aerosol studies. IEEE T Geosci Remote. 32(3): 672-683.

Kassawmar T, Eckertc S, Hurnic K , Zelekeb G, Hurnia H. 2018. Reducing landscape heterogeneity for improved land use and land cover (LULC) classification across the large and complex Ethiopian highlands. Geocarto International. 33(1): 53-69

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. 31: 321-326.

Otsu N, 1979. A threshold selection method from gray-level histograms. IEEE Trans.Syst. Man Cybern. 9(1): 62–66.

Patel N, Mukherjee R. 2015. Extraction of impervious features from spectral indices using artificial neural network. Arab J Geosci. 8(6): 3729–3741.

Piyoosh AK, Ghosh SK. 2018. Development of a modified bare-soil and urban index for Landsat 8 satellite data. Geocarto International. 33(4): 423-442. DOI:10.1080/10106049.2016.1273401.

Rwanga SS, Ndambuki JM. 2017. Accuracy Assessment of Land Use/Land Cover Classification Using Remote Sensing and GIS. International Journal of Geosciences 8(4): 611-622.

Sinha P, Verma NK, Ayele E. 2016. Urban Built-up Area Extraction and Change Detection of Adama Municipal Area using Time-Series Landsat Images. International Journal of Advanced Remote Sensing and GIS. 5(8):1886-1895.

Stathakis D, Perakis K, Savin I. 2012. Efficient segmentation of urban areas by the VIBI. Int J Remote Sens. 33(20): 6361-6377. DOI 10.1080/01431161.2012.687842.

Sun G, Chen X, Jia X, Yao Y,Wang Z. 2016. Combinational Build-Up Index (CBI) for Effective Impervious Surface Mapping in Urban Areas. IEEE Journal Of Selected Topics in Applied Earth Observations and Remote Sensing. 9(5):2081-2092.

Tucker CJ, Grant DM, Dykstra JD. 2004. NASA's Global Orthorectified Landsat Data Set. Photogrammetric Engineering & Remote Sensing. 70(3): 313–322.

Varshney A, Rajesh E. 2014. A comparative study of built-up index approaches for automated extraction of built-up regions from remote sensing data. Journal of the Indian Society of Remote Sensing. 42(3):659–663.

Waqar MM, Mirza JF, Mumtaz R, Hussain E. 2012. Development of new indices for extraction of built-up area and bare soil from landsat data. Open Access Scientific Reports. 1(1):1- 4.

Xu H. 2007. Extraction of Urban Built-up Land Features from Landsat Imagery Using a Thematic oriented Index Combination Technique. Photogrammetric Engineering & Remote Sensing. 73(12):1381–1391.

Xu H. 2008. A new index for delineating built-up land features in satellite imagery. Int J Remote Sens. 29(14): 4269-4276. 10.1080/01431160802039957.

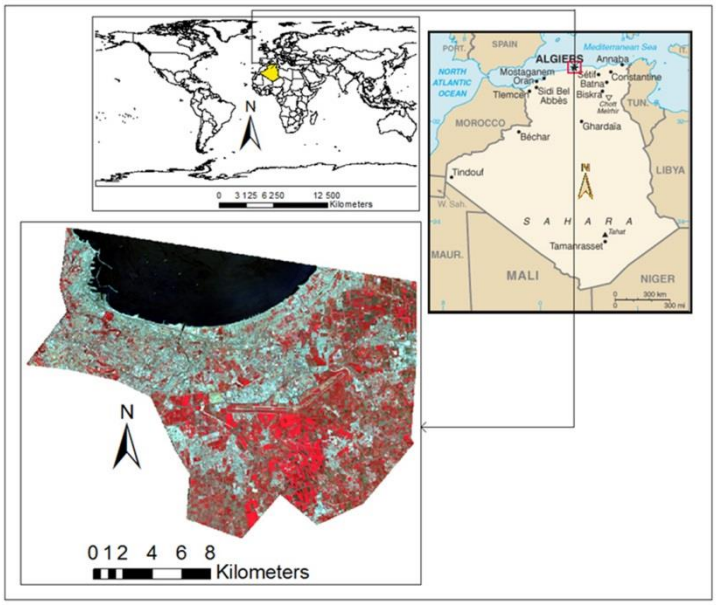
Xu Q, Zhang Y, Li B. 2014. Recent advances in pansharpening and key problems in applications. International Journal of Image and Data Fusion. 5(3): 175-195. DOI:10.1080/19479832.2014.889227.

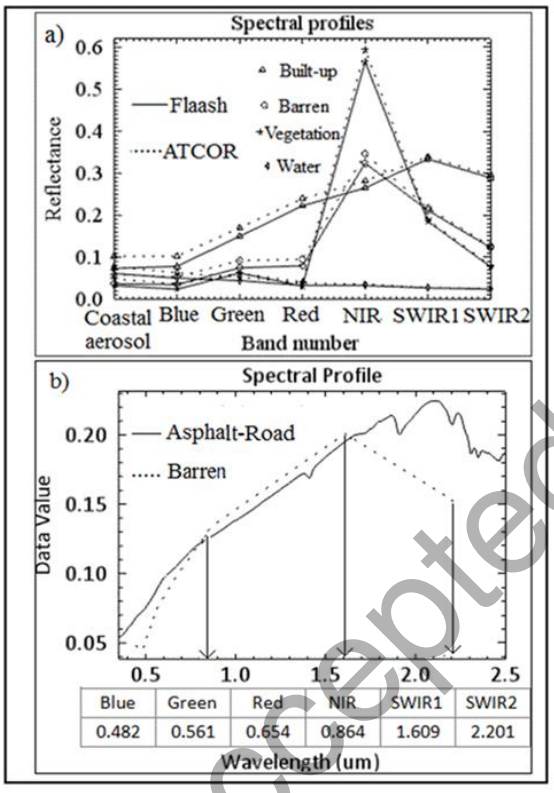
Zha GY, Gao J, Ni S. 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. International Journal of Remote Sensing, 24(3): 583–594. 10.1080/01431160304987.

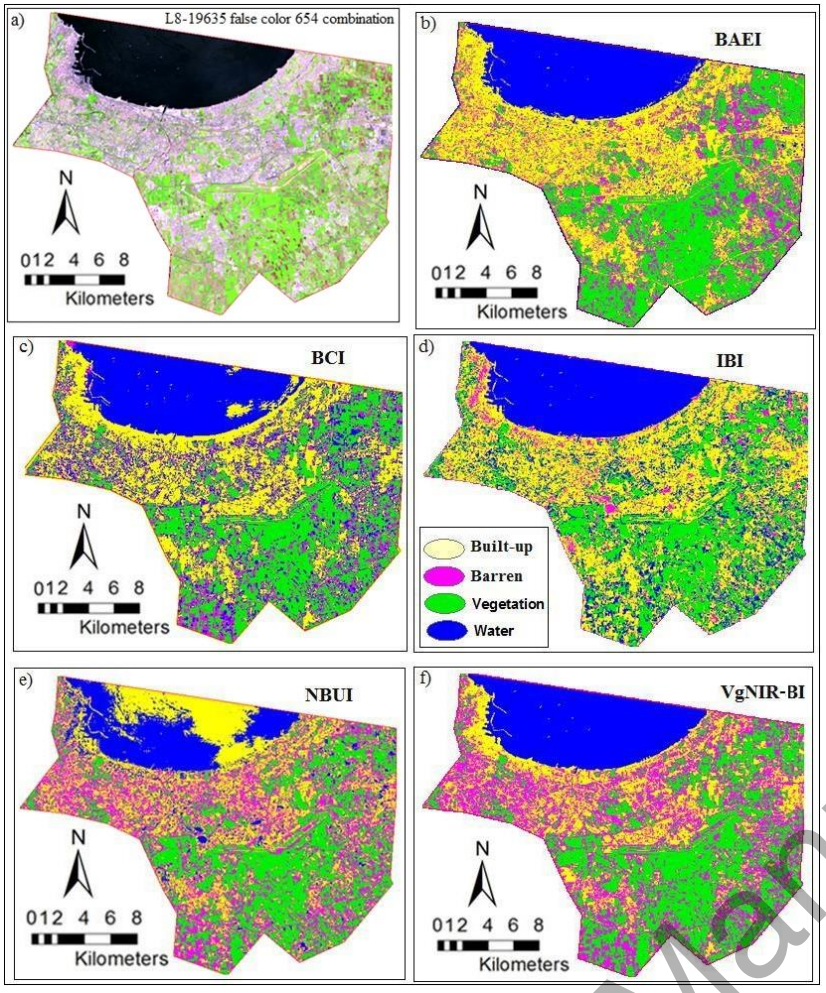
Zhang H, Lan Y, Lacey R, Hoffmann WC, Huang Y .2009. Analysis of vegetation indices derived from aerial multispectral and ground hyperspectral data. International Journal of Agricultural and Biological Engineering. 2(3):33–40.

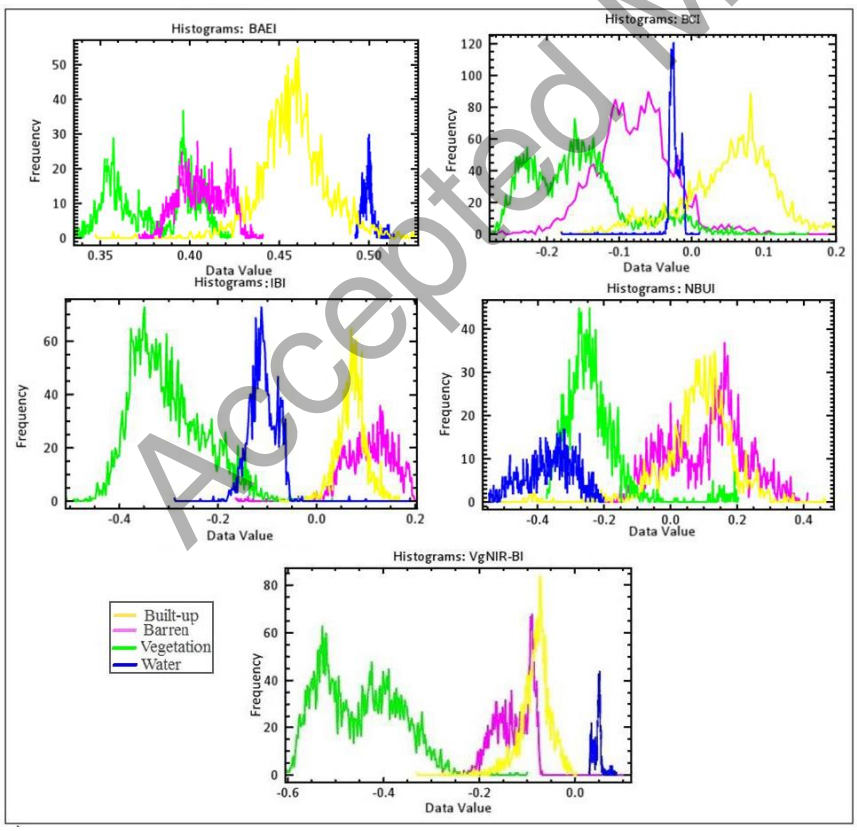
Zhang J, Li P, Wang J. 2014. Urban Built-Up Area Extraction from Landsat TM/ETM+ Images Using Spectral Information and Multivariate Texture. Remote Sens. 6(8): 7339-7359. doi: 10.3390/rs6087339.

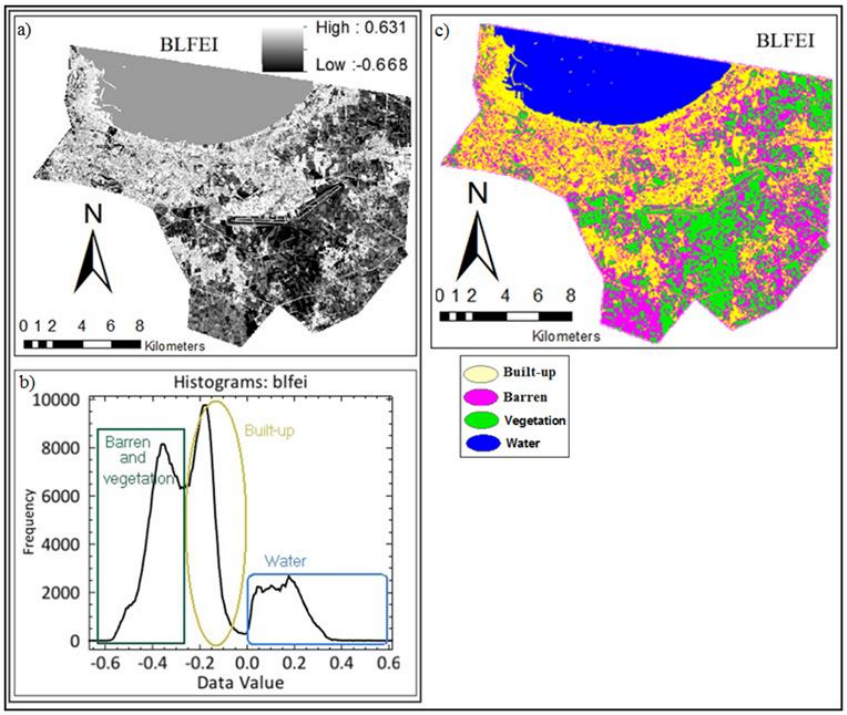
Zhou Y, Yanga G, Wanga S, Wang L, Wang F, Liu X. 2014. A new index for mapping built-up and bare land areas from Landsat-8 OLI data. Remote Sensing Letters. 5(10): 862–871. doi: 10.1080/2150704X.2014.973996.

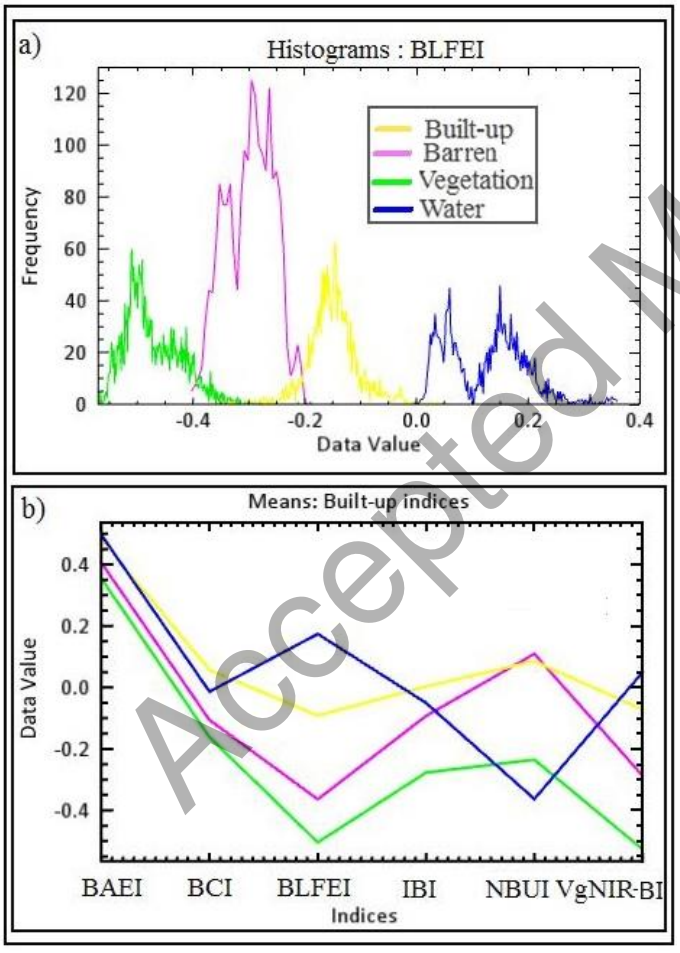


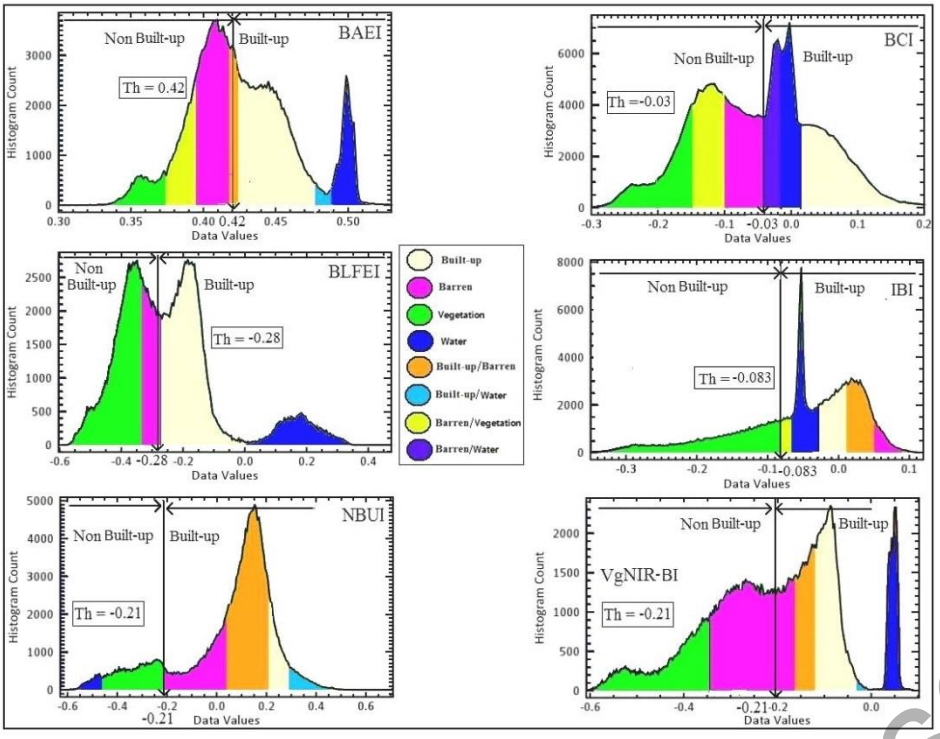


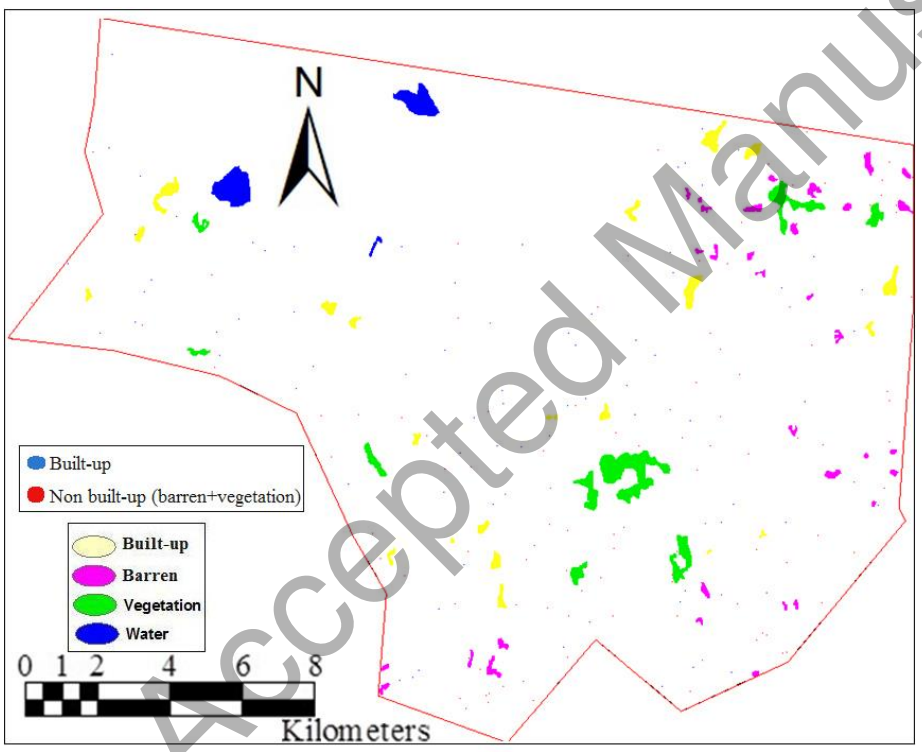


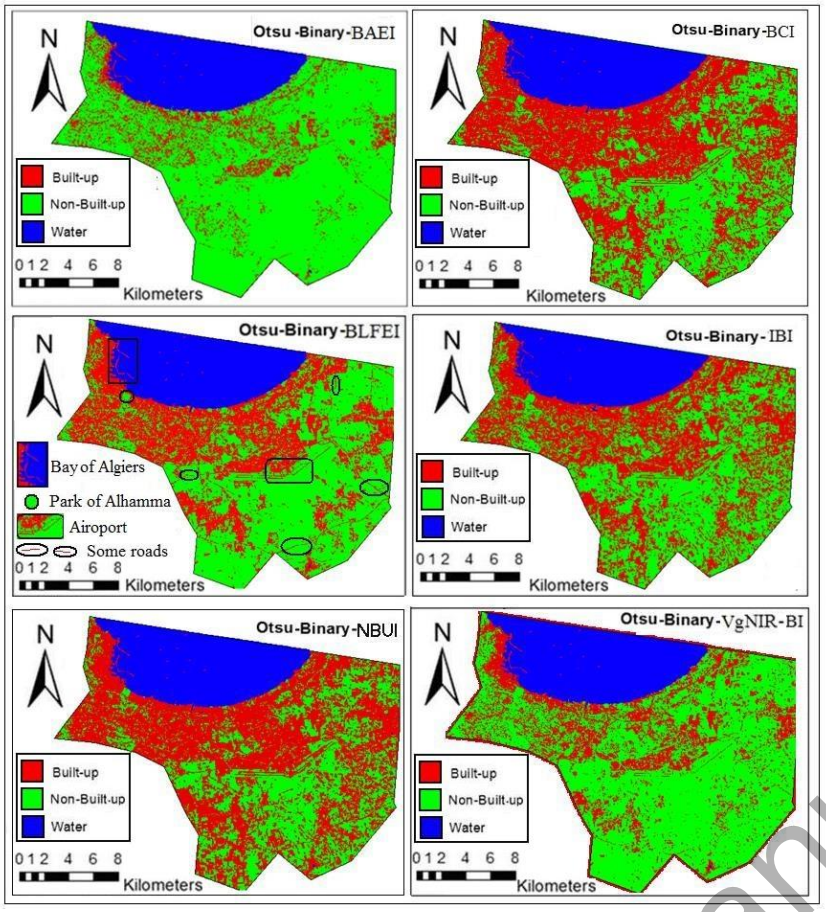












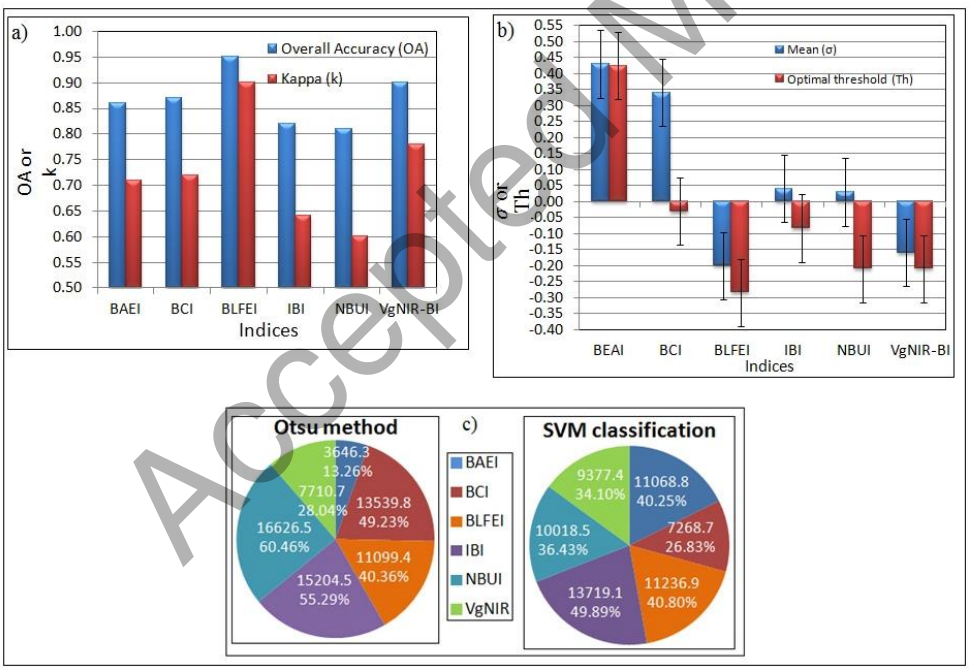


Table 1. Some of built-up extraction indices.

Title	Designation	Authors	Year	Formula	Satellite-Sensor	Overall Accuracy (OA%)	Kappa(k)
Urban Index	UI	Kawamura et al	1996	$\frac{(B_7-B_4)}{(B_7+B_4)}$ (1)	Landsat-TM	-	-
NDBI derived	$NDBI_b$-$NDVI_b$ *b: binary	Zha et al	2003	$(\frac{(B_5-B_4)}{(B_5+B_4)})_b - (\frac{(B_4-B_2)}{(B_4+B_2)})_b$ (2)	Landsat-TM	96.2	-
Improved NDBI	$NDBI_c$-$NDVI_c$ *c: continous	He et al	2010	$NDBI_c$-$NDVIc$ (3)	Landsat-TM	-	-
Vegetation Index Built-up Iindex	VIBI	Stathakis et al	2012	$\frac{NDVI}{(NDVI+NDBI)}$ (4)	Landsat-TM	91.13	0.56
Built-up Area Extraction Method	$BAEM_{OLI}$	Bhatti and Tripathi	2014	$NDBI_{OLI}$-$NDVI_{OLI}$-$MNDWI_{OLI}$ $NDBI_{OLI}=\frac{PCA(OLI_6,OLI_7)+PCA(OLI_{10},OLI_{11})-OLI_5}{PCA(OLI_6,OLI_7)+PCA(OLI_{10},OLI_{11})+OLI_5}$ (5)	Landsat8-OLI	80.5	0.59
Visible red - Based Built-up Index	VrNIR-BI	Estoque and Murayama	2015	$\frac{(\rho_{Red}-\rho_{NIR})}{(\rho_{Red}+\rho_{NIR})}$ (6)	Landsat ETM+ Landsat8-OLI	92.1	0.87
Vegetation and Water Masking Index	VWMI	Li et al	2015	$\frac{(NDVI-N_{SWIR1}-MNDWI)}{(NDVI+N_{SWIR1}-MNDWI)}$ (7) where N_{SWIR1} denotes the normalized values of SWIR1.	Landsat8-OLI	-	-
Normalized Ratio Urban Index	NRUI	Piyoosh and Ghosh	2016	$\frac{RUI-MNDSI}{RUI+MNDSI}$ (8) $RUI=\frac{BCI}{MNDSI}$ $MNDSI=\frac{OLI_7-OLI_8}{OLI_7+OLI_8}$	Landsat8-OLI/TIRS	-	-

*"-" OA and k are not mentioned by authors.

Table 2. Characteristic of Landsat 8 OLI/ TIRS sensors.

Satellite	Sensor	Band name	Band number	Spectral range (μm)	Spatial resolution(m)	Spectral resolution	Radiometric resolution	Satallite altitude
Landsat-8	OLI	Coastal Aerosol	1	0.43 - 0.45	30	9 bands	16 bits	705 Km
		Blue	2	0.45 - 0.51				
		Green	3	0.53 - 0.59				
		Red	4	0.64 – 0.67				
		Near Infrared	5	0.85 – 0.88				
		Shortwave Infrared 1	6	1.57 – 1.65				
		Shortwave Infrared 2	7	2.11 – 2.29				
		Panchromatic	8	0.50 – 0.68	15			
		Cirrus	9	1.36 – 1.38	30			
	TIRS	Thermal Infrared (TIRS) 1	10	10.60 – 11.19	100 * (30)	2 bands		
		Thermal Infrared (TIRS) 2	11	11.50 – 12.51	100 * (30)			

*The TIRS bands are resampled at 30 m in the data product delivered, in fact they are acquired at a resolution of 100 m. https://landsat.usgs.gov/what-are-band-designations-landsat-satellites

Table 3. Values of Spectral Discrimination Index (SDI) between land features for the BLFEI index and evaluated indices.

Indices	Built-up/Barren	Built-up/Water	Barren/Vegetation	Barren/Water	Built-up/Vegetation	Vegetation/Water
BAEI	1.25	1.33	0.31	3.00	2.00	4.00
BCI	1.08	1.00	0.13	0.86	1.57	1.75
BLFEI	2.10	1.85	1.75	4.81	3.25	3.00
IBI	0.40	4.00	0.03	3.33	3.05	3.42
NBUI	0.08	0.55	2.95	0.52	3.10	3.43
VgNIR-BI	1.69	3.25	2.00	3.60	3.40	4.63

*SDI decision: Poor (SDI<1), good (1≤SDI<3), excellent (SDI≥3).

Table 4. Means, standard deviations (SD) and optimal thresholds (Th) for BLFEI index and the built-up indices

Indices	BAEI	BCI	BLFEI	IBI	NBUI	VgNIR-BI
Mean	0.43	0.34	-0.20	0.04	0.03	-0.16
SD	2.27	0.51	0.19	0.31	0.21	0.18
Th	0.424	-0.03	-0.284	-0.083	-0.21	-0.21

Table 5. Overall accuracy (OA), Kappa coefficient (k), commission error and omission error values for the new index and the built-up indices.

Indices	L8-19635 classified images	L8-19634 and Google earth as reference data						
			b	nb	Total	Commission error (%)	Omission error (%)	OA (%)	k
BAEI		b	64	11	75	14.60	20.00		
		nb	16	109	125	12.80	9.16	86	0.71
		Total	80	120	200				
BCI		b	67	10	77	12.98	19.27		
		nb	16	107	123	13.08	8.50	87	0.72
		Total	83	117	200				
BLFEI		b	76	5	81	6.17	5.00		
		nb	4	115	119	3.36	4.16	95	0.90
		Total	80	120	200				
IBI		b	61	10	71	14.08	20.77		
		nb	16	103	129	12.40	8.84	82	0.64
		Total	77	113	200				
NBUI		b	64	16	80	20.00	25.58		
		nb	22	98	120	18.30	14.03	81	0.60
		Total	86	114	200				
VgNIR-BI		b	65	13	78	16.00	9.70		
		nb	7	115	122	5.70	10.15	90	0.78
		Total	72	128	200				

* b = Built-up nb = Non Built-up OA = Overall accuracy k = Kappa coefficient

Table 6. Scale used for describing the relative strength agreement of Kappa statistic.

Kappa statistic	strength of agreement
<0.00	Poor
0.00-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost perfect (Very good)