The Modified Normalized Urban Area Composite Index: A Satellite-Derived High-Resolution Index for Extracting Urban Areas

Feng Li 1, Xiaoyang Liu 1, Shunbao Liao 1,3 and Peng Jia 2,3,*

Abstract: The accurate and efficient extraction of urban areas is of great significance for better understanding of urban sprawl, built environment, economic activities, and population distribution. Night-Time Light (NTL) data have been widely used to extract urban areas. However, most of the existing NTL indexes are incapable of identifying non-luminous built-up areas. The high-resolution NTL imagery derived from the Luojia 1-01 satellite, with low saturation and the blooming effect, can be used to map urban areas at a finer scale. A new urban spectral index, named the Modified Normalized Urban Areas Composite Index (MNUACI), improved upon the existing Normalized Urban Areas Composite Index (NUACI), was proposed in this study, which integrated the Human Settlement Index (HSI) generated from Luojia 1-01 NTL data, the Normalized Difference Vegetation Index (NDVI) from Landsat 8 imagery, and the Modified Normalized Difference Water Index (MNDWI). Our results indicated that MNUACI improved the spatial variability and differentiation of urban components by eliminating the NTL blooming effect and increasing the variation of the nighttime luminosity. Compared to urban area classification from Landsat 8 data, the MNUACI yielded better accuracy than NTL, NUACI, HSI, and the EVI-Adjusted NTL Index (EANTLI) alone. Furthermore, the quadratic polynomial regression analysis showed the model based on MNUACI had the best $R^2$ and Root-Mean Square Error (RMSE) compared with NTL, NUACI, HSI, and EANTLI in terms of estimation of impervious surface area. It is concluded that MNUACI could improve the identification of urban areas and non-luminous built-up areas with better accuracy.

Keywords: nighttime light; Luojia 1-01; MNUACI; urban area; urban remote sensing

1. Introduction

Urban areas are the supporting systems of urban population, built-ups, transportation, and commercial corporations, as well as where urbanization takes place [1]. Urbanization and urban sprawl have important influence on the urban ecological environment, climate, public health, and socioeconomic development through the transformation of land use/cover types [2–6]. Urban expansion brings a series of urban problems, such as underground water pollution, traffic congestion, carbon emission increment, urban heat island effect, etc. Furthermore, urban nighttime light (NTL) causes disturbance to the human circadian rhythm and sleep disorders [7]. These urbanization issues increase the burden on the urban ecological system and impact the sustainable development of cities, especially in developing countries like China. Due to China’s opening and economic development strategies, the urbanization rate of China’s permanent population has increased from 18.96% in 1979 to 60.60% in 2019 [8], as shown in Figure 1. This approximately exponential urban population growth has greatly promoted the accelerated expansion of China’s cities. Faced with problems caused by rapid expansion of China’s cities, it is...
essential to design relevant analytical methods to solve these regular mapping problems of urban sprawl.

![Graph showing urbanization rate from 1979 to 2019.](image_url)

**Figure 1.** China’s urbanization rate from 1979 to 2019.

The NTL data have been most widely used to extract urban areas at regional and global scales, such as the Defense Meteorological Satellite Program/Operational Linescan System (DMSP/OLS) and the Suomi National Polar-orbiting Partnership satellite with Visible Infrared Imaging Radiometer Suite (NPP-VIIRS) data [9–12]. The main approaches for identification of urban areas through NTL data include edge-detection, supervised classification, and threshold-based segmentation. Assuming the existence of abrupt changes of NTL in urban and suburban transition zones, Tan used a light intensity gradient to delineate the boundaries of urban areas [13]. Xue et al. adopted an edge detection method to acquire urban boundaries based on the Vegetation Adjusted NTL Urban Index (VANUI) [14].

In supervised classification methods, many studies have proved that the support vector machine (SVM) method could provide high-precision classification results. Cao et al. developed an SVM-based region-growing algorithm to distinguish urban areas from non-urban background [15]. Jing et al. proved that results obtained by k-nearest-neighbors, SVM, and the random forests classification algorithm could achieve a better agreement for the purpose of urban area detection [16]. In addition, simple thresholding methods were also usually adopted to extract urban area extent [17,18].

The integration of multi-source remote sensing data and NTL data can mitigate the blooming effect (i.e., adjacent pixels of pervious surface are usually counted as impervious surface due to similar NTL values) of NTL and improve its performance. Integrating urban NTL with vegetation index and land surface temperature (LST) provides promising approaches to differentiate urban areas from non-urban areas. For example, Lu et al. developed a Human Settlement Index (HSI) with DMSP-OLS and the normalized difference vegetation index (NDVI) data to map urban settlements [19]. Zhang et al. proposed a simple and effective VANUI that could reduce the effects of NTL saturation and overcome the overcorrection issue of HSI [20]. Zhuo et al. proposed an Enhanced Vegetation Index (EVI) Adjusted NTL Index (EANTLI) to lessen the saturation problem of NTL data [21]. Liu et al. demonstrated that an LST and EVI Regulated NTL City Index (LERNCI) was more effective in delineating the urban spatial structures than VANUI [22]. However, when land cover types include water body and bare land, it is not enough to rely solely on vegetation index or land surface temperature to separate urban area from non-urban area. Therefore, Liu et al. established a Normalized Urban Areas Composite Index (NUACI) by combining the Normalized Difference Water Index (NDWI), NTL, and the EVI to estimate the urban impervious surface [23]. NUACI could degrade oversaturation using water body and vegetation indexes, but it might still ignore certain urban areas in low-luminous areas.

A Modified NDWI (MNDWI) was found to be more suitable for water feature recognition than the NDWI, as it can better suppress non-water land noise and better enhance
Thus, this study aimed to develop a Modified NUACI (MNUACI) on the basis of the MNDWI, NDVI, and NTL, to further improve the identification of urban areas. Moreover, the MNUACI has a higher spatial resolution than the previous DMSP/OLS- and NPP-VIIRS-based indexes by using NTL data from the Luojia 1-01 NTL satellite, designed and developed by Wuhan University in China, which has started providing nighttime imagery with a finer resolution of 130 m since 2018. The MNUACI would be useful for a wide array of urban studies, such as urban population health [25], urban spatial structure [26], and energy carbon emissions [27], where such an index has been urgently demanded.

2. Study Sites and Data Sources

2.1. Study Sites

As shown in Figure 2, four capital cities in China from north to south, Beijing, Nanjing, Guangzhou and Haikou, were chosen as study sites. Beijing, the political, science and technology innovation and cultural center of China, is surrounded by mountains in the west, north and northeast, and the North China Plain in the southeast [28]. It has sixteen municipal districts with a total area of 16,410 km$^2$, a resident population of 21.54 million and a GDP of 3032 billion CNY in 2018 [29]. With six downtown areas as the center and Tongzhou District as the sub-center, Beijing is expanding outward along east-west and north-south axes, and the urban secondary industry and a large labor force is beginning to migrate to the surrounding suburbs. Nanjing, a regional transportation hub, is in the lower reaches of the Yangtze River and is the only megacity in the Yangtze River Delta and East China. It includes twelve administrative districts, with a total area of 6587 km$^2$, a resident population of 8.44 million and a GDP of 1282 billion CNY in 2018 [30]. New urban areas and towns were built along two banks of the Yangtze River, and the central city expanded northward and southward to the suburbs, where the chemical industry was centrally located. Guangzhou, located at the northern edge of the pearl river delta, is an important transportation and logistics hub in South China. It consists of eleven administration districts, with a resident population of 14.9 million and a GDP of 2286 billion CNY in 2018, covering a total area of 7434 km$^2$ [31,32]. With the metropolitan area as the city center, Guangzhou built two new districts in the south and east regions, and three sub-centers in the north and east regions. The international tourist city of Haikou borders the Qiongzhou Strait in the north and serves as the core city of the China Free Trade Zone. Haikou covers a land area of 2290 km$^2$ and a sea area of 861 km$^2$, with a population of 2.3 million and a GDP of 151.1 billion CNY in 2018 [33]. Haikou has built the east and west coast tourism belt and the north-south tourism axis of Nandu River urban water system.

The traditional economic growth points are mostly located in the central areas of China cities. Based on the strong economic foundation, the latest urban planning and layouts of the above four cities guide and drive the flow of the large labor force to the new industrial agglomeration areas in suburbs, thus promoting the continuous urban expansion of these cities. Besides, these four cities have experienced tremendous socioeconomic development over the past 40 years, representing the natural and socioeconomic development levels of different cities in China. Therefore, they are suitable to study the dynamics of urbanization and its impact on urban ecosystems according to the extent of urban areas extraction.
2.2. Data Sources

The cloud-free Luojia 1-01 NTL images in Beijing, Nanjing, Guangzhou and Haikou, which are dated 6 September 2018, 15 July 2018, 4 September 2018 and 5 September 2018, respectively, as illustrated in Figure 3 and Table 1, were downloaded for free from the Hubei Data and Application Center of High-Resolution Earth Observation System website (http://59.175.109.173:8888, accessed on 1 June 2021). The Luojia 1-01 NTL satellite, designed and developed by Wuhan University in China, has started to provide nighttime imagery with a finer resolution of 130 m since 2018. This satellite sensor records with 14-bits radiometric resolution and improves on-board calibration functions, which demonstrates finer spatial detail and urban spatial structure than DMSP/OLS and NPP-VIIRS data [34,35]. Landsat 8 Operational Land Imager (OLI) images with minimum cloud cover for Beijing, Nanjing, Guangzhou and Haikou on 23 October 2017, 6 June 2018, 23 October 2017, and 17 May 2018 were obtained from the United States Geological Survey (USGS) website (https://glovis.usgs.gov, accessed on 1 May 2021), as illustrated in Table 1. Except for the coastal/aerosol band and the cirrus band, OLI inherits the seven bands of the Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) sensors, improving image measurement abilities and offering compatibility with the historical Landsat images. The Landsat 8 multi-spectral imagery ranging from Band 2 to Band 7 (Blue, Green, Red, NIR, SWIR1, SWIR2, respectively) were used for relevant vegetation index calculation, water body index calculation and image classification.

Figure 2. Geolocation map of study sites (Beijing, Nanjing, Guangzhou, Haikou).
Landsat 8 multi‐spectral imagery ranging from Band 2 to Band 7 (Blue, Green, Red, NIR, SWIR1, SWIR2, respectively) were used for relevant vegetation index calculation, water body index calculation and image classification.

Figure 3. Nighttime light images of Luojia 1-01 in (a) Guangzhou, (b) Nanjing, (c) Beijing, and (d) Haikou.

Table 1. Research dataset description.

<table>
<thead>
<tr>
<th>Image Type</th>
<th>Acquisition Date</th>
<th>Day/Night</th>
<th>Spatial Resolution (m)</th>
<th>Path &amp; Row</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luojia 1-01</td>
<td>6 September 2018</td>
<td>Night</td>
<td>130</td>
<td>1423 &amp; 28</td>
<td>Beijing</td>
</tr>
<tr>
<td>Luojia 1-01</td>
<td>15 July 2018</td>
<td>Night</td>
<td>130</td>
<td>8979 &amp; 18</td>
<td>Nanjing</td>
</tr>
<tr>
<td>Luojia 1-01</td>
<td>4 September 2018</td>
<td>Night</td>
<td>130</td>
<td>6005 &amp; 05</td>
<td>Guangzhou</td>
</tr>
<tr>
<td>Luojia 1-01</td>
<td>5 September 2018</td>
<td>Night</td>
<td>130</td>
<td>7644 &amp; 05</td>
<td>Haikou</td>
</tr>
<tr>
<td>Landsat 8</td>
<td>23 October 2017</td>
<td>Day</td>
<td>30</td>
<td>123 &amp; 32</td>
<td>Beijing</td>
</tr>
<tr>
<td>Landsat 8</td>
<td>6 June 2018</td>
<td>Day</td>
<td>30</td>
<td>120 &amp; 38</td>
<td>Nanjing</td>
</tr>
<tr>
<td>Landsat 8</td>
<td>23 October 2017</td>
<td>Day</td>
<td>30</td>
<td>122 &amp; 44</td>
<td>Guangzhou</td>
</tr>
<tr>
<td>Landsat 8</td>
<td>17 May 2018</td>
<td>Day</td>
<td>30</td>
<td>124 &amp; 46</td>
<td>Haikou</td>
</tr>
</tbody>
</table>

3. Methodology

Figure 4 demonstrates the methodological framework, which involves four main steps. First, basic preprocessing such as geometric rectification, reprojection, and atmospheric correction was executed for the Luojia 1-10 NTL and Landsat 8 OLI raw data. Next, the MNUACI model was developed by integrating NTL, HSI, NDVI and MNDWI data. Third, four classic classification methods were applied to NTL, HSI, NDVI and MNDWI data for mapping urban areas of study sites. Lastly, the reference urban mapping results based on Landsat 8 data were used to evaluate the accuracy of the MNUACI model.
3.1. Data Preprocessing

The positioning accuracy of Luojia 1-01 NTL imagery was reported as approximately 800 m by executing the on-board geometric calibration method [36], but its results are far from meeting the actual positioning requirements. For each Luojia 1-01 image, at least forty ground control points (GCPs) were selected at road intersections on Landsat 8 images, and a geometric correction was carried out on Luojia 1-01 images through an affine transformation. The geometric accuracy of the final correction error of each image was controlled within half of a Luojia 1-01 image pixel, namely 65 m. Landsat 8 OLI records not only the reflected and emitted radiation from the earth’s surface, but also the radiation scattered or emitted by the atmospheric layer. To quantize the real reflectance from the earth’s surface, the Radiometric Calibration Tool and the Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) module in the ENVI software were used to convert the DNs of raw images into surface reflectance values. The purpose of atmospheric correction is to eliminate the influence of atmosphere and solar illumination and obtain the correct surface reflectance parameters of the earth’s surface. After atmospheric correction, satellite images can improve the ability of data analysis. To integrate them better with images from Landsat 8 OLI, Landsat 8 images were resampled to the same 130 m-resolution as the Luojia 1-01 nighttime images after the images underwent atmospheric correction.

3.2. Modified Normalized Urban Area Composite Index (MNUACI)

Water and vegetation index can effectively differentiate water body and vegetation types from urban areas [37]. Based on the characteristics of the two indexes, the NUACI, established by integrating water index, vegetation index and NTL data, can be used to recognize urban areas, as shown in the following equation:

$$NUACI = \begin{cases} 0, & d > r, \\ \sqrt{(NDWI - \bar{NDWI})^2 + (EVI - \bar{EVI})^2}, & d \leq r \\ \end{cases} \quad (1 - d/r) * NTL_{norm}, \quad d \leq r$$
where $\bar{\text{NDWI}}$ and $\bar{\text{EVI}}$ indicate the average NDWI and EVI from the urban impervious surface, respectively; $d$ and $r$ denote the distance and maximum radius of the circle from urban core, respectively; and $\text{NTL}_{\text{norm}}$ represents the normalized NTL which is expressed with the following equation:

$$\text{NTL}_{\text{norm}} = \frac{\text{NTL} - \text{NTL}_{\text{min}}}{\text{NTL}_{\text{max}} - \text{NTL}_{\text{min}}}$$  \hfill (2)

where $\text{NTL}_{\text{min}}$ and $\text{NTL}_{\text{max}}$ are the minimum and maximum values of NTL DN values, respectively.

Equation (1) reveals that the integration of NDWI and EVI can eliminate the blooming effect of NTL when $d$ is greater than $r$, while the saturation effect of NTL can be mitigated when $d$ is less than $r$. As illustrated in Figure 5, in the absence of NTL, NUACI is unable to detect impervious surfaces, such as buildings and roads in urban areas. To make up for the drawback of NUACI, HSI was introduced to reinforce the NTL effect of urban impervious surfaces. In addition, compared with NDWI, MNDWI is more suitable for water identification of water bodies under urban background due to its merits of suppressing noises from bare soil and built-up areas [38]. Considering the advantages of HSI and MNDWI, a modified urban index MNUACI is constructed by integrating them, as expressed in the following equation:

$$\text{MNUACI} = \begin{cases} 
0, & d > r, \\
\frac{(\text{MNDWI} - \bar{\text{MNDWI}})^2 + (\text{NDVI} - \bar{\text{NDVI}})^2}{(1 - d/r) \times \text{HSI}}, & d \leq r 
\end{cases}$$  \hfill (3)

where $\bar{\text{MNDWI}}$ and $\bar{\text{NDVI}}$ indicate the average MNDWI and NDVI from urban impervious surfaces, respectively. HSI can be expressed with following equation:

$$\text{HSI} = \frac{(1 - \text{NDVI}_{\text{norm}}) + \text{NTL}_{\text{norm}}}{1 - \text{NTL}_{\text{norm}} + \text{NDVI}_{\text{norm}} + \text{NDVI}_{\text{norm}} \times \text{NTL}_{\text{norm}}}$$  \hfill (4)

where $\text{NDVI}_{\text{norm}}$ represents the normalized NDVI, and its normalized method is the same as $\text{NTL}_{\text{norm}}$.

The equation of MNDWI is expressed as follows:

$$\text{MNDWI} = \frac{G - \text{MIR}}{G + \text{MIR}}$$  \hfill (5)

where $G$ and $\text{MIR}$ denote the green band and mid-infrared band of Landsat 8 OLI images, separately.
Remote Sens. 2021, 13, x FOR PEER REVIEW

Figure 5. Urban areas in the Temple of Heaven Park in Beijing, extracted from (a) Bands 5, 4, 2 composite image of Landsat 8, (b) normalized Night-Time Light (NTL), (c) Enhanced Vegetation Index Adjusted Nighttime Light Index (EANTLI), (d) Human Settlement Index (HSI), (e) Normalized Urban Areas Composite Index (NUACI), and (f) Modified Normalized Urban Areas Composite Index (MNUACI).

3.3. Accuracy Analysis Methods

A confusion (error) matrix is an effective quantitative method of characterizing accuracies of land use/land cover types in image classification results. The Commission Errors (CE) are mistakes where results erroneously included in consideration when they should be excluded. The Omission Errors (OE) are mistakes where results are erroneously excluded from consideration when they should have been included. Overall Accuracy (OA) is essentially what percentage of all reference data is correctly classified. The Kappa Coefficient (KC) is a statistic measure of inter-rater reliability or intra-rater reliability for qualitative (categorical) items [39–41]. The Jaccard Similarity coefficient (JSC) refers to a statistic used for gauging the similarity and diversity of sample sets [42].

By calculating the CE, OE, OA, KC and JSC of the reference data and the user classification data, the consistency between both datasets can be evaluated. The detailed equations are as follows:
\[ OE = 1 - \frac{n_{tt}}{n_{rt}} \quad (6) \]

\[ CE = 1 - \frac{n_{tt}}{n_{ut}} \quad (7) \]

\[ OA = \frac{\sum n_{tt}}{N} \quad (8) \]

\[ KC = \frac{N \sum_{t=1}^{c} n_{tt} - \sum_{t=1}^{c} n_{rt}n_{ut}}{N^2 - \sum_{t=1}^{c} n_{rt}n_{ut}} \quad (9) \]

\[ JSI = \frac{|U \cap R|}{|U| + |R| - |U \cup R|} \quad (10) \]

where \( n_{tt} \) refers to pixel numbers correctly classified in type \( t \); \( n_{ut} \) refers to pixel numbers of type \( t \) in user classification data; \( n_{rt} \) refers to pixel numbers of type \( t \) in the reference data; \( c \) refers to the number of all types; \( N \) refers to total pixel numbers in all types; and \( U \) and \( R \) refer to the user classification dataset and reference dataset, respectively.

### 3.4. Estimation of Urban Impermeable Surface

The Impervious Surface Area (ISA) is considered to be an important indicator to measure the degree of urbanization. Previous studies have confirmed a positive correlation between ISA and urban NTL data \cite{43,44}, therefore, ISA can be used as an evaluation indicator for extraction results of urban areas. Using the blue and near-infrared bands of Landsat 8 images, the Perpendicular Impervious Surface Index (PISI) was derived and used to represent the ISA \cite{45}. The extraction accuracy for the ISA based on PISI ranged from 89.51% to 96.50% in the four China cities, which demonstrated a better separability for ISA and bare soil. The ISA can be derived by following equation:

\[ ISA = 0.8192\rho_{Blue} - 0.5735\rho_{NIR} + 0.0750 \quad (11) \]

where \( \rho_{Blue} \) and \( \rho_{NIR} \) denote the reflectance of the blue band and near-infrared band from a Landsat 8 image.

### 4. Results

#### 4.1. Urban Area Extraction by the MNUACI

Landsat 8 multi-spectral reflectance data from the four capital cities in China were adopted to calculate MNDWI and NDVI. The MNUACI was then derived by integrating MNDWI, NDVI and Luojia 1-01 NTL. Before performing the calculation for MNUACI, the parameters \( a_{MNDWI} \) and \( b_{NDVI} \) were determined by Equation (3) based on samples collected from the urban cores. The parameter \( r \) was calculated according to the farthest distance between \( (a_{MNDWI}, b_{NDVI}) \) and \( (a_{MNDWI}, b_{NDVI}) \) from the sample data. The parameters \( (a_{MNDWI}, b_{NDVI}, r) \) from Beijing, Nanjing, Guangzhou and Haikou are (0.73, 0.54, 0.35), (0.37, 0.55, 0.51), (0.41, 0.46, 0.32) and (0.36, 0.36, 0.49), respectively.

MNUACI is used to distinguish between light-intensity differences in urban core areas, and therefore, to improve pixel resolution in light-saturated areas and allow recognition of urban core structures. The Temple of Heaven Park in Beijing and Hongcheng Lake in Haikou were selected to evaluate the effectiveness of MNUACI. As illustrated in Figure 5, the urban areas (cyan) extracted from Landsat 8 were regarded as reference data, to which NTL, EANTLI, HSI, NUACI, and MNUACI were compared. It can be seen that NTL and EANTLI have similar results: neither buildings nor roads are recognized in the middle of the park due to the lack of nighttime luminosity. Although NUACI shows more fractions of urban areas, only a small amount of impervious surface in the park can be recognized, due to the lack of a sufficient luminous condition. HSI and MNUACI identified detail structures of impervious surfaces, but MNUACI extracted impervious surfaces more accurately.
As illustrated in Figure 6, NTL and EANTLI mistakenly identify most urban areas as pervious surfaces and increase omission errors. Although HSI identifies more urban areas, it recognizes the lake as urban areas by mistake, resulting in many commission errors. The recognition results of urban areas from NUACI and MNUACI are similar, both showing the detailed urban structure. However, MNUACI exhibits higher accuracy of urban areas extraction resulting from the reduction of the impact of water bodies on urban areas using MNDWI.

Figure 6. Urban areas in the Hongcheng Lake in Haikou, extracted from (a) Bands 5, 4, 2 composite image of Landsat 8, (b) normalized Night-Time Light (NTL), (c) Enhanced Vegetation Index Adjusted Nighttime Light Index (EANTLI), (d) Human Settlement Index (HSI), (e) Normalized Urban Areas Composite Index (NUACI), and (f) Modified Normalized Urban Areas Composite Index (MNUACI).

Taking Nanjing as an example, Figure 7 illustrates a latitudinal transect of NTL, NUACI and MNUACI. These three types of curve variation are similar, but DN values of MNUACI and NUACI in urban areas are distinctly higher than those of NTL, which suggests that MNUACI and NUACI can enhance the NTL effect in urban areas. For urban areas, MNUACI has higher peaks and lower valleys than NUACI, which reflects more characteristics of inner-urban variability and spatial differentiation. This suggests an easier process of urban area extraction using MNUACI. For peri-urban areas, NUACI and NTL present similar low DN values, proving it difficult to identify small towns with them. In contrast, DN values are higher when MNUACI extracts urban areas in suburban regions. In addition, NTL cannot eliminate blooming effects due to a small quantity of luminosity...
values occurring in water and vegetation areas, while MNUACI and NUACI solve these blooming problems by introducing vegetation and water indexes.

![Image](https://via.placeholder.com/150)

**Figure 7.** Night-Time Light (NTL), Normalized Urban Areas Composite Index (NUACI), and Modified Normalized Urban Areas Composite Index (MNUACI) along a longitudinal transection in Nanjing.

### 4.2. Performance Assessment of the MNUACI

In terms of urban area recognition, a combination of NTL and auxiliary data is better than the use of NTL alone. Different extraction methods for urban areas demonstrate different performances on the same composited NTL index [46]. To assess the feasibility and effectiveness of MNUACI, several supervised and unsupervised classification approaches were separately applied to identify urban areas on NTL, EANTLI, HSI, NUACI and MNUACI images. Because the optimal thresholding method is time-consuming and laborious, the genetic algorithm (GA) was used instead of automatically determining the image segmentation threshold for the extraction of urban areas [47]. Deep learning (DL), GA, fuzzy C-means (FCM) and SVM methods were used to extract urban areas from NTL, EANTLI, HSI, NUACI and MNUACI images of Beijing, Nanjing, Guangzhou and Haikou. Moreover, urban area references of the four sample cities were obtained from Landsat 8 images using the maximum likelihood classifier (MLC) method. Corresponding urban areas from each city were derived from NTL, EANTLI, HSI, NUACI and MNUACI images associated with DL, GA, FCM and SVM approaches. The point-to-point comparison method was applied to test images and reference images. Then, precision indicators such as the Kappa coefficient, overall accuracy and the Jaccard similarity index were calculated to analyze the performance of the combination of different indexes and approaches.
The OA and KC of reference images from Beijing, Nanjing, Guangzhou and Haikou are (93.42%, 0.87), (95.88%, 0.92), (98.16%, 0.96) and (96.94%, 0.94), respectively, which suggests that the classification accuracies of Landsat 8 images from the four cities are reliable.

As illustrated in Tables 2–5, based on the OA, KC and JSC of five NTL indexes, the order of accuracy of urban area classification in Beijing, Nanjing, Guangzhou and Haikou respectively are: MNUACI > HSI > NUACI > NTL > EANTLI, MNUACI > NUACI > HSI > EANTLI, MNUACI > NUACI > HSI > NTL > EANTLI. MNUACI has a higher classification accuracy than the other four NTL indexes of the four capital cities using four classification approaches. Based on the OA, KC and JSC of the four classification methods in MNUACI, each SVM demonstrates the highest classification accuracy in the four urban area classification methods. Except the classification accuracy of NUACI which ranks third in Beijing, each NUACI accuracy from the other three cities follows the corresponding MNUACI. For MNUACI, the accuracy relationship of the four urban area extraction approaches in Beijing, Nanjing, Guangzhou and Haikou are as follows: SVM > GA > FCM > DL, SVM > DL > GA > FCM, SVM > GA > FCM > DL. The SVM method is superior to other methods with the GA method being the second, the FCM method the third, and the DL method the last.

Table 2. Accuracy comparison among various methods of urban area extraction methods using different nighttime light indexes in Beijing.

<table>
<thead>
<tr>
<th>Method</th>
<th>MNUACI</th>
<th>NUACI</th>
<th>HSI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OA</td>
<td>KC</td>
<td>JSC</td>
</tr>
<tr>
<td>DL</td>
<td>91.31%</td>
<td>0.826</td>
<td>0.843</td>
</tr>
<tr>
<td>GA</td>
<td>92.33%</td>
<td>0.847</td>
<td>0.861</td>
</tr>
<tr>
<td>FCM</td>
<td>92.50%</td>
<td>0.850</td>
<td>0.864</td>
</tr>
<tr>
<td>SVM</td>
<td>93.36%</td>
<td>0.867</td>
<td>0.883</td>
</tr>
<tr>
<td></td>
<td>EANTLI</td>
<td>NTL</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OA</td>
<td>KC</td>
<td>JSC</td>
</tr>
<tr>
<td>DL</td>
<td>56.56%</td>
<td>0.149</td>
<td>0.156</td>
</tr>
<tr>
<td>GA</td>
<td>64.74%</td>
<td>0.306</td>
<td>0.319</td>
</tr>
<tr>
<td>FCM</td>
<td>80.07%</td>
<td>0.604</td>
<td>0.632</td>
</tr>
<tr>
<td>SVM</td>
<td>84.50%</td>
<td>0.691</td>
<td>0.722</td>
</tr>
</tbody>
</table>

OA: Overall Accuracy; KC: Kappa Coefficient; JSC: Jaccard Similarity Index.

Table 3. Accuracy comparison among various methods of urban area extraction methods using different nighttime light indexes in Nanjing.

<table>
<thead>
<tr>
<th>Method</th>
<th>MNUACI</th>
<th>NUACI</th>
<th>HSI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OA</td>
<td>KC</td>
<td>JSC</td>
</tr>
<tr>
<td>DL</td>
<td>94.71%</td>
<td>0.893</td>
<td>0.888</td>
</tr>
<tr>
<td>GA</td>
<td>93.99%</td>
<td>0.879</td>
<td>0.878</td>
</tr>
<tr>
<td>FCM</td>
<td>80.07%</td>
<td>0.604</td>
<td>0.632</td>
</tr>
<tr>
<td>SVM</td>
<td>84.50%</td>
<td>0.691</td>
<td>0.722</td>
</tr>
<tr>
<td></td>
<td>EANTLI</td>
<td>NTL</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OA</td>
<td>KC</td>
<td>JSC</td>
</tr>
<tr>
<td>DL</td>
<td>72.60%</td>
<td>0.430</td>
<td>0.415</td>
</tr>
<tr>
<td>GA</td>
<td>65.87%</td>
<td>0.284</td>
<td>0.272</td>
</tr>
<tr>
<td>FCM</td>
<td>90.14%</td>
<td>0.800</td>
<td>0.791</td>
</tr>
<tr>
<td>SVM</td>
<td>89.66%</td>
<td>0.790</td>
<td>0.781</td>
</tr>
</tbody>
</table>

OA: Overall Accuracy; KC: Kappa Coefficient; JSC: Jaccard Similarity Index.
Table 4. Accuracy comparison among various methods of urban area extraction methods using different nighttime light indexes in Guangzhou.

<table>
<thead>
<tr>
<th>Method</th>
<th>MNUACI</th>
<th>NUACI</th>
<th>HSI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OA</td>
<td>KC</td>
<td>JSC</td>
</tr>
<tr>
<td>DL</td>
<td>86.22%</td>
<td>0.729</td>
<td>0.759</td>
</tr>
<tr>
<td>GA</td>
<td>94.06%</td>
<td>0.879</td>
<td>0.901</td>
</tr>
<tr>
<td>FCM</td>
<td>93.11%</td>
<td>0.860</td>
<td>0.885</td>
</tr>
<tr>
<td>SVM</td>
<td>94.06%</td>
<td>0.879</td>
<td>0.901</td>
</tr>
</tbody>
</table>

OA: Overall Accuracy; KC: Kappa Coefficient; JSC: Jaccard Similarity Index.

Table 5. Accuracy comparison among various methods of urban area extraction methods using different nighttime light indexes in Haikou.

<table>
<thead>
<tr>
<th>Method</th>
<th>MNUACI</th>
<th>NUACI</th>
<th>HSI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OA</td>
<td>KC</td>
<td>JSC</td>
</tr>
<tr>
<td>DL</td>
<td>57.01%</td>
<td>0.213</td>
<td>0.255</td>
</tr>
<tr>
<td>GA</td>
<td>83.61%</td>
<td>0.670</td>
<td>0.740</td>
</tr>
<tr>
<td>FCM</td>
<td>72.21%</td>
<td>0.465</td>
<td>0.538</td>
</tr>
<tr>
<td>SVM</td>
<td>76.72%</td>
<td>0.545</td>
<td>0.616</td>
</tr>
</tbody>
</table>

OA: Overall Accuracy; KC: Kappa Coefficient; JSC: Jaccard Similarity Index.

After applying the SVM method, the spatial distribution of the extraction accuracy of urban areas, commission errors and omission errors from Beijing, Nanjing, Guangzhou and Haikou are displayed in Figure 8. The results of EANTLI and NTL produce a great deal of omission errors on some peri-urban areas lacking in nighttime luminosity. This might be due to the Luojia 1-01 satellite imaging time set at 2:00–3:00 a.m. local time. The primary errors of HSI for the extraction of urban areas are commission errors caused by a large number of water bodies. Although NUACI improves the accuracy of urban area extraction by integrating vegetation and water bodies, NUACI, like EANTLI and NTL, still has difficulty identifying unlit urban areas due to the use of NTL alone. All results of MNUACI in the term of extraction of urban areas illustrate lower commission errors and omission errors contrasting to results of NUACI, HSI, EANTLI and NTL. Moreover, the spatial distribution type of MNUACI results is also closer to the MLC results.
Remote Sens. 2021, 13, x FOR PEER REVIEW ..., and even HSI mistakenly identifies rivers as urban areas. In contrast, MNUACI and NUACI do not generate such errors.

Figure 8. Accuracy comparison of urban area extraction using an SVM method on the basis of the Modified Normalized Urban Areas Composite Index (MNUACI), Normalized Urban Areas Composite Index (NUACI), Human Settlement Index (HSI), EVI-Adjusted NTL Index (EANTLI), and Night-Time Light (NTL) in (a) Beijing, (b) Nanjing, (c) Guangzhou, and (d) Haikou.

The extraction results of urban areas in the Tongzhou District of Beijing based on MNUACI, NUACI, HSI, EANTLI and NTL by the SVM method are shown in Figure 9, and a Landsat 8 false color composite image (Figure 9a) is used as a visual reference for urban areas. For two central city areas, MNUACI and HSI show specific spatial distribution patterns and inner-urban differentiation. NUACI and EANTLI extracted non-vegetation and illuminated regions as urban areas, while NTL extracted only illuminated regions as urban areas. For two town areas, NTL, EANTLI and NUACI merely identify road areas within them, missing most town areas, especially for Town II, while MNUACI and HSI recognize more urban areas. For bare land area, NTL, EANTLI and NUACI merely identify minor bare lands within them while MNUACI and HSI recognize most bare lands. For construction sites, NTL almost identifies the whole construction sites as urban areas correctly, among which MNUACI have the best extraction effect. For village areas, the results of urban areas identified by the five indexes are similar, but for Village I, NTL, EANTLI and NUACI, lead to a large number of omission errors, while the results of HSI and MNUACI generate minimum omission errors. Moreover, NTL and EANTLI produce slight commission errors over the river area, and even HSI mistakenly identifies rivers as urban areas. In contrast, MNUACI and NUACI do not generate such errors.
4.3. Correlation between MNUACI and Urban Impervious Surface

Furthermore, one thousand sample points from an ISA image and corresponding MNUACI, NUACI, HSI, EANTLI and NTL images in each city were randomly selected by using the Create Random Points tool as well as the Extract Multi Values to Points tool of the ArcGIS software. The quadratic polynomial regression models were subsequently established with MNUACI, NUACI, HSI, EANTLI and NTL for estimation of ISA. Correlation coefficients and Root-Mean Square Error (RMSE) were employed together to evaluate the performance of the established regression models.

As shown in Table 6, the average $R^2$ and RMSE of MNUACI, NUACI, HSI, EANTLI and NTL in Beijing, Nanjing, Guangzhou and Haikou are (0.74, 0.13), (0.49, 0.18), (0.44, 0.19), (0.21, 0.22) and (0.24, 0.22), respectively. According to correlation coefficients and the RMSE of quadratic polynomial regression models, the results of EANTLI and NTL have similar lower fitting accuracy, and the result of HSI is better than that of the previous two indexes. Apart from the result of Beijing, model regression effects of NUACI in the other three cities are better than EANTLI and NTL. In contrast, MNUACI shows the highest correlation coefficients and the lowest RMSE in all four cities. This suggests that the regression model of MNUACI could enormously decrease the blooming effect of Luojia 1-01 NTL and improve identification accuracy for non-luminous ISA better than for other models. As illustrated in Figure 10, the scatter plots indicate that regression models between MNUACI and ISA in Beijing, Nanjing, and Guangzhou demonstrate the form of a quadratic polynomial regression model, whereas the polynomial regression model at Haikou is closer to a linear regression model. The NTL of urban core areas in developed metropolitan cities, such as
Beijing, can contribute to the saturated MNUACI value. The ISA corresponding to PISI might not be the highest, because the differentiation between the blue and the near-infrared band during the daytime is weakened in the urban core area. On the contrary, the NTL of urban core areas in developing cities, such as Haikou, might rarely generate a saturated MNUACI value, which can present a good linear correspondence to ISA derived from a multispectral image.

Table 6. Correlation coefficients and RMSE of regression models for estimating impervious surface areas.

<table>
<thead>
<tr>
<th>Cities</th>
<th>MNUACI</th>
<th>NUACI</th>
<th>HSI</th>
<th>EANTLI</th>
<th>NTL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
<td>RMSE</td>
<td>$R^2$</td>
<td>RMSE</td>
<td>$R^2$</td>
</tr>
<tr>
<td>Beijing</td>
<td>0.78</td>
<td>0.12</td>
<td>0.39</td>
<td>0.20</td>
<td>0.58</td>
</tr>
<tr>
<td>Nanjing</td>
<td>0.73</td>
<td>0.11</td>
<td>0.51</td>
<td>0.15</td>
<td>0.39</td>
</tr>
<tr>
<td>Guangzhou</td>
<td>0.72</td>
<td>0.14</td>
<td>0.52</td>
<td>0.18</td>
<td>0.51</td>
</tr>
<tr>
<td>Haikou</td>
<td>0.75</td>
<td>0.13</td>
<td>0.52</td>
<td>0.18</td>
<td>0.27</td>
</tr>
</tbody>
</table>

Figure 10. The quadratic polynomial regression models established based on the Modified Normalized Urban Areas Composite Index (MNUACI) and impervious surface area (ISA) in (a) Beijing, (b) Nanjing, (c) Guangzhou, and (d) Haikou.

5. Discussion

In this study, the MNUACI was proposed to improve the capability of delineating spatial structures of inner-urban areas using vegetation coverage and water body index via Luojia 1-01 NTL data. Four China cities with different development levels were chosen to evaluate the performance of MNUACI. To some extent, MNUACI expressed the specific spatial distribution patterns and inner-urban differentiation of urban areas. It also tackled the problems of urban area extraction in areas with low- and non-luminosity.
5.1. Comparison with Previous Indexes

It is difficult for a single NTL to correctly identify urban areas. Through the introduction of a vegetation index, EANTLI, HSI and NUACI all improved the differentiation within the city core area, but HSI mistakenly identifies water bodies as urban areas, and EANTLI also brings about commission errors because of the NTL blooming effect; NUACI applied the water body and vegetation indexes further enhance the inner-city variability and differentiation. However, NUACI is still subject to the NTL blooming effect, generating commission errors in the urban core area. For bare land and suburban towns, NTL, EANTLI and NUACI all generate numerous omission errors due to the lack of luminosity information. In contrast, by using smaller NDVI values (bare land, sand land and built-ups), HSI increases the recognition rate of bare land and urban areas, so that HSI and MNUACI can identify urban areas even under lower luminous conditions. As a regulated version of NDWI, MNDWI can effectively reduce the misclassified built-ups and their shadow information in urban water bodies, while HSI can strengthen the light index value in urban and suburban areas. By introducing improved water index and HSI, MNUACI decreased the size of saturated urban areas and increased the spatial differentiation and variability of inner-urban ones. In addition, with the introduction of HSI, the MNUACI significantly improves the identification ability of urban areas without NTL; it especially reduces the commission errors of urban areas in suburban areas. Adjusted by MNDWI and HSI, MNUACI can not only accurately express the spatial differentiation related to urban spatial structure, but can also increase spatial variation in NTL outside saturated urban areas more than NUACI.

The accuracy of urban area extraction was compared through four classification methods on four sample datasets under different natural and socioeconomic conditions. Overall accuracy, Kappa coefficient and Jaccard similarity index were introduced to assess the accuracy of MNUACI, NUACI, HSI, EANTLI and NTL in terms of urban area mapping. Although vegetation index was added to the NTL data to reduce the blooming effect of NTL, the problem of omission errors caused by unlit or low-lit areas was not solved for the high-resolution light data such as for Luojia 1-01. Therefore, NTL and EANTLI had a lower extraction accuracy for urban areas. HSI improves the identification accuracy of urban areas in unlit or low-lit areas because HSI uses a smaller NDVI. However, HSI wrongly identifies the water body as the urban area, resulting in large commission errors. The integration of the water and vegetation index allows NUACI to reduce NTL saturation and the blooming effect, but there still exists the problem of omission errors associated with no-lit or low-lit data. Among the five indexes, the urban area results extracted by MNUACI exhibit the highest accuracy and robustness. This is due to MNUACI dramatically reducing the omission error caused by the unlit area based on small NDVI values and eliminating the blooming effect of NTL through vegetation and the water body index.

For MNUACI, the SVM method has the best performance, followed by the GA method. In third is the FCM method and the DL method shows the lowest urban classification accuracy. Due to the lack of extensive shape and texture information from urban objects, the DL method failed to achieve the desired high accuracy of urban area classification. Although the GA method can obtain higher classification accuracy of urban areas by simulating the natural evolution process to determine the optimal segmentation threshold, the SVM supervised classification approach using training samples shows higher accuracy in terms of Kappa coefficient, overall accuracy and Jaccard similarity index. In addition, considering the coefficient of regression model and RMSE, the correlation degree between the NTL index and ISA is as follows: MNUACI>NUACI>HSI>EANTLI>NTL. The above results suggest that the MNUACI model is robust and reliable for extraction of urban areas. Therefore, with the global coverage of Luojia 1-01 NTL and NDVI data, the approach we proposed can also be applied to the study of urban socio-economic, and environmental issues in other countries and regions around the world.
5.2. Limitations of the Method

The proposed MNUACI model is proven to be effective and accurate in analyzing and identifying urban areas. However, there remain several shortcomings of MNUACI that could be further improved in the future study. Firstly, though the estimation errors of MNUACI were the smallest in five indexes, there still exist large omission errors in unlit and low-lit area, especially in peri-urban areas. More efforts in improving this model quality and integrating more ancillary data should be made. For instance, POI (Point of Interest), land surface temperature and population data can make up the defect caused by unlit and low-lit areas [48–50]. Also, co-registration errors of Luojia 1-01 NTL images with Landsat 8 images would be transmitted to MNUACI through NDVI and MNDWI, resulting in some urban area misidentification. Once the positioning accuracy of Luojia 1-01 images is improved, the fusion of NTL and Landsat 8 images will improve the performance of MNUACI. Moreover, it is complicated and difficult to derive accurate parameters \( d \) and \( r \) by statistical sample data of urban areas. Inaccurate parameters may lead to the inability to exclude the impact of vegetation and water. Finally, the performance of MNUACI was tested only using Luojia 1-01 NTL data, and its applicability and feasibility need to be further evaluated by using other low-resolution NTL data, such as NPP-VIIRS and DMSP-OLS.

6. Conclusions

Accurate and timely information on the spatial extent and spatial distribution of urban areas, particularly at the regional and global levels, is crucial and important for environmental and ecological issues. NTL data are valuable for regional and global urban mapping and for analysis of urban human activities. The Luojia 1-01 satellite usually captures NTL images before dawn, when urban area lights may be turned off. Thereby, NTL data might ignore certain important urban characteristics. In this research, a new urban index is proposed, combining information from Luojia 1-01 NTL data, NDVI and MNDWI of Landsat 8 data for more detailed characterizations of inner-urban variations in nighttime luminosity. In comparison with NUACI, HSI, VANTLI and NTL, MNUACI was superior in identification of inner-city forms. Then, the performance of SVM, GA, FCM and DL methods for extraction of urban areas were evaluated in four Chinese cities according to the five urban indexes mentioned. MNUACI based on the SVM method exhibits the best performance in urban area extraction, attributed to the integration of HSI, NDVI and MNDVI information. The regression models based on the five NTL indices were respectively established to map ISA using the urban fraction obtained by Landsat 8 images as the reference data. The validation results reveal a closer goodness-of-fit relationship with both MNUACI and corresponding ISA. The average correlation coefficient and the RMSE of the four cities are 0.74 and 0.13, respectively. In conclusion, combining with multi-source remotely sensed data, MNUACI has the ability to mitigate NTL pixel saturation and eliminate blooming effects, and provides a promising approach for identification of urban areas by enhancing inner-urban spatial differentiation and spatial differentiation.

Author Contributions: F.L. designed the MNUACI method for extracting urban areas and drafted the manuscript. X.L. carried out the processing of nighttime light data and land use classification using a maximum likelihood classifier based on Landsat 8 images. S.L. analyzed and assessed results of the MNUACI method. P.J. supervised the study and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Hebei Province of China (Grant No. D2018512002), the Fundamental Research Funds for the Central Universities of China (Grant No. ZY20180101) and the Key Project of Science and Technology Research for Universities of Hebei Province (Grant No. ZD2020407).

Data Availability Statement: Direct request for these materials may be made to the provider as indicated in the Acknowledgments. The processed data presented in this study are available on request from the first author.
Acknowledgments: We thank the High Resolution Earth Observation System of Hubei Data Application Center for providing Luojia 1-01 satellite imagery, and the International Institute of Spatial Life course Epidemiology (ISLE) for the research support.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References


38. Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int J. Remote Sens.* 2006, 27, 3025–3033. [CrossRef]


42. Dalponte, M.; Frizzera, L.; Gianelle, D. Individual tree crown delineation and tree species classification with hyperspectral and LiDAR data. *Peer J.* 2019, 6, 6227. [CrossRef]

43. Pok, S.; Matsushita, B.; Fukushima, T. An easily implemented method to estimate impervious surface area on a large scale from MODIS time-series and improved DMSP-OLS nighttime light data. *ISPRS J. Photogramm.* 2017, 133, 104–115. [CrossRef]


48. Li, F.; Yan, Q.; Bian, Z.; Liu, B.; Wu, Z. A POI and LST Adjusted NTL urbanindex for urban built-up area extraction. *Sensors* 2020, 20, 2918. [CrossRef]
