Mapping Impervious Surface Distribution with Integration of SNPP VIIRS-DNB and MODIS NDVI Data

Wei Guo 1, Dengsheng Lu 2,*, Yanlan Wu 1 and Jixian Zhang 3

1 School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China; E-Mails: guowei_rs@163.com (W.G.); wylmq@sina.com (Y.W.)
2 Zhejiang Provincial Key Laboratory of Carbon Cycling in Forest Ecosystems and Carbon Sequestration, School of Environmental & Resource Sciences, Zhejiang A & F University, Lin’an 311300, China
3 Institute of Photogrammetry and Remote Sensing, Chinese Academy of Surveying and Mapping, Beijing 100830, China; E-Mail: zhangjx@casm.ac.cn

* Author to whom correspondence should be addressed; E-Mail: luds@zafu.edu.cn; Tel./Fax: +86-571-6374-6366.

Academic Editors: Janet Nichol and Prasad S. Thenkabail

Received: 20 May 2015 / Accepted: 15 September 2015 / Published: 22 September 2015

Abstract: Data from the U.S. Defense Meteorological Satellite Program’s Operational Line-scan System are often used to map impervious surface area (ISA) distribution at regional and global scales, but its coarse spatial resolution and data saturation produce high inaccuracy in ISA estimation. Suomi National Polar-orbiting Partnership (SNPP) Visible Infrared Imaging Radiometer Suite’s Day/Night Band (VIIRS-DNB) with its high spatial resolution and dynamic data range may provide new insights but has not been fully examined in mapping ISA distribution. In this paper, a new variable—Large-scale Impervious Surface Index (LISI)—is proposed to integrate VIIRS-DNB and Moderate Resolution Imaging Spectroradiometer (MODIS) normalized difference vegetation index (NDVI) data for mapping ISA distribution. A regression model was established, in which LISI was used as an independent variable and the reference ISA from Landsat images was a dependent variable. The results indicated a better estimation performance using LISI than using a single VIIRS-DNB or MODIS NDVI variable. The LISI-based approach provides accurate spatial patterns from high values in core urban areas to low values in rural areas, with an overall root mean squared error of 0.11. The LISI-based approach is recommended for fractional ISA estimation in a large area.
1. Introduction

Impervious surface area (ISA) is defined as any man-made places such as buildings, highways, streets, and parking lots where water cannot penetrate [1–4]. Population increases and economic conditions have led to rapid ISA expansion in the past three decades, especially in China [5–8]. The rapid ISA increment has produced serious ecological and environmental problems [9,10] such as vegetation loss [11], water pollution [2], climate change [12–14], and urban heat islands [9,15] at local, regional, and even global scales. Therefore, increasing attention has focused on mapping ISA distribution in the past two decades [3,16,17].

Many studies on ISA mapping focused on individual cities using high or medium spatial resolution images such as QuickBird, IKONOS, Landsat, and ASTER [3]. Considering the labor intensity and cost in regional or global ISA mapping, coarse spatial resolution images such as the U.S. Defense Meteorological Satellite Program’s Operational Line-scan System (DMSP-OLS) and Moderate Resolution Imaging Spectroradiometer (MODIS) have been used in recent years [18–20]. Two types of variables—vegetation abundance and nighttime light data—are often used. Because of the inverse correlation between ISA and vegetation indices, MODIS normalized difference vegetation index (NDVI) has been used to map ISA in a large area [21–24]. However, vegetation distribution in an urban landscape is influenced by many factors such as terrain, climate, population, economic conditions, and cultures; thus, using these data alone for large-scale ISA mapping may generate high inaccuracy [25]. Since nighttime light data such as DMSP-OLS and the U.S. Suomi National Polar-orbiting Partnership (SNPP) Visible Infrared Imaging Radiometer Suite’s Day/Night Band (VIIRS-DNB) can effectively record human activities in a large area, they have been used for estimating population, energy consumption, economic data, and mapping ISA distribution [26–31]. Like vegetation data, the nighttime light data are also influenced by different factors such as economic conditions, mixed pixel problems, and data saturation (especially for DMSP-OLS) [32,33]. In previous research, DMSP-OLS was often used for mapping ISA at regional and global scales because these data have been available since 1992 at no cost [34–37]. However, the original DMSP-OLS data have a spatial resolution of 2.7 km and data range of 0–63. The mixed-pixel problem and data saturation in DMSP-OLS make it inaccurate in mapping ISA distribution in a large area. The SNNP VIIRS DNB, which was launched in October 2011, has many advantages over the DMSP-OLS data: improved spatial resolution (375 m and 750 m at nadir for VIIRS DNB, compared with 2.7 km for DMSP-OLS) and enlarged data range (14 bit for VIIRS DNB compared with 6 bit for DMSP-OLS), co-location with multispectral measurements on VIIRS and other NPOESS sensors, and elimination of cross-track pixel size variation [33,38,39]. The new generation of nighttime light data from VIIRS-DNB was first released by NOAA/NGDC in early 2013. Since then, this dataset has been employed to estimate electric power consumption and economic conditions [29,40–42]. It is believed that these data will be
valuable for mapping ISA distribution in a large area and may provide more accurate ISA estimation than DMSP-OLS.

Much previous research explored the thresholding approach of the DMSP-OLS data for mapping ISA distribution in a large area [37,43–45]. Since Lu et al. used the combination of DMSP-OLS and MODIS NDVI data to estimate ISA in South and East China [25], scientists have explored the combined use of DMSP-OLS and MODIS NDVI or SPOT VGT and confirmed the promise of improving ISA mapping performance [6,46–50]. However, the coarse spatial resolution in MODIS and DMSP-OLS cannot be effectively used to produce details of spatial patterns. The improved spatial resolution in VIIRS-DNB and MODIS NDVI will provide new insights for ISA mapping in a large area. Therefore, the objective of this research was to propose a new variable through integrating VIIRS-DNB and MODIS NDVI data to improve ISA mapping performance in a large area.

2. Study Area and Datasets

2.1. Study Area

Since 2011, over half of China’s population has lived in cities (National Bureau of Statistics of China, 2012). China has been experiencing rapid urbanization during the past three decades [5,51,52]. It is necessary to update ISA distribution frequently because of its rapid dynamic change and the requirements of urban planning and management. We chose all of China as a study area because of its large area with wide variations in population densities and economic conditions, as well as different landscape patterns and cultural customs. We chose six typical cities—Beijing, Chengdu, Kunming, Shanghai, Wuhan, and Urumqi—as sample sites to verify variables’ versatility and accuracy (see Figure 1). Beijing and Shanghai are the two biggest megacities in China with the largest populations and gross domestic products (GDPs); Wuhan and Chengdu are megacities located in Central and West China, respectively; Kunming and Urumqi are located in Southwest and Northwest China, respectively. Table 1 provides a summary of populations, GDPs, and areas of the six cities.

![Figure 1. The study area—all of China and six selected cities.](image)
Table 1. Summary of populations, gross domestic products (GDPs), and total areas for the six cities.

<table>
<thead>
<tr>
<th>Cities</th>
<th>Population (Million)</th>
<th>GDP (Billion RMB)</th>
<th>Area (km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beijing in North China</td>
<td>20.69</td>
<td>1780.1</td>
<td>16,800.00</td>
</tr>
<tr>
<td>Shanghai in East China</td>
<td>23.80</td>
<td>2010.1</td>
<td>6340.50</td>
</tr>
<tr>
<td>Wuhan in Central China</td>
<td>10.12</td>
<td>800.3</td>
<td>8494.41</td>
</tr>
<tr>
<td>Chengdu in Central China</td>
<td>11.73</td>
<td>813.8</td>
<td>12,390.00</td>
</tr>
<tr>
<td>Kunming in Southwest China</td>
<td>6.53</td>
<td>301.1</td>
<td>21,001.28</td>
</tr>
<tr>
<td>Urumqi in Northwest China</td>
<td>3.35</td>
<td>206.0</td>
<td>15,173.13</td>
</tr>
</tbody>
</table>

2.2. Datasets

Three types of remote sensing data—VIIRS-DNB (two-month composite product), MODIS 16-day NDVI composite (MOD13Q1), and Landsat 8 OLI (Operational Land Imager) data—were used in this research (see Table 2). The VIIRS-DNB and MODIS NDVI data were acquired in 2012, and Landsat 8 OLI images were acquired in 2013 because no Landsat data were available in 2012.

Table 2. Remote sensing data used in research.

<table>
<thead>
<tr>
<th>Data</th>
<th>Acquisition Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIIRS-DNB</td>
<td>Two-month composite product in April and October 2012</td>
<td>A spectral range of 500–900 nm; highly sensitive to very low levels of visible light at night with zero moonlight; spatial resolution of 743 m.</td>
</tr>
<tr>
<td>MODIS NDVI (MOD13Q1)</td>
<td>16-day MODIS NDVI composite between April and October 2012</td>
<td>Gridded level-3 product with 250 m spatial resolution.</td>
</tr>
<tr>
<td></td>
<td>h23v04-h23v05, h24v04-h24v05, h25v03-h25v06, h26v03-h26v06, h27v04-h27v06, h28v05-h28v07, h29v06; total number of scenes: 247</td>
<td></td>
</tr>
<tr>
<td>Landsat 8 OLI imagery</td>
<td>path/row: acquisition date</td>
<td>Six multispectral bands with 30 m and one panchromatic band with 15 m spatial resolution were used. Two thermal bands with 100 m spatial resolution were not used due to their relatively coarse spatial resolution.</td>
</tr>
<tr>
<td></td>
<td>123/32: 1 September 2013</td>
<td></td>
</tr>
<tr>
<td></td>
<td>118/38: 29 August 2013</td>
<td></td>
</tr>
<tr>
<td></td>
<td>123/39: 12 May 2013</td>
<td></td>
</tr>
<tr>
<td></td>
<td>129/39: 20 April 2013</td>
<td></td>
</tr>
<tr>
<td></td>
<td>129/43: 20 April 2013</td>
<td></td>
</tr>
<tr>
<td></td>
<td>143/29: 28 August 2013</td>
<td></td>
</tr>
</tbody>
</table>

Note: VIIRS-DNB data were downloaded from National Geophysical Data Center [53]; MODIS NDVI time-series data were downloaded from the NASA Goddard Space Flight Center (GSFC) [54]; and Landsat 8 OLI imagery was downloaded from the United States Geological Survey (USGS) (http://earthexplorer.usgs.gov/).

3. Methods

Figure 2 illustrates the framework for mapping ISA in a large area using a combined use of VIIRS-DNB and MODIS NDVI data. As a comparison, single VIIRS-DNB and MODIS NDVI variables
were also used with the same samples. The major steps include: (1) produce ISA reference data from Landsat 8 OLI data for the selected cities; (2) develop a new index based on the combination of VIIRS-DNB and MODIS NDVI data; (3) establish regression models for ISA estimation; and (4) evaluate ISA estimates.

Figure 2. Framework of mapping ISA distribution using the Large-scale Impervious Surface Index (LISI).

3.1. Produce ISA Reference Data from Landsat 8 OLI Imagery

Landsat 8 OLI imagery covers 11 bands, including eight reflective bands (e.g., visible, near infrared, and shortwave infrared) with 30 m spatial resolution, one panchromatic band with 15 m spatial resolution, and two thermal infrared bands with 100 m spatial resolution [55]. Band 1 (violet-deep blue), band 9 (Cirrus), and bands 10 and 11 (thermal infrared) were not used in this research. The Universal Transverse Mercator (UTM) coordinate system in the Landsat 8 OLI data was re-projected to Albers Conical Equal Area projection. A combination of thresholding and cluster analysis was used to extract ISA data, as illustrated in Figure 3. The major steps include (1) conduct data fusion to produce a new dataset with improved spatial resolution; (2) produce vegetation indices to mask out vegetation and water; (3) extract spectral signatures for the remaining pixels and conduct cluster analysis; and (4) merge the clusters into ISA and others and evaluate ISA results.

The spatial resolution of remote sensing data is an important factor in effectively extracting ISA data because of the complex land-cover composition in urban landscapes [3]. Previous research has indicated that improved spatial resolution in Landsat data is valuable in ISA mapping and the wavelet-merging technique is an effective tool to integrate multispectral and panchromatic data into a new dataset [56] and thus this technique is used in this research.
Many vegetation indices have been developed [21,57,58]. NDVI may be the most commonly used index for vegetation-related studies and NDWI (normalized difference water index) for water extraction [59,60]. Thus, they are used in this research to extract vegetation and water from the fused multispectral imagery. Based on analysis of vegetation samples, a threshold of 0.3 in NDVI data was selected to extract the vegetation pixels. Meanwhile, a threshold of 0.1 in NDWI was used to extract water pixels. After masking the vegetation and water pixels in the fused multispectral imagery, the remaining pixels are mainly ISA and bare soils. The spectral signatures from the fused imagery were extracted for the remaining pixels, and a cluster analysis (ISODATA in this research) was conducted to classify the remaining pixels into 100 clusters. The analyst was then to merge the clusters into ISA and others based on analysis of their spectral signature and QuickBird images in Google Earth. Finally, ISA results with 15 m spatial resolution were produced. Although no quantitative accuracy assessment for the ISA results was conducted, visual examination of the final ISA images by overlaying them on the corresponding Landsat 8 OLI color composite indicates the reliability and good quality; thus, the extracted ISA data from Landsat were used in modeling and accuracy assessment for the ISA estimation in a large area.

3.2. Develop Large-Scale Impervious Surface Index Data through a Combination of VIIRS-DNB and MODIS NDVI Data

The VIIRS-DNB data were re-projected from geographic (Lat/Lon) system into Albers Conical Equal Area projection and resampled to a cell size of 750 m by 750 m using the nearest-neighbor algorithm. Since the VIIRS-DNB data contain fires, gas flares, volcanoes, and background noises, they
must be removed before using the data for ISA mapping. In this research, we used a threshold value of 0.5 to remove them. As the majority of data values were less than 65 (some had extremely high values) we set the threshold of 65 as the maximum, which means that all the pixel values greater than 65 were assigned to 65. In order to keep all data sources in the same range, between 0 and 1, the VIIRS-DNB image was normalized with Equation (1):

$$DNB_{nor} = \frac{DBN - DNB_{min}}{DNB_{max} - DNB_{min}}$$  \hspace{1cm} (1)$$

where $DNB_{nor}$ is a fractional VIIRS-DNB image having data ranges between 0 and 1, $DNB_{min}$ and $DNB_{max}$ are the minimum and maximum values, respectively.

The MODIS NDVI data (MOD13Q1 product here: 16-day composite with 250 m spatial resolution) were re-projected from sinusoidal projection to Albers Conical Equal Area projection, and the nearest-neighbor resampling algorithm was used during the reprojection procedure. In theory, NDVI imagery has values ranging from −1 to +1. Due to the coarse spatial resolution in MODIS NDVI, the land surface covers, not including water bodies, in a large area have data ranges between 0 and 1 during the growing season.

Since non-vegetation land covers such as farmlands (e.g., bare soils after harvest), water, and ISA have similar NDVI values, it is necessary to reduce the confusion among them. One effective approach is to produce a maximum NDVI image from multi-temporal NDVI images [25]:

$$NDVI_{max} = \text{MAX}[NDVI_1, NDVI_2, ..., NDVI_n]$$  \hspace{1cm} (2)$$

where $NDVI_1$, $NDVI_2$, ..., $NDVI_n$ are the multitemporal MOD13Q1 NDVI images acquired in 2012. Another important role for Equation (2) is to remove the impact of cloud contamination. Therefore, the final $NDVI_{max}$ imagery is cloud-free and has a data range between 0 and 1.

In urban landscapes, vegetation abundance is closely related to the patterns of settlements [61]; thus, vegetation indices have been used to estimate ISA. However, non-vegetation land covers, such as bare soils and water bodies within or outside the urban areas have NDVI values similar to ISA. Therefore, individual NDVI data are difficult to use directly for separating ISA from water and bare soils. Oppositely, the nighttime light data represent the urban area and have features considerably different from data for other regions, but these data are also influenced by different economic conditions [44,62]. Therefore, a combination of nighttime light and NDVI data has complementary features, as previous research has confirmed [25,46].

Two common combination variables—human settlement index [25] and vegetation adjusted normalized urban index [6,46]—have been proposed for mapping ISA distribution in a large area. Both variables are based on DMSP-OLS and MODIS NDVI with the same spatial resolution data (1000 m). Here we present a new combination variable called Large-scale Impervious Surface Index (LISI) based on VIIRS-DNB data with spatial resolution of 750 m and MODIS NDVI$\text{max}$ with spatial resolution of 250 m:

$$\text{LISI} = (1 - NDVI_{max}) \times \sqrt{\text{DNB}_{nor}}$$  \hspace{1cm} (3)$$

Since $NDVI_{max}$ is negatively related to ISA, $1-NDVI_{max}$ is used to keep the values between 0 and 1 and is positively related to ISA. Compared to DMSP-OLS data, VIIRS-DNB data have high spatial resolution and a richly dynamic data range. Since some objects, such as airports and new, tall buildings [41] have very high DNB values, the expression $\sqrt{\text{DNB}_{nor}}$ can smooth this effect. The
combination of $\sqrt{\text{DNB}_{\text{nor}}}$ and $(1 - \text{NDVI}_{\text{max}})$ not only highlights urban ISA but also improves spatial resolution. As a comparison of these datasets, Figure 4 illustrates the improvement by combining both features into the new dataset.

**Figure 4.** A comparison of four datasets in Beijing City; (a) VIIRS-DNB with 750 m spatial resolution; (b) 1-NDVI$_{\text{max}}$ with 250 m spatial resolution; (c) LISI with improved spatial resolution (250 m); and (d) Landsat 8 OLI color composite with 15 m spatial resolution.

3.3. *Map ISA Distribution with Regression Models*

Regression analysis was used to develop ISA estimation models. The dependent variable—fractional ISA data—was obtained from the Landsat 8 OLI images, which were aggregated from a cell size of 15 m to 250 m and 750 m, respectively, using a mean algorithm to match the cell size of LISI and 1-NDVI$_{\text{max}}$ (250 m) and DNB$_{\text{nor}}$ (750 m). The independent variable is DNB$_{\text{nor}}$, 1-NDVI$_{\text{max}}$, and LISI, respectively. In order to develop the regression models, a random sampling technique was used to collect samples for each selected city. A total of 4800 samples were extracted from six reference images with 800 samples for each. The values for the same sample locations were extracted from DNB$_{\text{nor}}$, 1-NDVI$_{\text{max}}$, and LISI. Of these samples, 3600 samples were used for modeling and the remaining 1200 samples were used to evaluate ISA estimates. The coefficient of determination ($R^2$) was used to evaluate the fitness of regression models. Meanwhile, three typical cities—Beijing,
Wuhan, and Urumqi—were selected for a comparative analysis of their ISA spatial patterns, which were developed from the three regression models based on DNB$_{nor}$, 1-NDVI$_{max}$, and LISI.

3.4. Conduct Evaluation of ISA Estimates

Overall accuracy, kappa coefficient, and producer’s and user’s accuracies are often used for pixel-level classification evaluation [63]. However, these approaches are not suitable for the evaluation of fractional ISA estimates [3]. As in previous research [64], we used correlation coefficient (R) and root mean squared error (RMSE) to evaluate the results. In addition to the evaluation of ISA estimates in overall China, the accuracy assessment was also conducted at five ISA groups—very low, low, medium, high, and very high—based on 0.2 intervals of the ISA reference data (between 0 and 1), and at individual cities. The objectives of different accuracy assessment methods are to understand the error sources, whether they are from different levels of ISA values or from different spatial locations due to various population densities and economic conditions.

4. Results

4.1. Analysis of ISA Spatial Distribution

A comparison of R$^2$ values among three regression models (see Figure 5) indicate that the LISI-based model has the highest performance and the 1-NDVI$_{max}$-based model has the lowest performance. This is reasonable because NDVI is influenced by environmental and geographic factors such as bare soils, moisture, and compositions of different land-cover types. This study indicates that individual MODIS NDVI data are not suitable for ISA mapping in China because of the considerable difference of vegetation conditions in Eastern and Western China. Although using VIIRS DNB provides better estimation performance than NDVI, its estimation variation is wide due to the impacts of different economic conditions on nighttime light data. Figure 6 illustrates the ISA distribution using the LISI-based model, indicating that large amounts of ISA are distributed along the coastal regions and in the central metropolis of major cities, and much smaller amounts of ISA in Western and Northwestern China. This figure also clearly shows the spatial pattern differences between a high ISA proportion in core urban areas and a low ISA proportion in rural regions (see Figure 6a–c).

![Figure 5](image-url)
Figure 6. ISA distribution in China using the LISI-based model, highlighting the ISA spatial patterns from high ISA proportions in core urban areas to low proportions in rural regions; (a) Urumqi; (b) Beijing; and (c) Wuhan.

In order to better explain the ISA estimation performances, three cities—Beijing, Wuhan, and Urumqi—having different economic conditions and locations (see Table 1) were selected as examples. Beijing and Wuhan are located in North and Central China, respectively, with much better vegetation distributions than Urumqi in the west of China due to the latter’s dry weather. A comparison of ISA distribution (see Figure 7) indicates that the LISI-based model highlights the ISA spatial patterns from the highest values in core urban areas to the lowest values in rural areas, and the NDVI-based approach cannot effectively extract ISA distribution in urban regions, resulting in considerable underestimation when ISA density is high in an urban area. Figure 7 shows that the LISI-based method provides better ISA estimation performance than DNB- and NDVI-based approaches for each city.

4.2. Comparative Analysis of ISA Estimates

The overall accuracy assessment results in Table 3 also confirm that the LISI-based approach provides the best accuracy with R value of 0.81 and RMSE value of 0.11, much improved compared to the 1-NDVI\textsubscript{max}-based approach. The scatterplots between estimates and reference data (Figure 8) indicate a reasonably good estimation performance using the DNB\textsubscript{hor}-based approach and best estimation performance using the LISI-based approach. In contrast, the 1-NDVI\textsubscript{max}-based approach has very poor estimation performance because it cannot effectively estimate ISA values when the ISA proportion in a pixel is higher than 0.5.
Figure 7. A comparison of the ISA distributions from three datasets for three cities, highlighting the better spatial patterns from the LISI-based model than the other two results; (a1–a3) represent ISA distribution using the DNB\textsubscript{nor} in Beijing, Wuhan, and Urumqi; (b1–b3) represent ISA distribution using 1-NDVI\textsubscript{max} in Beijing, Wuhan, and Urumqi; and (c1–c3) represent ISA distribution using LISI in Beijing, Wuhan, and Urumqi.

Table 3. A comparison of overall accuracy assessments among three datasets.

<table>
<thead>
<tr>
<th>Variable</th>
<th>R</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNB\textsubscript{nor}</td>
<td>0.729</td>
<td>0.132</td>
</tr>
<tr>
<td>1-NDVI\textsubscript{max}</td>
<td>0.563</td>
<td>0.160</td>
</tr>
<tr>
<td>LISI</td>
<td>0.812</td>
<td>0.113</td>
</tr>
</tbody>
</table>
Figure 8. The relationships between ISA estimates and corresponding reference data among three datasets; (a) ISA estimates from DNB\textsubscript{nor}; (b) ISA estimates from 1-NDVI\textsubscript{max}; and (c) ISA estimates from LISI.

The analysis above is based on overall estimation performance, but may not reveal some details caused by different environmental or economic conditions. Table 4 summarizes RMSE results, which are grouped into five ISA levels—very low, low, medium, high, and very high—based on reference data ranges: <0.2, [0.2–0.4), [0.4–0.6), [0.6–0.8), and ≥0.8. These results confirmed that the LISI-based approach produces better performance than the other two. Comparing the RMSE values at different ISA levels indicates that the LISI-based approach improves ISA estimation performance when ISA is relatively low or high, but when ISA is at medium level (i.e., between 0.4 and 0.6), the DNB\textsubscript{nor}-based approach provides better estimation performance than the LISI-based approach. Table 4 indicates that a single 1-NDVI\textsubscript{max} variable is not suitable for ISA estimation in China, DNB\textsubscript{nor} is valuable when ISA is at medium level, and LISI is recommended for ISA estimation in a large area.

Table 4. Comparison of RMSE results among three estimation approaches at five ISA levels.

<table>
<thead>
<tr>
<th>Group</th>
<th>Data Range</th>
<th>RMSE DNB\textsubscript{nor}</th>
<th>RMSE 1-NDVI\textsubscript{max}</th>
<th>RMSE LISI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very low</td>
<td>&lt;0.2</td>
<td>0.100</td>
<td>0.121</td>
<td>0.076</td>
</tr>
<tr>
<td>Low</td>
<td>0.2–0.4</td>
<td>0.217</td>
<td>0.188</td>
<td>0.181</td>
</tr>
<tr>
<td>Medium</td>
<td>0.4–0.6</td>
<td>0.182</td>
<td>0.461</td>
<td>0.195</td>
</tr>
<tr>
<td>High</td>
<td>0.6–0.8</td>
<td>0.226</td>
<td>Null</td>
<td>0.179</td>
</tr>
<tr>
<td>Very high</td>
<td>≥0.8</td>
<td>0.310</td>
<td>Null</td>
<td>0.215</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>0.132</td>
<td>0.160</td>
<td>0.113</td>
</tr>
</tbody>
</table>

Note: Null represents there are no data in ISA estimates.

R and RMSE were also used to evaluate the estimation performance at individual cities, and the results are summarized in Table 5. Again, the LISI-based approach provided better estimation performance than the other two variables in all cities, implying the robustness of the LISI variable in ISA estimation. The 1-NDVI\textsubscript{max} variable has good performance in Beijing and Chengdu, but very poor performance in Urumqi, implying that the 1-NDVI\textsubscript{max} variable is not suitable for ISA estimation in cities where vegetation accounts for a limited proportion and bare soils/desert have important impacts.
on the NDVI values. The DNB$_{nor}$-based approach provides reliable estimation in different cities, but relatively poorer performance than the LISI-based approach.

**Table 5.** Comparison of accuracy analysis results from three variables at different cities.

<table>
<thead>
<tr>
<th>City</th>
<th>DNB$_{nor}$</th>
<th>1-NDVI$_{max}$</th>
<th>LISI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>RMSE</td>
<td>R</td>
</tr>
<tr>
<td>Beijing</td>
<td>0.744</td>
<td>0.170</td>
<td>0.845</td>
</tr>
<tr>
<td>Shanghai</td>
<td>0.804</td>
<td>0.105</td>
<td>0.661</td>
</tr>
<tr>
<td>Wuhan</td>
<td>0.793</td>
<td>0.096</td>
<td>0.740</td>
</tr>
<tr>
<td>Chengdu</td>
<td>0.801</td>
<td>0.104</td>
<td>0.801</td>
</tr>
<tr>
<td>Kunming</td>
<td>0.739</td>
<td>0.109</td>
<td>0.764</td>
</tr>
<tr>
<td>Urumqi</td>
<td>0.814</td>
<td>0.076</td>
<td>0.233</td>
</tr>
</tbody>
</table>

**5. Discussions**

DMSP-OLS data are often used for ISA estimation in a large area using a thresholding-based approach [44,45,65], but a pixel-based approach produces high inaccuracy in ISA estimation due to the mixed-pixel problem (2.7 km spatial resolution for the original data) and data saturation (6 bits) [33]. Since the DMSP-OLS data are seriously influenced by different economic conditions, no suitable thresholds can be used to accurately extract ISA data, thus, the spatial patterns of extracted ISA distribution are often poor [25]. The VIIRS-DNB data considerably reduce the problems in DMSP-OLS because of the improvement in spatial resolution and data ranges; thus, the ISA estimation using VIIRS-DNB for ISA estimation in China has reasonably good estimation performance, with an overall RMSE of 0.13, as shown in this research.

Due to the problems inherited in DMSP-OLS data, previous research has explored the combination of DMSP-OLS and MODIS NDVI data for ISA mapping. Lu et al. (2008) first proposed the human settlement index based on the combination of DMSP-OLS and MODIS NDVI data for mapping human settlements in China and obtained much better estimation performance (including the improved spatial patterns and estimation accuracy) than using individual DMSP-OLS data [25]. Zhang et al. (2013) further proposed the vegetation adjusted normalized urban index to improve the ISA estimation [46]. The combined indices are based on coarse spatial resolution images (e.g., 1 km spatial resolution) and small data ranges (0–63 in DMSP-OLS data), thus, the improvement is influenced by the data limitations themselves. This research proposed the LISI variable—an integration of VIIRS-DNB and MODIS NDVI—which has provided much improvement in ISA estimation because of the integration of the improved spatial resolution and data ranges in VIIRS-DNB data and MODIS NDVI with 250 m spatial resolution. The LISI variable combined the advantages of both VIIRS-DNB and MODIS NDVI data; that is, VIIRS-DNB can effectively reflect the difference between urban and non-urban regions, and MODIS NDVI can better capture the differences within the urban inner structure (e.g., the different compositions of ISA, vegetation, and water). Additionally, this research indicates that the LISI-based approach can provide robust ISA estimation in different cities, and it is especially valuable in improving the ISA estimation when the ISA proportion in a pixel is relatively low or high.

Previous research using DMSP-OLS for ISA mapping has indicated that unbalanced economic conditions in a large area affect the nighttime light values; therefore, multiple thresholds are often used
to extract ISA data for different cities (or locations) [44]. However, this research shows that the VIIRS-DNB reduced this problem. Even when using the same DNB-based model for ISA mapping in China, the different economic conditions such as in Shanghai, Chengdu, and Urumqi did not considerably affect the ISA estimation performance, as shown in Table 5. When combining the RMSE (Table 5) and economic conditions (Table 1) for the same city, a scatterplot showing the relationships between RMSE and GDP indicates that economic condition is indeed an important factor affecting the estimation errors (see Figure 9). For example, Beijing and Shanghai have the largest populations and GDPs in China, but their different characteristics in spatial patterns of urban landscapes and different land-cover compositions affect the ISA estimation performance. Use of individual VIIRS-DNB data has relatively high estimation errors when economic condition is very high or relatively low, as shown in Figure 9. However, the LISI can considerably reduce this problem and is, thus, recommended for ISA mapping in a large area. We encourage researchers to explore the use of LISI in different countries. Furthermore, this research only explores the ISA estimation using a linear regression model, and more research is needed in the future to explore nonparametric algorithms such as a support vector machine or neural network for developing ISA estimation models.

![Figure 9](image.png)

**Figure 9.** The relationship between RMSE and economic conditions, showing the effects of different economic conditions on ISA estimation performance through the comparison of different variables.

6. Conclusions

Through comparative analysis of VIIRS-DNB, MODIS NDVI, and the proposed LISI variable for ISA mapping in a large area, we obtained the following conclusions:
(1) VIIRS-DNB can be used for ISA mapping in a large area with an overall RMSE of 0.13. However, the areas having higher ISA proportion produced higher errors. Additionally, very high or very low economic conditions influenced ISA estimation performance. This implies that individual VIIRS-DNB data may produce inaccurate spatial patterns of ISA distribution if the study area covers urban landscapes having considerably different economic conditions;

(2) Individual MODIS NDVI is not a good variable for ISA mapping in a large area, especially in areas with very low vegetation covers, such as Western China. However, in some large cities such as Chengdu and Kunming in this research, NDVI can produce ISA estimates with similar to or even better performance than VIIRS-DNB. This implies that MODIS NDVI is valuable, but it is critical to properly use it in ISA estimation;

(3) The proposed LISI variable combined advantages of both VIIRS-DNB and MODIS NDVI features and provided much improved ISA estimation performance, especially the improved spatial patterns. Overall, the LISI-based approach has an RMSE of 0.11 and has much-improved estimation performance when ISA proportion is high, compared to the other two datasets. Therefore, LISI is recommended for ISA estimation in a large area.

Acknowledgments

The study was financially supported by the Zhejiang A & F University’s Research and Development Fund for the talent startup project (2013FR052) and facility support from the Center for Global Change and Earth Observations at Michigan State University. The authors would like to thank Changming Zhu for his assistance in data processing.

Author Contributions

Wei Guo is responsible for the data collection and processing, resultant analysis, and preparation of manuscript. Dengsheng Lu is responsible for the design of the study, technical support, resultant analysis, and preparation of the manuscript. Yanlan Wu and Jixian Zhang contributed to the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

References


© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).