Automated Built-Up Extraction Index: A New Technique for Mapping Surface Built-Up Areas Using LANDSAT 8 OLI Imagery

Mohammad Karimi Firozjaei 1,2, Amir Sedighi 1, Majid Kiavarz 1, Salman Qureshi 2,* Dagmar Haase 2 and Seyed Kazem Alavipanah 1,2

1 Department of Remote Sensing and GIS, University of Tehran, Tehran 1417853933, Iran
2 Department of Geography, Humboldt University Berlin, Unter den Linden 6, 10099 Berlin, Germany
* Correspondence: salman.qureshi@geo.hu-berlin.de

Received: 29 June 2019; Accepted: 16 August 2019; Published: 21 August 2019

Abstract: Accurate built-up area extraction is one of the most critical issues in land-cover classification. In previous studies, various techniques have been developed for built-up area extraction using Landsat images. However, the efficiency of these techniques under different technical and geographical conditions, especially for bare and sandy areas, is not optimal. One of the main challenges of built-up area extraction techniques is to determine an optimum and stable threshold with the highest possible accuracy. In many of these techniques, the optimum threshold value fluctuates substantially in different parts of the image scene. The purpose of this study is to provide a new index to improve built-up area extraction with a stable optimum threshold for different environments. In this study, the developed Automated Built-up Extraction Index (ABEI) is presented to improve the classification accuracy in areas containing bare and sandy surfaces. To develop and evaluate the accuracy of the new method for built-up area extraction with Landsat 8 OLI reflective bands, five test sites located in the Iranian cities (Babol, Naqadeh, Kashmar, Bam and Masjed Soleyman), eleven European cities (Athens, Brussels, Bucharest, Budapest, Ciechanow, Hamburg, Lyon, Madrid, Riga, Rome and Porto) and high resolution layer imperviousness (HRLI) data were used. Each site has varying environmental and complex surface coverage conditions. To determine the optimal weights for each of the Landsat 8 OLI reflective bands, the pure pixel sets for different classes and the improved gravitational search algorithm (IGSA) optimization were used. The Kappa coefficient and overall error were calculated to evaluate the accuracy of the built-up extraction map. Additionally, the ABEI performance was compared with the urban index (UI) and normalized difference built-up index (NDBI) performances. In each of the five test sites and eleven cities, the extraction accuracy of the built-up areas using the ABEI was higher than that using the UI and NDBI (P-value of 0.01). The relative standard deviations of the optimal threshold values for the ABEI and UI were 27 and 155% (at five test sites) and were 16 and 37% (at eleven European cities), respectively, which indicates the stability of the ABEI threshold value when the location and environmental conditions change. The results of this study demonstrated that the ABEI can be used to extract built-up areas from other land covers. This index is effective even in bare soil and sandy areas, where other indices experience major challenges.

Keywords: built-up extraction; threshold stability; improved gravitational search algorithm; Landsat 8

1. Introduction

Urban areas generally cover a small portion of land worldwide. However, the high population density and intensity of utilizing resources relative to their surroundings turn these areas into highly important land areas [1,2]. Understanding the spatial distribution and growth pattern of urban built-up
areas is vital for urban planning, natural resource management, climate change monitoring, etc., and the ecological, economic and social effects of built-up lands have made the mapping of built-up areas very important [2–6].

Satellite images are a valuable source for built-up mapping in areas with different environmental conditions [7–12]. Remote sensing technology provides an integrated, comprehensive view of urban areas that is difficult to achieve through surveying. Another advantage of using remote sensing data for urban studies is the availability of archives of past periods that can be used to review, monitor, and prepare a distribution map of built-up areas over time [13,14]. Over recent years, satellite imagery capabilities have been used for studies on urban expansion [15], effective factors on urban development [16], and the negative effects of urban growth [17].

Generally, in previous studies for mapping built-up areas usually coarse-resolution satellite imagery such as moderate resolution imaging spectroradiometer (MODIS) [18,19] and defense meteorological program / operational line-scan system (DMSP/OLS) [20], medium-resolution satellite imagery such as Landsat [21] and Sentinel-2 [22], and high-resolution satellite imagery such as Quick Bird [23] and WorldView [24] have been used. Landsat imagery has the potential to provide a historical record for mapping built-up areas, as Landsat has been acquiring images since the early 1980s. Landsat imagery has been used frequently to extract built-up areas in urban studies [25]. Its large coverage, suitable spatial resolution (30 m), historical observations and free access have proved that Landsat imagery is useful for mapping built-up areas [1,2,25].

Currently, satellite imagery is increasingly being used automatically and semi-automatically to generate various maps of land surface characteristics, such as vegetation, snow, water and built-up land characteristics [26–30]. Several techniques have been developed for this purpose, which are grouped into two general categories. The first group relates to image classification techniques, which have been widely developed based on pixel-based and object-oriented classification methods [13,31]. The second group involves information extraction from images based on specific indices [2]. Both groups of techniques have a set of advantages and disadvantages. The accuracy of classification methods may be higher than indices, but, in general, the use of indices in a case of certain land-cover extraction according to the amount of information, need to train data, and time required for processing is more efficient than the use of classification techniques.

Different indices have been developed for built-up area extraction using satellite imagery [2,7–9,11,29]. The basis of the development of these indices is the spectral behaviour of the built-up land and various features associated with the wavelengths of the electromagnetic spectrum in terms of absorption or reflection. Different bands of the satellite image represent the different wavelengths of the electromagnetic spectrum. Accordingly, these indices have been developed with a special combination of image bands. Investigating and studying the spectral signatures of various features are necessary to create and develop various indices related to the identification and extraction of these phenomena [26].

In previous studies, different indices, such as the index-based built-up index (IBI) [12], urban index (UI) [32,33], normalized difference bareness index (NDBaI) [33], normalized difference built-up index (NDBI) [29], modified NDBI [9], normalized built-up area index (NBAI) and band ratio for built-up area (BRBA) [11], new built-up index (NBI) [8], built-up index (BUI) [10], bare soil index (BI) [34], enhanced built-up and bareness index (EBBI) [7], and built-up area extraction method (BAEM) [2], have been developed for built-up and urban area mapping.

Kawamura et al. (1996) proposed the UI for mapping built-up areas. The UI is computed by using the bands 7 and 4 from Landsat Thematic Mapper (TM) imagery [32]. Zha et al. (2003) developed the NDBI for the automatic mapping of built-up areas. They extracted built-up areas through arithmetic calculation using the NDVI and NDBI derived from Landsat TM images. In the binary images created by the NDBI and NDVI images via this technique, it is assumed that positive values of the NDBI indicate built-up areas and that positive NDVI values indicate vegetation covers. Zha et al. (2003) reported that their method is incapable of separating built-up areas from bare land [29].
et al. (2012) developed the EBBI for mapping built-up and bare land in urban areas. They extracted built-up and bare land areas through single calculation using near infrared (NIR), short wave infrared (SWIR), and thermal infrared (TIR) bands simultaneously [7]. Waqar et al. (2012) developed NBAI and BRBA Indices for mapping built-up area and bare soil covers using Landsat TM images. They extracted built-up areas through NBAI using the bands 2, 5 and 7 of Landsat TM images. Also, they extracted built-up areas through BRBA using the ratio of bands 3 and 5 of Landsat TM images [11]. Kaimaris and Patias (2016) developed the BUI based on the combination of the bands 3, 5 and 7 of Landsat Enhanced Thematic Mapper Plus (ETM+) images. In BUI, unlike other indices, the NIR band was not used for mapping built-up areas [10]. The IBI was distinguished from other built-up indices because of its first-time use of thematic indices rather than using original bands. The soil adjusted vegetation index (SAVI), modified normalized difference water index (MNDWI) and NDBI indices were used in IBI [12].

The NDBI and UI provide quick mapping of built-up areas or bare land. However, these two indices are unable to separate built-up areas from bare land effectively [9,29,35]. He et al. [9] noted in their study that this incapability was due to the complex spectral response generated by the combination of the spectral responses of built-up areas, vegetation and bare land, especially in terms of mixed pixels in areas with heterogeneous features.

Having reviewed previous studies in this field of study, it is clear that most of the developed indices are effective in mapping bare land and built-up areas, but these indices are not able to completely isolate built-up and bare lands in the same area [7]. In different areas, the performances of these indices are affected by the type of land use in the region. Therefore, these indices have problems in distinguishing between built-up and other land use areas that have similar spectral behaviours. Even the spectral heterogeneity between pixels related to built-up areas reduces the accuracies of the indices. Generally, the accuracies of these indices for built-up area extraction are low [10]. For this reason, in some studies, several different indices have been combined with built-up area extraction [2].

Simplified techniques, such as two-band indices and single-band thresholding, may not accurately separate built-up pixels from non-built-up bright surfaces, particularly in scenes where non-built-up bright surfaces are seen. Current built-up extraction methods have accuracy problems in distinguishing built-up areas from bare land and rock surfaces [2,7,8,29]. It is worth mentioning that no existing built-up index has been capable of automatically separating built-up and bare lands using Landsat 8 OLI (Operational Land Imager) images. Therefore, in this study, a multiple-band index, called the Automated Built-up Extraction Index (ABEI), was introduced with two main objectives: (1) increase the accuracy of built-up extraction mapping by automatically separating built-up areas from other non-built-up bright surfaces and (2) test the novel technique under various environmental conditions and assess its accuracy in comparison with other techniques.

The major distinctions between the present study and other studies in this field are (1) the development of a new index for built-up extraction using all of the Landsat 8 OLI reflective bands, (2) the use of the improved gravitational search algorithm (IGSA) optimization to determine the relevant weights in different reflective bands and (3) the use of the stable optimum threshold value for the classification of built-up areas under different environmental conditions and various surfaces.

2. Test Sites and Cities

To perform an accuracy assessment with the new technique for automatic built-up extraction from Landsat 8 OLI reflective bands, 5 test sites with different environmental conditions and complex land covers were selected. Test sites from Babol, Naqadeh, Kashmar, Bam and Masjed Soleyman are included and located in different geographical locations with different climatic conditions in Iran. The characteristics of the test sites are shown in Table 1.
Table 1. General characteristics of the test sites and cities.

<table>
<thead>
<tr>
<th>Test Site and City</th>
<th>Centre Point Coordinate (Lon, Lat-WGS84)</th>
<th>Country</th>
<th>Area (Km²)</th>
<th>Mean Alt. (m)</th>
<th>Climate</th>
<th>Population (2019)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test sites</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Babol</td>
<td>52.68, 36.54</td>
<td>Iran</td>
<td>2.0</td>
<td></td>
<td>Humid</td>
<td>&gt;5600</td>
</tr>
<tr>
<td>Naqadeh</td>
<td>45.38, 36.95</td>
<td>Iran</td>
<td>0.5</td>
<td>1325</td>
<td>Semi-humid</td>
<td>&gt;1500</td>
</tr>
<tr>
<td>Masjed Soleyman</td>
<td>49.28, 31.96</td>
<td>Iran</td>
<td>3.4</td>
<td>240</td>
<td>Dry and semi-arid</td>
<td>&gt;6500</td>
</tr>
<tr>
<td>Kashmar</td>
<td>58.46, 35.24</td>
<td>Iran</td>
<td>5.4</td>
<td>1060</td>
<td>Warm and dry</td>
<td>&gt;7000</td>
</tr>
<tr>
<td>Bam</td>
<td>58.33, 29.10</td>
<td>Iran</td>
<td>1.3</td>
<td>1083</td>
<td>Warm and dry</td>
<td>&gt;2800</td>
</tr>
<tr>
<td>European cities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rome</td>
<td>12.45, 41.85</td>
<td>Italy</td>
<td>631.7</td>
<td></td>
<td>Mediterranean</td>
<td>&gt;4,234,000</td>
</tr>
<tr>
<td>Riga</td>
<td>24.09, 56.94</td>
<td>Latvia</td>
<td>558.0</td>
<td>10</td>
<td>Humid continental</td>
<td>&gt;633,000</td>
</tr>
<tr>
<td>Madrid</td>
<td>−3.70, 40.41</td>
<td>Spain</td>
<td>2332.3</td>
<td>650</td>
<td>Mediterranean and semi-arid</td>
<td>&gt;6,559,000</td>
</tr>
<tr>
<td>Athens</td>
<td>23.72, 37.97</td>
<td>Greece</td>
<td>2142.2</td>
<td>80</td>
<td>Mediterranean</td>
<td>&gt;3,154,000</td>
</tr>
<tr>
<td>Brussels</td>
<td>4.38, 50.85</td>
<td>Belgium</td>
<td>608.1</td>
<td>60</td>
<td>Oceanic</td>
<td>&gt;2,065,000</td>
</tr>
<tr>
<td>Porto</td>
<td>−8.60, 41.16</td>
<td>Portugal</td>
<td>481.4</td>
<td>80</td>
<td>Mediterranean</td>
<td>&gt;1,309,000</td>
</tr>
<tr>
<td>Lyon</td>
<td>4.83, 45.76</td>
<td>France</td>
<td>1143.6</td>
<td>175</td>
<td>Humid subtropical</td>
<td>&gt;1,704,000</td>
</tr>
<tr>
<td>Ciechanow</td>
<td>20.60, 52.82</td>
<td>Poland</td>
<td>81.1</td>
<td>151</td>
<td>Humid subtropical</td>
<td>&gt;44,000</td>
</tr>
<tr>
<td>Hamburg</td>
<td>10.02, 53.60</td>
<td>Germany</td>
<td>1097.5</td>
<td>10</td>
<td>Oceanic</td>
<td>&gt;1,791,000</td>
</tr>
<tr>
<td>Budapest</td>
<td>19.07, 47.59</td>
<td>Hungary</td>
<td>3664.3</td>
<td>120</td>
<td>Oceanic and Humid subtropical</td>
<td>&gt;1,763,000</td>
</tr>
<tr>
<td>Bucharest</td>
<td>26.10, 44.42</td>
<td>Romania</td>
<td>1385.7</td>
<td>85</td>
<td>Humid continental</td>
<td>&gt;1,812,000</td>
</tr>
</tbody>
</table>

The test sites are selected in the form of complex features in terms of spectral properties, such as soil, sand, and rock properties. The Babol and Naqadeh test sites are selected due to their humid and semi-humid climates, including built-up, agricultural, water, soil, and dense vegetation covers. The Kashmar and Bam test sites are selected due to their warm and dry climates, including built-up, bare land, sandy soil and vegetation covers. The test site of Masjed Soleyman is a mountainous region with a semi-arid climate, including built-up, bare land and rock surfaces.

To broadly evaluate the ABEI performance in different geographical locations with different environmental conditions, the ABEI were implemented on 11 European cities including Athens, Brussels, Bucharest, Budapest, Ciechanow, Hamburg, Lyon, Madrid, Riga, Rome and Porto. The cities are selected to cover different geographical locations and climatic conditions across Europe. Among the selected cities, Athens, Riga, Rome and Porto are Coastal cities. The areas around Athens and Madrid include bare land. The Brussels, Budapest, Bucharest, Ciechanow, Hamburg, Lyon, Riga and Porto have high vegetation cover due to humid climates. Above all, the main reason for selecting these cities in Europe was the absence of cloud-free Landsat 8 OLI images in 2015. In summary, the characteristics of the European cities are listed in Table 1. The geographical locations of the test sites and European cities are shown in Figure 1.
3. Data and Methods

3.1. Landsat Images and Reference Data

In the present study, the reflective bands of the satellite images acquired by Landsat 8 OLI were used. The satellite images with no cloud cover and no rainfall within 2 days before satellite overpass were selected. The images were georeferenced in the UTM coordinate system. The downloaded data from USGS included the highest quality Level-1 Precision Terrain (L1TP) data, which were considered suitable for time-series analysis. The geo-registration was consistent, and the root mean square error (RMSE) was less than 12 m ($\leq 0.5$ pixel) [36]. For each of the five test sites, high-spatial-resolution Google Earth Images were used as reference data. The times for the reference data and Landsat 8 OLI images used in this study were very similar at the test sites. To digitize and extract built-up areas from Google Earth images, a photointerpretation method was used. The derived built-up map was used to accurately assess the different build-up extraction methods.

Additionally, the high resolution layer imperviousness (HRLI) data have been used as reference data to evaluate ABEI performance in European cities. The HRLI data provided by the Copernicus Global Land Service (CGLS) had a pixel resolution of 20 m. The HRLI data were generated based on IRS-P6/Resourcesat-2 LISS-III, SPOT 5 and Landsat 8 data sets for different regions of Europe. These reference data for 2006, 2009, 2012 and 2015 are available at the https://land.copernicus.eu/ website. The results of applying the ABEI on the Landsat 8 OLI imagery from European cities were evaluated with the HRLI data of 2015 for each city. Details of the Landsat images and reference data are presented in Table 2.
Table 2. Description of the Landsat 8 OLI scenes and corresponding reference data.

<table>
<thead>
<tr>
<th>Test Site</th>
<th>Landsat Scene</th>
<th>Reference Data and Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babol</td>
<td>1 May 2015</td>
<td>Path 164, Row 35 Google Earth™ image acquired on 29 April 2015, ©Digital Globe</td>
</tr>
<tr>
<td>Naqadeh</td>
<td>13 August 2014</td>
<td>Path 169, Row 34 Google Earth™ image acquired on 10 August 2014, ©Digital Globe</td>
</tr>
<tr>
<td>Masjed Soleyman</td>
<td>15 August 2016</td>
<td>Path 165, Row 38 Google Earth™ image acquired on 14 August 2016, ©Digital Globe</td>
</tr>
<tr>
<td>Kashmar</td>
<td>14 October 2016</td>
<td>Path 160, Row 35 Google Earth™ image acquired on 14 August 2016, ©Digital Globe</td>
</tr>
<tr>
<td>Bam</td>
<td>6 May 2018</td>
<td>Path 159, Row 40 Google Earth™ image acquired on 9 May 2018, ©Digital Globe</td>
</tr>
</tbody>
</table>

Cities

<table>
<thead>
<tr>
<th>Cities</th>
<th>Acquisition Date</th>
<th>Path</th>
<th>Row</th>
<th>Reference Data and Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rome</td>
<td>17 July 2015</td>
<td>191</td>
<td>031</td>
<td>High Resolution Layer</td>
</tr>
<tr>
<td>Riga</td>
<td>9 May 2015</td>
<td>188</td>
<td>020</td>
<td>Imperviousness (HRLI) for the 2015 reference year, European Environment Agency</td>
</tr>
<tr>
<td>Madrid</td>
<td>7 July 2015</td>
<td>201</td>
<td>032</td>
<td></td>
</tr>
<tr>
<td>Athens</td>
<td>26 August 2015</td>
<td>180</td>
<td>031</td>
<td></td>
</tr>
<tr>
<td>Brussels</td>
<td>11 September 2015</td>
<td>199</td>
<td>025</td>
<td></td>
</tr>
<tr>
<td>Porto</td>
<td>12 July 2015</td>
<td>204</td>
<td>031</td>
<td></td>
</tr>
<tr>
<td>Lyon</td>
<td>12 August 2015</td>
<td>197</td>
<td>028</td>
<td></td>
</tr>
<tr>
<td>Ciechanow</td>
<td>4 August 2015</td>
<td>189</td>
<td>023</td>
<td></td>
</tr>
<tr>
<td>Hamburg</td>
<td>21 August 2015</td>
<td>196</td>
<td>023</td>
<td></td>
</tr>
<tr>
<td>Budapest</td>
<td>29 August 2015</td>
<td>188</td>
<td>027</td>
<td></td>
</tr>
<tr>
<td>Bucharest</td>
<td>3 August 2015</td>
<td>189</td>
<td>029</td>
<td></td>
</tr>
</tbody>
</table>

3.2. Methods

3.2.1. Image Preprocessing

Atmospheric correction of satellite imagery was performed using the Fast Line-of-sight Atmospheric Analysis of Hypercubes (FLAASH) algorithm [37]. This module employs the MODTRAN 6 model [38]. To this end, some parameters, such as the satellite overpass time, the sensor altitude, the geographical location, the specific atmospheric model related to the region, and the solar zenith angle, are considered. For each of the five test sites, the aerosol optical depth (AOD) and total water vapor content in the atmosphere used in the FLAASH algorithm were derived from the Terra MODIS atmospheric profile product, which was collected by the Terra platform [18]. The overpass times of Terra and Landsat 8 approximately match each other (less than a 2 h difference). There was no need for manual co-registration because the co-registration between the Landsat and Google Earth images at the test sites was accurate.

3.2.2. Formulation of the Automated Built up Extraction Index (ABEI)

Landsat 8 OLI reflective bands (7 bands) were used to develop the ABEI. The purpose of the ABEI is to create maximum similarity between the pixel values of the built-up class and the maximum difference between the pixel values of the built-up class and other land-cover classes. For this purpose, the combination of the reflective bands of Landsat 8 OLI and the weight of each band are used, as shown in Equation (1).

\[
ABEI = \mu_1 \text{Ref}_1 + \mu_2 \text{Ref}_2 + \mu_3 \text{Ref}_3 + \mu_4 \text{Ref}_4 + \mu_5 \text{Ref}_5 + \mu_6 \text{Ref}_6 + \mu_7 \text{Ref}_7
\]  

(1)

In this study, IGSA optimization was applied to determine the weights of the reflective bands (\(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6\) and \(\mu_7\)). To implement IGSA optimization for determining the optimal weights of
different bands applicable to all areas, a pure pixel data set of surface covers and a specific objective function were used. The objective function in the IGSA optimization to determine the scale of the feature space considered (weights of different bands) is defined in Equation (2), where $\Delta ABEI_{M_{i}}$ is the distance (difference) between the ABEI value of the i-th built-up training sample and the built-up cluster centre ($M_{i}$) (Mean ABEI values of built-up samples) and $\Delta ABEI_{M_{nearest}}$ is the distance (difference) between the ABEI value of the i-th built-up training sample and the nearest neighbour cluster centre ($M_{nearest}$) (Mean ABEI values of each land cover samples were calculated (cluster centre for each land cover was obtained). The difference of ABEI value of a built-up sample and other land cover cluster centre was calculated. The minimum value of difference determines the nearest neighbour cluster centre to a built-up sample).

$$F_{obj} = \sum_{i=1,T} \frac{\Delta ABEI_{M_{i}}}{\Delta ABEI_{M_{nearest}}}$$  \hspace{1cm} (2)

In Equation (2), $T$ is the total number of selected training samples for the built-up area. The purpose of IGSA optimization is to search the space for all possible states that allow the coefficient values of $\mu_1$, $\mu_2$, $\mu_3$, $\mu_4$, $\mu_5$, $\mu_6$ and $\mu_7$ to reach the optimal value (minimum) for the objective function. This objective function establishes a balance in the feature space in such a way that the maximum homogeneity of ABEI values between built-up pixels and the maximum difference of the ABEI values between built-up pixels and other land-cover classes can be established. Additionally, by finding the optimum value for this objective function and using the pure pixels of different land-cover classes, the threshold values are also close to each other for different dates and regions. Therefore, the effective feature vectors and their scales can be predicted and we can use them to achieve the best accuracy of the built-up area classification. Main reasons for employing the IGSA algorithm instead of other optimization algorithms in this study are: (a) in many optimization algorithms, the initial population does not cover some parts of problem space. Therefore, optimization algorithms face challenges to find the optimum location in problem space. However, IGSA is able to remove this problem with negative mass; (b) The process time to achieve the optimum coefficients in IGSA compared to other algorithms is short; (c) the IGSA properties such as the memory and using the negative mass significantly decrease the occurrence possibility of local minima and search in a more efficient way the problem space.

3.2.3. Pure-Pixel Selection

To extract the surface reflectance values for 8 different land covers, 7 reflective bands of the Landsat 8 OLI images of the test sites and surrounding areas were sampled. The types of surface covers include built-up, vegetation, bare land, sandy soil, rock, agricultural and water land covers. Given that the spectral values related to the pure pixels of surface covers directly affect the accuracy of the results of the built-up extraction, this stage is one of the most important stages of the study. For each of the surface covers, 150 pure pixels were extracted from 7 reflective bands of the Landsat 8 OLI images.

The set of methods includes photointerpretation from Google Earth images, ground sampling, the pixel purity index (PPI) and selection from the feature space scatter plot used to extract pure pixels of the land-cover types. Spectral separability of the selected land-cover classes was tested by the Jeffries–Matusita distance separability measure [39]. All pairs of land-cover classes, with values ranging from 1.76 to 2.01 were separable from each other.

3.2.4. Improved Gravitational Search Algorithm

In the IGSA, optimization was accomplished through the application of gravitational laws and motion in a synthetic discrete-time system. The system environment is the same as the definition of the problem. According to the law of gravity, every object is able to understand the location and situation of other objects through the law of gravitational attraction. Therefore, this force can be utilized as an instrument to exchange information. The designed optimum finder can be used to solve every optimization problem in which each solution is defined as a situation in space, and its similarity to
the other solutions was expressed as a distance. The mass amounts were determined by the objective function. More detail about IGSA optimization can be found in [40].

3.2.5. Classification, Threshold Optimization and Per-Pixel Accuracy Assessment

To compare the accuracy of the ABEI with those of other indices, in the present study, indices such as the NDBI, modified NDBI, IBI, UI, NBAI, BRBA, NDBaI, BI, EBBI, BAEM, NBI and BUI were applied. The results of applying these indices on the data from the test sites proved that the UI has been more efficient in built-up extraction. For this reason, only the UI performance was compared with the ABEI performance in this study.

According to the ABEI, a default value of the threshold equal to zero is used to separate built-up and non-built-up pixels. Pixels with a value greater than zero are built-up areas, and pixels with values less than zero are non-built-up areas. However, due to different environmental conditions in different regions, the threshold values for both the ABEI and UI indices were determined according to the reference map for test sites and based on the overall error parameter. For this purpose, the threshold value was changed to 0.001, and with each change, the value of the overall error parameter was calculated. Finally, the optimal threshold for each test site was calculated with the condition of minimizing the overall error value. To assess the stability of the ABEI and UI threshold values, the changes in the optimal threshold values for the two indices applied to the test sites were investigated based on relative standard deviation criteria. To compare the accuracies of the ABEI, and UI the overall errors and Kappa coefficients were calculated based on confusion matrices. To compare the ABEI and UI accuracies, the results from the optimal threshold values of each index were used.

Additionally, to investigate whether the ABEI remarkably improves accuracy compared to the UI method at these test sites, the McNemar statistical test was applied. According to Equation (3), McNemar’s continuity-corrected chi-square statistic was calculated as [41]:

\[ X^2 = \frac{(|f_{12} - f_{21}| - 1)^2}{f_{12} - f_{21}} \]  

where \( f_{12} \) and \( f_{21} \) define the frequencies of pixels that were correctly classified by one method but incorrectly classified by the other, respectively.

3.2.6. Mixed Pixel (Heterogeneous Region) Accuracy Assessment

The sensitivity and efficiency of different classifiers for the accurate classification of mixed pixels of built-up and non-built-up lands were evaluated using the overall accuracy assessment for mixed pixels and boundary regions. For this assessment, 1800 mixed pixels of the 5 test sites were selected. Using Google Earth Images for each test site, the percentages of built-up and non-built-up areas for each pixel were computed. Then, the pixels were classified into 10 classes based on the percentage of built-up area. Finally, the percentages of pixels of each class classified by each of the three methods (ABEI and UI) for built-up areas were calculated. Ideally, if the percentage of built-up areas in a pixel is higher than 50%, this pixel should be classified in the built-up class and, conversely, the opposite is classified in the non-built-up class.

3.2.7. The Performance Evaluation of ABEI Using Cross-Validation Analysis

In this study, HR LI reference data were used for the final evaluation of the ABEI performance for mapping surface built-up areas. The results of ABEL were compared with HRLI reference maps of 2015 for European cities. Finally, the overall accuracy of the built-up map extracted by ABEI was computed based on reference map for each city. Also, the performance of ABEI was compared with NDBI and UI for mapping built-up areas in European cities based on reference maps. Finally, the changes of optimal threshold values for built-up extraction with ABEI, UI and NDBI were analyzed for the different European cities.
4. Results

4.1. ABEI

The aim of the ABEI is to create the greatest similarity between pixel values within the built-up class and the greatest difference in pixel values between the built-up class and other classes. In this study, using IGSA optimization, the objective function and pure pixels of different surface land covers and the optimal coefficients for each reflective band were obtained. The obtained ABEI is shown in Equation (4).

\[
ABEI = 0.312\text{Ref}_1 + 0.513\text{Ref}_2 - 0.086\text{Ref}_3 - 0.441\text{Ref}_4 + 0.052\text{Ref}_5 - 0.198\text{Ref}_6 + 0.278\text{Ref}_7
\]  

(4)

According to Equation (4), the coefficient value for each spectral band indicates the extent of the effect of each spectral band on built-up area extraction. It should be noted that in this technique, simultaneously with applying NDVI, water pure pixels (NDVI < 0) and vegetation pure pixels (NDVI > 0.5) were classified into the non-built-up class and in the built-up extraction process with ABEI were not included.

4.2. Built-Up Extraction Maps

The results of the two classification methods for the five test sites were shown in Figure 2. The visual assessment of Figure 2 demonstrates that the accuracy of the built-up extraction map extracted by the ABEI is better than that extracted by the UI. In each of the five test sites, the results of the UI are noisy. However, for the Naqadeh and Babol test sites, there are fewer differences between the classification methods in terms of built-up extraction. However, for the Bam, Kashmar and Masjed Soleyman test sites, due to their bare land, sandy soil and rocky surfaces, there is a remarkable difference between the results of the ABEI and those of the other methods. A visual examination of the results indicates that the increase in ABEI efficiency compared to the UI efficiencies at the arid test sites, including bare land and sandy areas, is more than that at the humid test sites. The main reason is the limitations that other methods have with bare lands. At the Bam site, due to the complexity of the region and the presence of bare lands, sandy soils and desert surfaces, the UI has a very low efficiency in built-up extraction.

<table>
<thead>
<tr>
<th>Site</th>
<th>RGB (7 5 2)</th>
<th>Reference map</th>
<th>ABEI</th>
<th>UI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babol</td>
<td><img src="image1" alt="Babol" /></td>
<td><img src="image2" alt="Babol Reference" /></td>
<td><img src="image3" alt="Babol ABEI" /></td>
<td><img src="image4" alt="Babol UI" /></td>
</tr>
<tr>
<td>Naqadeh</td>
<td><img src="image5" alt="Naqadeh" /></td>
<td><img src="image6" alt="Naqadeh Reference" /></td>
<td><img src="image7" alt="Naqadeh ABEI" /></td>
<td><img src="image8" alt="Naqadeh UI" /></td>
</tr>
</tbody>
</table>

Figure 2. Cont.
Figure 2. Comparison of the built-up extraction results using three classifiers at the five test sites.

The visual assessment of these results illustrates the higher efficiency of the ABEI for built-up extraction under bare land and sandy soil conditions.

4.3. Classification Accuracy and Edge Pixel Effects

The Kappa coefficient and overall accuracy of the built-up maps obtained for each test site were shown in Table 3. For all test sites, the accuracy obtained by the ABEI is higher than that obtained by the UI. According to Figure 3, the average total omission and commission errors at the five test sites by the ABEI, and UI are 15.21, and 58.2%, respectively. Additionally, the standard deviations in the average of total omission and commission errors by the ABEI, and UI methods are 2.9, and 26.08%, respectively. The standard deviation values indicate that the ABEI method error compared to the mean error has a small difference. However, UI errors have large differences. In other words, the results of the accuracy assessment indicate the high stability of the accuracy of the ABEI at all test sites with different environmental conditions and surfaces in contrast to the UI method.
The difference between the accuracies and errors of the built-up extraction methods among the test sites is different. This is because of the differences in the complexities of the surface land cover at various test sites. The difference among the overall accuracies of the built-up extractions by three methods at the Babol and Naqadeh test sites is low, but for the Kashmar and Bam test sites, this difference is very high.

A summary of the classification accuracies of the various methods by test site is shown in Table 4.

Table 4. McNemar chi-square test results for the five test sites.

<table>
<thead>
<tr>
<th>Test Sites</th>
<th>Classifier</th>
<th>$\chi^2$</th>
<th>$p$-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babol</td>
<td>ABEI</td>
<td>26.03</td>
<td>0</td>
</tr>
<tr>
<td>Naqadeh</td>
<td>ABEI</td>
<td>20.04</td>
<td>0</td>
</tr>
<tr>
<td>Masjed Soleyman</td>
<td>ABEI</td>
<td>288</td>
<td>0</td>
</tr>
<tr>
<td>Kashmar</td>
<td>ABEI</td>
<td>749</td>
<td>0</td>
</tr>
<tr>
<td>Bam</td>
<td>ABEI</td>
<td>27.03</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 3. Total classification error (combined commission and omission errors).

A summary of the classification accuracies of the various methods by test site is shown in Table 4.
The McNemar chi-square test results for the five test sites were shown in Table 4. According to the Table 4, the improvement in the accuracy of the ABEI was significant (p-value < 0.01 at the 99% confidence level) compared to those of the UI at all test sites. For test sites where the difference in the built-up extraction accuracy of the ABEI compared to those of other methods is low, the McNemar chi-square value is also low. The results of the subpixel accuracy assessment using various methods were shown in Figure 4.

**Figure 4.** Cumulative frequency of mixed edge pixels classified as built-up pixels.

The results indicate that the abilities of the ABEI, and UI to correctly classify pixels with various percentages of built-up areas are different. In general, with the ABEI, 60% of the mixed pixels were classified correctly into the built-up class. Using the UI method, due to the lack of completely distinguishing bare lands, sandy soils, and rock surfaces from built-up areas, an average of 24.5% of the pixels with less than 50% built-up areas were classified in the built-up class. Additionally, using the UI, an average of 30% of pixels with more than 50% built-up areas were classified into the non-built-up class. However, because of the objective function defined in the optimization algorithm that determines the optimal weight of each spectral band in the ABEI, the capability of this index is high for separating built-up areas from bare lands, sandy soils and rock surfaces. For this reason, with this method, only 4.6% of pixels with less than 50% built-up areas are classified into the built-up class, and on average, 11.2% of pixels with more than 50% built-up areas were classified into the non-built-up class.

4.4. Optimal Threshold and Its Variability

The comparison results for the stability of the optimal threshold values for the ABEI and UI were shown in Figure 5.
The results indicate that the optimal threshold values for the different test sites using the UI compared to those using the ABEI experience more changes. For example, at the Babol and Naqadeh test sites, the optimal threshold value for the UI has changed from \(-0.1\) to \(-0.02\), and that for the ABEI has changed from \(0.3\) to \(0.232\). Generally, the relative standard deviations of the optimal threshold values at the test sites using the UI and ABEI are 155 and 27\%, respectively. The lower relative standard deviation of the optimal threshold values for the ABEI compared to those for the UI represents more stability of the threshold value with spatial and temporal variations.

The advantages of the ABEI include a simple method and systematic technique for the precise separation of built-up areas from other surface land-cover classes, such as bare land, sandy soil and rock surfaces. The advantages of the new index in comparison with other indices in this field are the relatively high stability of the optimum threshold value with respect to the changes in spatial, temporal, and environmental conditions and surface land cover for the separation of built-up areas from non-built-up areas.

### 4.5. Cross-Validation Analysis

The results of the ABEI, UI and NDBI indices for the eleven European cities are shown in Figure 6. Based on the HRLI reference map of each city, the results of indices were evaluated. The blue pixels in each map represent omission errors and the green pixels in each map represent commission errors.
Figure 6. **Cont.**
The visual assessment of Figure 6 demonstrates that the accuracy of the built-up extraction map extracted by the ABEI is better than that extracted by the UI and NDBI. Additionally, there is a significant difference between the results of the ABEI and those of the other indices in case of separation between built-up and bare land areas. A summary of the overall accuracies of the three indices by city was shown in Figure 7.

The mean value of overall accuracy at the eleven European cities using the ABEI, UI and NDBI is 89.3, 72.4 and 64.7%, respectively. The results indicate the high efficiency of ABEI for mapping built-up areas in absolute and relative terms according to UI and NDBI results. The most highest overall error of the UI and NDBI is seen in Madrid (48 and 67%, respectively). In general, the overall accuracies of the built-up extraction using the ABEI are low in areas that covers with bare land. However, ABEI has good performance for separating built-up areas from other surface land-cover classes. The comparison results for the stability of the optimal threshold values for the ABEI, UI and NDBI were shown in Figure 8.
For the five test sites considered in this study, the standard deviations of the overall errors for the ABEI, UI and NDBI are 2.9, 19.66 and 16.08%, respectively. The results indicate that the optimal threshold values for the different European cities using the ABEI compared to those using UI and NDBI experience less changes. The lower changes of the optimal threshold values for the ABEI represent more stability of the threshold value over the different European cities. The results of the performance evaluation of ABEI based on reference data confirmed the high efficiency of this index for mapping built-up areas.

5. Discussion

In previous studies, different indices have been developed for built-up area extraction using satellite imagery [8,9,42]. In different regions, the performances of these indices are affected by the type of surface covers in the region. These indices have problems in distinguishing between surface covers that have similar spectral behaviours. Therefore, these indices are not able to completely isolate built-up and bare lands in the same area [9,29,36]. Additionally, the performance of these indices in regions with dry climates is lower than in humid regions. Even the spectral heterogeneity between pixels related to built-up areas reduces the accuracies of the indices. On the other hand, the optimum threshold value in the previous indices to distinguish built-up areas under different environmental conditions and various surfaces was not stable.

Built-up and bare lands have similar spectral behaviours. On the other hand, the spectral heterogeneity due to the diversity and complexity of urban structures is enormous. Therefore, the indices that only use two bands have problems in distinguishing completely between built-up and bare lands. For these reasons, the development of a new index to consider the spectral diversity of urban structures in built-up extraction process is necessary. In this study, the ABEI was presented for mapping surface built-up areas using Landsat 8 OLI imagery. In ABEI, all of the Landsat 8 OLI reflective bands were employed and optimal weights for each of the bands were obtained based on IGSA optimization. The proposed index has high efficiency for the classification of built-up areas under different environmental conditions.

For the five test sites considered in this study, the standard deviations of the overall accuracy for the ABEI, UI, and NDBI indices...
were 2.2, 10.4 and 15.7, respectively. The difference of the accuracy of the built-up area extraction for different sites using ABEI is very low, but for other methods, is high. In other words, the stability of the ABEI method in different environmental conditions and surface covers is higher than other methods. Because of the defined objective function in the IGSA optimization to determine the optimal weight of each spectral band in ABEI, the performance of this index is high for the separation of built-up area from bare, sand, and rock surfaces (Figures 2 and 6). The values of these coefficients are constant for different regions. Unlike most previous indices, the seven reflective bands of Landsat 8 OLI were used in ABEI to produce built-up maps. For this reason, the ABEI was less influenced by spectral similarities between built-up and bare lands and has better performance compared to previous indices.

The major advantage of the ABEI in addition to its high performance, is a simple implementation. The relative standard deviations of optimal threshold values at 5 test sites for ABEI and UI were 27 and 155%, respectively. Additionally, the relative standard deviations of optimal threshold at eleven European cities for ABEI, UI and NDBI were 16, 37 and 125%, respectively. The lower value of the relative standard deviation of the optimum threshold values for ABEI to the UI and NDBI represents the higher stability of the threshold value under spatial and temporal variations. Additionally, due to the simple implementation of ABEI, this index can be used easily to extract built-up area in different regions. In previous studies, it has been shown that thermal bands can also be used to improve the performance of methods for the built-up areas extraction [42]. For this reason, future studies may consider using the thermal bands in ABEI to improve the performance of this index and comparing the results with our index.

6. Conclusions

The main objective of this study was to develop an automatic technique that improves the accuracy of built-up extraction by increasing spectral separability between built-up and non-built-up surfaces, specifically in regions with bare land and sandy soils, which are often major causes of low classification accuracy. A new Automated Built-up Extraction Index (ABEI) using Landsat 8 OLI images was designed. The aim of the ABEI is to create the greatest similarity between pixel values within the built-up class and the greatest difference between the pixel values in the built-up class and other classes. Therefore, using IGSA optimization, the objective function and pure pixels of different surface land covers, optimal coefficients for each reflective band are obtained. The accuracies and threshold stabilities of the new index are compared with those of the UI, and NDBI methods. The ABEI remarkably improved accuracy in regions where bare lands and other bright surfaces were the major sources of classification errors. A mixed pixel assessment of errors at the edges of the built-up regions demonstrated that the ABEI was more accurate in classifying edge pixels compared to the UI method. Furthermore, the optimal threshold values of the ABEI were indicated to be less changeable under different environmental conditions, locations and times compared to those of the UI. Accordingly, the ABEI is suggested as an improved built-up index, particularly in extracting built-up areas from regions where imprecise results are anticipated due to bare lands, sandy soils and rock surfaces. Additionally, the ABEI technique would be appropriate for built-up area change detection analysis because it separates edge pixels with high accuracy and an optimum threshold. The other advantages of the ABEI include a simple method and systematic technique for the precise separation of built-up areas from other surface land covers, such as bare lands, sandy soils and rock surfaces. Landsat 8 thermal bands are proposed to be used in future studies to increase the accuracy and efficiency of the ABEI index. However, the accuracy of the ABEI can be negatively affected by deep and saltiness water body, and density vegetation areas. Hence, Landsat 8 thermal bands are proposed to be used in future studies to increase the accuracy and efficiency of the ABEI index. Also, we suggest that the efficiency of ABEI and deep learning algorithms be investigated for built-up extraction in future studies.
Author Contributions: M.K.F., A.S., M.K. and S.K.A. conceived and designed the research of the first draft; M.K.F. and A.S. performed data analysis and wrote the first draft; M.K. and S.K.A. edited the pre-draft; S.Q. and D.H. re-designed the research, revised and edited the paper; all authors contributed to and approved the final manuscript.

Funding: This research received no external funding.

Acknowledgments: We are grateful to editor, and three anonymous reviewers for their valuable comments and suggestions to improve this study. Also, the authors deeply thank International Relationship Program, University of Tehran, for their great support in conducting this research.

Conflicts of Interest: The authors declare no conflict of interest.

References


24. Kumar, A.; Pandey, A.C.; Jeyaseelan, A. Built-up and vegetation extraction and density mapping using worldview-ii. *Geocarto Int.* 2012, 27, 557–568. [CrossRef]


30. Moghaddam, M.H.R.; Sedighi, A.; Fayyazi, M.A. Applying mdwi index and linear directional mean analysis for morphological changes in the zarrin-rüd river. *Arab. J. Geosci.* 2015, 8, 8419–8428. [CrossRef]


