

Article

Urban Land Extraction Using VIIRS Nighttime Light Data: An Evaluation of Three Popular Methods

Yinyin Dou ^{1,2,†}, Zhifeng Liu ^{1,2,†}, Chunyang He ^{1,2,*} and Huanbi Yue ^{1,2}

¹ Center for Human-Environment System Sustainability (CHESS), State Key Laboratory of Earth Surface Processes and Resource Ecology (ESPRE), Beijing Normal University, 19 Xinjiekouwai Street, Beijing 100875, China; douyy@mail.bnu.edu.cn (Y.D.); Zhifeng.Liu@bnu.edu.cn (Z.L.); yuehuanbi@163.com (H.Y.)

² Academy of Disaster Reduction and Emergency Management, Faculty of Geographical Science, Beijing Normal University, 19 Xinjiekouwai Street, Beijing 100875, China

* Correspondence: hcy@bnu.edu.cn; Tel.: +86-10-5880-4498

† These authors contributed equally to this work.

Academic Editors: Bailang Yu, Yuyu Zhou, Xiaofeng Li, James Campbell and Prasad S. Thenkabail

Received: 15 December 2016; Accepted: 15 February 2017; Published: 20 February 2017

Abstract: Timely and accurate extraction of urban land area using the Suomi National Polar-orbiting Partnership Visible Infrared Imaging Radiometer Suite (VIIRS) nighttime light data is important for urban studies. However, a comprehensive assessment of the existing methods for extracting urban land using VIIRS nighttime light data remains inadequate. Therefore, we first reviewed the relevant methods and selected three popular methods for extracting urban land area using nighttime light data. These methods included local-optimized thresholding (LOT), vegetation-adjusted nighttime light urban index (VANUI), integrated nighttime lights, normalized difference vegetation index, and land surface temperature support vector machine classification (INNL-SVM). Then, we assessed the performance of these methods for extracting urban land area based on the VIIRS nighttime light data in seven evaluation areas with various natural and socioeconomic conditions in China. We found that INNL-SVM had the best performance with an average kappa of 0.80, which was 6.67% higher than the LOT and 2.56% higher than the VANUI. The superior performance of INNL-SVM was mainly attributed to the integration of information on nighttime light, vegetation cover, and land surface temperature. This integration effectively reduced the commission and omission errors arising from the overflow effect and low light brightness of the VIIRS nighttime light data. Additionally, INNL-SVM can extract urban land area more easily. Thus, we suggest that INNL-SVM has great potential for effectively extracting urban land with VIIRS nighttime light data at large scales.

Keywords: VIIRS nighttime light data; urban land extraction; normalized difference vegetation index; land surface temperature; support vector machine; local-optimized thresholding

1. Introduction

The world has been experiencing dramatic urban land area growth, primarily due to rapid economic development and population growth [1–3]. From 2000 to 2010, the global built-up area was estimated to have increased from 0.60 to 0.87 million km², with an average annual growth of 3.73% [4]. At this growth rate, the global built-up area will increase to 3.11 million km² by 2050, which is close to the total terrestrial area of India [4]. Urban expansion causes numerous ecological and environmental effects worldwide [5–8]. For example, urban expansion in China has resulted in a natural habitat loss of 8647 km² from 1992 to 2012 [5], and urban expansion in the pan-tropics contributes to a loss in vegetation biomass of 1.38 PgC or approximately 5% of emissions from tropical deforestation and land-use change [9]. Moreover, urban expansion aggravates water shortage, air pollution, and the urban heat island effect, bringing stresses to regional sustainability around the world [1]. To address

these issues in urban studies, it is important to extract urban land area using timely and effective approaches [2,3,10].

Recently, the Suomi National Polar-orbiting Partnership Visible Infrared Imaging Radiometer Suite (VIIRS) nighttime light data have begun to be a novel data source for the timely and accurate extraction of urban land area [11,12]. In 2013, the Earth Observation Group in the National Oceanic and Atmospheric Administration's National Geophysical Data Center (NOAA/NGDC) of the United States published the global VIIRS nighttime light test data for 2012. In 2014, it released version 1, and in 2016 nighttime light data became available on a monthly basis between March 2014 and September 2016. The VIIRS nighttime light data are suitable for extracting urban land area due to the effective low nighttime light-detecting capability, with high radiometric resolution, large-scale coverage, and short revisiting period [13–16]. Specifically, VIIRS can detect nighttime lights within the radiation range from $3 \text{ nW} \cdot \text{cm}^{-2} \cdot \text{sr}^{-1}$ to $0.02 \text{ W} \cdot \text{cm}^{-2} \cdot \text{sr}^{-1}$, with a span from 75°N latitude to 65°S , a 3000 km swath, and a 12-h revisiting period [17]. Moreover, compared with Defense Meteorological Satellite Program—Operational Linescan System (DMSP-OLS) nighttime light data published earlier, VIIRS nighttime light data are of much higher quality [17,18]. The spatial resolution improved from 1 to 742 m. The digital number (DN) values also changed from relative light intensity values with a 6-bit quantization limit to calibrated radiance values with a 14-bit quantization, which effectively resolved the saturation problem [17]. Recently, some researchers have used VIIRS data to extract urban land area at regional, national, and global scales. For example, Shi et al. [11] obtained urban land information on 12 major Chinese cities in 2012. Xu et al. [8] extracted urban land area in China in 2015. Sharma et al. [19] extracted worldwide urban land areas in 2014.

However, challenges still exist for using VIIRS nighttime light data to extract urban land area. First, VIIRS nighttime light data are affected by the overflow effect as non-coherent light radiates in all directions from its source (i.e., the dispersion of light into surrounding areas) [20], resulting in the overestimation of urban land area [5,8]. Second, the VIIRS overpass time is after midnight, i.e., near 1:30 a.m., when most urban areas exhibit some decline in nighttime light brightness, which can lead to an underestimation of urban land area [17]. Thus, further studies are needed for accurate urban land area extraction using VIIRS nighttime light data [5,8,11,19].

Currently, 11 methods within three categories, i.e., thresholding, index, and supervised classification, have been used to extract urban land area using the VIIRS nighttime light data (Table 1). Among them, there are three widely used methods, including local-optimized thresholding (LOT), the vegetation-adjusted nighttime light urban index (VANUI), and the integrated nighttime lights, normalized difference vegetation index (NDVI) and land surface temperature (LST) support vector machine (SVM) classification (INNL-SVM) (Table 1). Generally, LOT determines an optimal threshold according to ancillary data (e.g., socioeconomic data, medium- to high-resolution remote sensing data, etc.) and extracts areas with nighttime light brightness greater than the optimal threshold as urban areas [21]. Using this approach, Xie and Weng [22] obtained urban land information for 2000, 2005, and 2010 in China. The VANUI is an index developed by combining nighttime light data with NDVI data and is used to extract urban land area based on an optimal threshold determined by ancillary data. Using the VANUI, Li et al. [23] extracted urban land area in the southeastern United States from 1992 to 2013. INNL-SVM is a supervised classification approach that integrates nighttime light data, NDVI data, and LST data. It utilizes the SVM classifier to extract urban land area based on training samples [5,24]. Using INNL-SVM, Xu et al. [8] extracted urban land area in China from 1992 to 2015. These three methods largely differ in theory, data source, applied region, and accuracy (Table 1). However, an application assessment of these three methods for the VIIRS nighttime light data is insufficient, limiting broad utilization of the VIIRS nighttime light data for urban land extraction.

Table 1. An overview of the methods for extracting urban land area using nighttime light data.

Categories	Methods	Study Area	Year	Kappa Index	Overall Accuracy	Nighttime Light Data	Reference
Thresholding	Empirical thresholding (ET)	United States	1994–1995	**	**	DMSP-OLS	[25,26]
	City-lights polygon perimeter-based thresholding (CPT)	United States	1994–1995	**	**	DMSP-OLS	[27]
	Brightness gradient-based thresholding (BGT))	274 Chinese cities	1992, 2002, 2012	**	**	DMSP-OLS	[28]
		Southeastern United States	1992, 1993, 2000	**	**	DMSP-OLS	[29]
		San Francisco, Beijing and Lhasa	1994–1995	***	68.0%–99.2%	DMSP-OLS	[21]
		China	1992, 1996, 1998	***	79.5%–82.9%	DMSP-OLS	[30]
	Local-optimized thresholding (LOT)	China	1992–2008	0.60	86.3%	DMSP-OLS	[31]
Index		Global	2000	0.55	87.0%	DMSP-OLS	[32]
		China	2000, 2005, 2010	0.53–0.56	95.2%–94.1%	DMSP-OLS	[22]
Supervised classification		12 Chinese cities	2012	***	89.6%–96.3%	VIIRS	[11]
	Enhanced built-up index (EUBI)	The whole world	2014	**	**	VIIRS	[19]
Index	Vegetation-adjusted nighttime light urban index (VANUI)	Southeastern United States	1992–2013	***	85.0%	DMSP-OLS	[23,33]
	Classification and regression tree (CART)	99 Chinese cities	2010	0.53	94.4%	DMSP-OLS	[34]
Supervised classification	K-nearest-neighbors (KNN)	99 Chinese cities	2010	0.58	96.2%	DMSP-OLS	[34]
	Random forests (RF)	99 Chinese cities	2010	0.60	96.4%	DMSP-OLS	[34]
Supervised classification	Integrated nighttime lights and NDVI SVM classification (INNL-SVM) *	25 Chinese cities	2000	0.62	90.5%	DMSP-OLS	[35]
		China	2008	0.69	90.0%	DMSP-OLS	[36]
		99 Chinese cities	2010	0.57	96.7%	DMSP-OLS	[36]
Supervised classification	Integrated nighttime lights, NDVI and LST SVM classification (INNL-SVM) *	China	1992–2012	0.66	95.2%	DMSP-OLS and VIIRS	[5]
		55 Chinese cities	2006	0.33–0.80	73.4%–98.4%	DMSP-OLS	[24]
		China	1992–2015	0.60	92.62%	DMSP-OLS and VIIRS	[8]

* NDVI: normalized difference vegetation index; SVM: support vector machine; LST: land surface temperature; ** The accuracy assessment was not performed; *** The Kappa index was not used in the accuracy assessment.

The objective of this study was to assess the LOT, VANUI, and INNL-SVM methods for extracting urban land area using the VIIRS nighttime light data. To achieve this goal, we first selected seven evaluation areas with various natural and socioeconomic conditions in China, which is experiencing rapid urbanization. Then, we extracted year 2015 urban land area for these evaluation areas using the LOT, VANUI, and INNL-SVM methods. Finally, we assessed the accuracy of these methods, and a discussion of their merits and disadvantages is provided.

2. Study Area and Data

2.1. Study Area

We selected seven evaluation areas with diverse natural and socioeconomic conditions based on the major biomes in China to represent regional discrepancies (Figure 1, Table 2). Each evaluation area is approximately 2.5 thousand km² (100 × 100 pixels) and includes one core city with an urban population exceeding 1 million (Table 2). Among these areas, the Kunming area in a tropical and sub-tropical moist broadleaf forests biome has the highest mean annual temperature of 14–16 °C and the largest mean annual precipitation of 800–1000 mm. In 2014, the core city had an urban population of 3.55 million and a gross domestic product (GDP) per capita of 71.75 thousand RMB yuan. The Jiamusi area, which is located in a flooded grasslands and savannas biome, has the lowest mean annual temperature of 2–4 °C. In 2014, the core city had an urban population of 1.36 million and a GDP per capita of 43.37 thousand RMB yuan. The Urumqi area, which is in a desert and xeric shrubland biome, has the lowest mean annual precipitation of 100–200 mm. In 2014, the core city had an urban population of 2.58 million and a GDP per capita of 94.84 thousand RMB yuan.

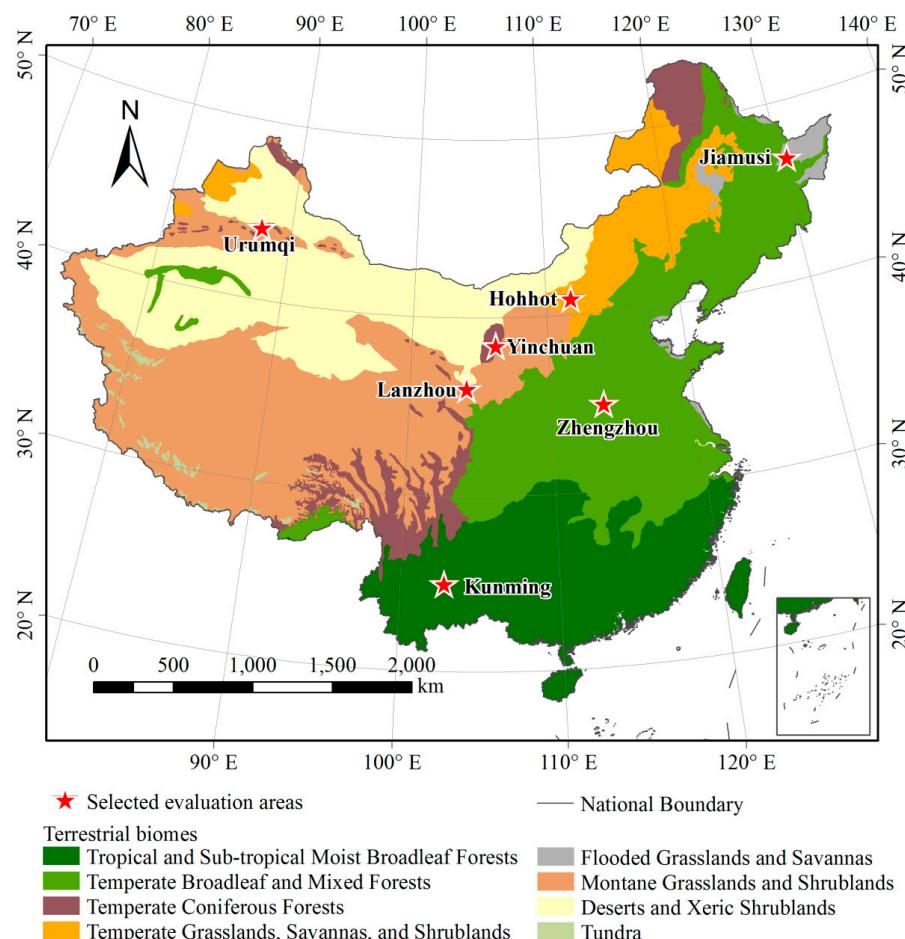


Figure 1. Study area.

Table 2. Description of the selected evaluation areas.

Evaluation Area *	Biome	Mean Annual Temperature (°C)	Mean Annual Precipitation (mm)	City	Urban Population in 2014 (Thousand Persons) **	GDP Per Capita in 2014 (Thousand·RMB·Yuan) **
Kunming	Tropical and sub-tropical moist broadleaf forests	14–16	800–1000	Kunming	3549.20	71.75
				Anning	262.80	71.34
				Yiliang	174.50	32.52
				Jining	116.10	35.40
				Congming	115.30	31.14
				Fumin	51.90	36.26
				Chengjiang	19.40	37.16
Zhengzhou	Temperate broadleaf and mixed forests	13–15	600–800	Zhengzhou	4078.30	60.11
				Xinzheng	436.20	81.81
				Xinmi	410.90	75.08
				Zhongmou	381.30	77.90
				Xingyang	306.30	92.69
				Wuzhi	222.70	41.23
				Yuanyang	193.10	16.54
Yinchuan	Temperate coniferous forests	8–10	200–300	Yinchuan	1247.68	61.95
				Helan	112.66	44.00
				Yongning	97.98	49.89
Hohhot	Temperate grassland, savanna, and shrubland	4–6	300–400	Hohhot	1790.68	134.07
Jiamusi	Flooded grassland and savanna	2–4	400–600	Jiamusi	1355.50	43.37
Lanzhou	Montane grassland and shrubland	8–10	300–400	Lanzhou	1852.90	80.73
Urumqi	Desert and xeric shrubland	4–6	100–200	Urumqi	2578.00	94.84
				Changji	372.90	74.48

* The name of the largest city was used as the name of each evaluation area; ** The urban population and GDP were obtained from the statistical yearbooks for corresponding provinces [37].

2.2. Data

The VIIRS nighttime light data for 2015 were obtained from the NOAA/NCEI (formerly NGDC) website [38]. These data include monthly average radiance composite images from January to August 2015 produced using nighttime data from the VIIRS Day/Night Band (DNB) acquired from the visible infrared radiometer onboard the Suomi NPP satellite. Prior to averaging, the DNB data are filtered to exclude data impacted by stray light, lightning, lunar illumination, and cloud-cover [39]. The spatial resolution is 15 arc-seconds (742 m), with a spectral range from 0.5 to 0.9 μm [17]. The DN unit is $\text{nW}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$, and the geographic coordinate system is WGS-1984. To make full use of the information derived from eight monthly composites for 2015, we followed the monthly composition approach to produce a mean annual value composite of the VIIRS nighttime light data and resampled it to a spatial resolution of 500 m [17]. The VIIRS nighttime light data have a high quality with temporal consistency (Figure 2a) and without saturation effect (Figure 2b).

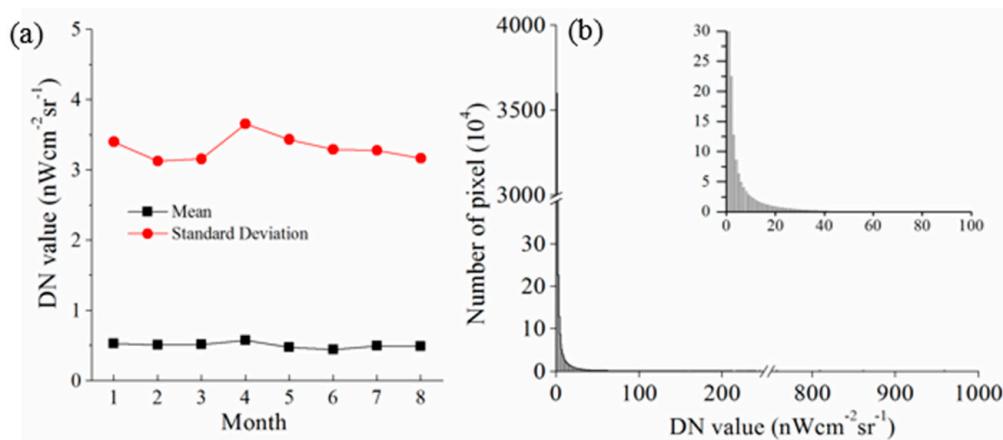


Figure 2. Quality assessment of VIIRS nighttime light data. (a) The mean and standard deviation of the VIIRS nighttime light data in the region with consistent observations in China from January to August in 2015; (b) distribution of DN values in the annually composited VIIRS nighttime light data for China in 2015.

Moderate-resolution Imaging Spectroradiometer (MODIS) 16-day composite images of the NDVI in 2015 were obtained from the National Aeronautics and Space Administration (NASA) [40]. These data included NDVI images from January to August of 2015, and each image was a composite of the maximum NDVI values observed over 16 days at a resolution of 500 m. These data were radiometrically calibrated, precisely georeferenced, and corrected for atmospheric effects before distribution. Furthermore, based on the work of Lu et al. [41], we produced an annual maximum value composite (MVC) image of the NDVI for 2015. Although an annual MVC image of the NDVI conceals land-cover changes that occur during the year, it can efficiently reduce cloud contamination while capturing changes in spatial vegetation characteristics [42].

MODIS eight-day composite images of nighttime LST (MOD11A2) in 2015 were also obtained from NASA [40]. These images were available from January to August of 2015 at a resolution of 1 km. The nighttime LST data were retrieved based on the generalized split window algorithm with thermal infrared data obtained from the MODIS sensors. The nighttime LST data were preprocessed with geometric correction and cloud removal and then composited based on average nighttime LST values over an eight-day period [43]. Studies have shown that nighttime LST data are better suited for distinguishing between urban and non-urban land areas compared with daytime surface temperature data [44]. Therefore, nighttime LST data were used. According to Mildrexler, Zhao, and Running [45], we produced an annual nighttime LST image for 2015 using the MVC method.

Landsat-8 Operational Land Imager (OLI) images for the evaluation areas in 2015 with a spatial resolution of 30 m (Table 3) were obtained from the USGS website [46].

Table 3. Landsat-8 OLI data.

Evaluation Area	Path/Row	Date
Kunming	129/43	5 January 2015
Zhengzhou	124/36	15 September 2015
Yinchuan	129/33	1 September 2015
Hohhot	126/32	30 October 2015
Jiamusi	115/28	15 September 2015
Lanzhou	130/35	10 October 2015
Urumqi	143/29	24 September 2015

Biome boundaries were obtained from the Terrestrial Ecoregions Database provided by the World Wildlife Fund for Nature (WWF) [47]. Moreover, Chinese administrative boundaries (1:4,000,000 scale) were obtained as GIS files from the National Geomatics Center of China [48].

3. Method

3.1. Extraction of Urban Land Area based on Different Methods

3.1.1. Local-Optimized Thresholding (LOT)

Following the methods suggested by Henderson et al. [21], Cao et al. [35], and Liu et al. [31], we first extracted urban land area in 2015 for each evaluation area using the Landsat-8 OLI data through visual interpretation. Then, we determined the optimal thresholds for each evaluation area using the urban land area extracted from the Landsat-8 OLI data as reference data. Specifically, referencing Bhatti and Tripathi's 2014 analysis [49], we used bands 4, 5, and 7 of the Landsat-8 OLI data to produce false color images of each city, and utilized the high-spatial resolution imageries in Google Earth to identify the standards for interpretation (i.e., color, shape, and texture). Based on these standards, we extracted urban land with Landsat-8 OLI data. After that, optimal thresholds were selected based on the time that the urban land area extracted using the VIIRS nighttime light data best matched the reference data with regard to the spatial extent. The selection of optimal thresholds can be summarized using the following equation:

$$\begin{aligned} & \text{Maximize } Kappa_j = f(T_j) \\ & \text{Subject to } T_j \in [VIIRS_j^{\min}, VIIRS_j^{\max}] \end{aligned} \quad (1)$$

where T_j is the threshold for nighttime light in the j th evaluation area, $Kappa_j$ is the kappa coefficient calculated using the urban land area extracted based on the T_j threshold from the VIIRS nighttime light and reference data, and $VIIRS_j^{\min}$ and $VIIRS_j^{\max}$ are nighttime light minimum and maximum values for the j th evaluation area in the VIIRS nighttime light data, respectively. Finally, we extracted areas with values greater than optimal thresholds in the VIIRS nighttime light data and regarded them as urban land for each evaluation area in 2015.

3.1.2. Vegetation-Adjusted Nighttime Light Urban Index (VANUI)

Following Zhang et al. [33] and Li et al. [23], we calculated the VANUI based on both the VIIRS nighttime light data and NDVI data. The equation for calculating the VANUI is as follows:

$$VANUI_{(i,j)} = (1 - NDVI_{(i,j)}) \times VIIRS_{(i,j)}^{nor} \quad (2)$$

where $VANUI_{(i,j)}$ is the VANUI value for the i th pixel in the j th evaluation area, $NDVI_{(i,j)}$ is the NDVI value for the i th pixel, and $VIIRS_{(i,j)}^{nor}$ is the normalized DN value of nighttime light for the i th pixel. $VIIRS_{(i,j)}^{nor}$ is calculated using the following equation:

$$VIIRS_{(i,j)}^{nor} = \frac{VIIRS_{(i,j)} - VIIRS_j^{\min}}{VIIRS_j^{\max} - VIIRS_j^{\min}}, \quad (3)$$

where $VIIRS_{(i,j)}$ is the original DN value of nighttime light for the i th pixel in the j th evaluation area and $VIIRS_j^{\min}$ and $VIIRS_j^{\max}$ are minimum and maximum values of nighttime light for the j th evaluation area in the VIIRS nighttime light data, respectively.

We followed the approach reported by Li et al. [23] to extract urban land area based on the VANUI. Specifically, optimal VANUI thresholds were determined using urban land area extracted from the Landsat-8 OLI data as reference data. The selection of optimal thresholds can be expressed using the following equation:

$$\begin{aligned} & \text{Maximize } Kappa_j = f(T_j) \\ & \text{Subject to } T_j \in [VANUI_j^{\min}, VANUI_j^{\max}] \end{aligned} \quad (4)$$

where T_j is the VANUI threshold in the j th evaluation area, $Kappa_j$ is the kappa coefficient calculated using urban land area extracted based on the T_j threshold from the VANUI and reference data, and $VANUI_j^{\min}$ and $VANUI_j^{\max}$ are minimum and maximum values of the VANUI for the j th evaluation area, respectively. Pixels with VANUI values exceeding the optimal thresholds were extracted as urban pixels for each evaluation area in 2015.

3.1.3. Integrated Nighttime Lights, Normalized Difference Vegetation Index, and Land Surface Temperature Support Vector Machine Classification (INNL-SVM)

Following He et al. [5] and Xu et al. [8], we performed INNL-SVM to extract urban land area based on three steps, including training sample selection, SVM classification, and post-classification. First, we selected urban and non-urban land training samples for each evaluation area using urban land characteristics based on the VIIRS nighttime light data, NDVI data, and LST data. According to Yang et al. [36] and He et al. [5], we recognized pixels with nighttime light exceeding $1 \text{ nW} \cdot \text{cm}^{-2} \cdot \text{sr}^{-1}$ as potential urban land and calculated average and standard deviation values of potential urban land nighttime light data, NDVI data, and LST data. Because urban land has stronger nighttime light, lower NDVI values, and higher LST values, the following equation was used to obtain urban and non-urban land training samples based on average and standard deviation values:

$$DN_{(i,j)} = \begin{cases} 1 & VIIRS_{(i,j)} > T_j^{VIIRS} \& NDVI_{(i,j)} < T_j^{NDVI} \& LST_{(i,j)} > T_j^{LST} \\ 0 & VIIRS_{(i,j)} < T_j^{VIIRS} \& NDVI_{(i,j)} > T_j^{NDVI} \& LST_{(i,j)} < T_j^{LST} \end{cases}, \quad (5)$$

where $DN_{(i,j)}$ is the class value of 1 (urban) or 0 (non-urban) at the i th pixel in the j th region in the training sample image. $VIIRS_{(i,j)}$, $NDVI_{(i,j)}$, and $LST_{(i,j)}$ are nighttime light, NDVI, and LST values at the i th pixel in the j th region, respectively. T_j^{VIIRS} , T_j^{NDVI} , and T_j^{LST} denote threshold VIIRS, NDVI, and LST values in the j th region, respectively, which were calculated using the following formulas:

$$T_j^{VIIRS} = \bar{X}_j^{VIIRS} \bar{X}_j^{VIIRS} + S_j^{VIIRS}/2 \quad (6)$$

$$T_j^{NDVI} = \bar{X}_j^{NDVI} \bar{X}_j^{NDVI} - S_j^{NDVI}/2 \quad (7)$$

$$T_j^{LST} = \bar{X}_j^{LST} + S_j^{LST}/2 \quad (8)$$

where \bar{X}_j^{VIIRS} , \bar{X}_j^{NDVI} , and \bar{X}_j^{LST} represent the average values of the potential urban land nighttime light, NDVI, and LST data in the j th region, respectively. S_j^{VIIRS} , S_j^{NDVI} , and S_j^{LST} denote the standard deviation of the potential urban land nighttime light, NDVI, and LST data in the j th region, respectively.

Second, we extracted urban land area using the general SVM classifier based on the training samples. This method classifies potential urban land into urban and non-urban land using a hyperplane with maximal margin, which was derived by solving the following constrained quadratic programming problem:

$$\begin{aligned} \text{Maximize } W(\alpha) &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ \text{Subject to } &\left\{ \sum_{i=1}^n \alpha_i y_i = 0 \text{ and } 0 \leq \alpha_i \leq T \text{ for } i = 1, 2, \dots, n \right\}, \end{aligned} \quad (9)$$

where $x_i \in R_d$ represents the training sample vectors, and $y_i \in \{-1, +1\}$ represents the corresponding class label. $K(x_i, x_j)$ represents the kernel function, for which the radial basis function was selected with the only free parameter σ set to 1.0 in the following formula:

$$K(x_i, x_j) = e^{(-|x_i - x_j|^2 / 2\sigma^2)}. \quad (10)$$

After classification was complete, we further performed post-classification, in which we recognized pixels with nighttime light values less than the regional average as non-urban pixels. The post-classification process can be summarized by the following formula:

$$DN_{(i,j)} = \begin{cases} 0 & VIIRS_{(i,j)} < \bar{X}_j^{VIIRS} \\ DN_{(i,j)} & otherwise \end{cases}, \quad (11)$$

where $DN_{(i,j)}$ is the class value of 1 (urban) or 0 (non-urban) at the i th pixel in the j th region in the 2015 urban land image. After performing the three steps, we obtained the urban land extent in each evaluation area.

3.2. Accuracy Assessment of Urban Land

According to Liu et al. [31] and He et al. [5], we used the urban land area extracted from Landsat-8 OLI images to assess the accuracy of urban land area extracted using the LOT, VANUI, and INNL-SVM methods, respectively. The kappa coefficient was used as the primary indicator, whereas overall accuracy (OA), commission error (CE), omission error (OE), quantity of disagreement (QD), and allocation of disagreement (AD) were used as auxiliary indicators. Specifically, the kappa coefficient was calculated using the following formula [50]:

$$Kappa_j^m = \frac{TA_j \cdot (N_j^m + O_j^m) - [(N_j^m + L_j^m) \cdot (N_j^m + V_j^m) + (O_j^m + L_j^m) \cdot (O_j^m + V_j^m)]}{TA_j^2 - [(N_j^m + L_j^m) \cdot (N_j^m + V_j^m) + (O_j^m + L_j^m) \cdot (O_j^m + V_j^m)]}, \quad (12)$$

where $Kappa_j^m$ is the kappa coefficient for the m th method in the j th evaluation area, TA_j is the total area, N_j^m is the non-urban land area extracted using both the Landsat-8 OLI image and the m th method, O_j^m is the urban land area extracted using both the Landsat-8 OLI image and the m th method, L_j^m is the urban land area only extracted using the Landsat-8 OLI image, and V_j^m is the urban land area only extracted using the m th method. The OA, CE, and OE were calculated using the following equations:

$$OA_j^m = \frac{N_j^m + O_j^m}{TA_j} \quad (13)$$

$$CE_j^m = \frac{V_j^m}{V_j^m + O_j^m} \quad (14)$$

$$OE_j^m = \frac{L_j^m}{L_j^m + O_j^m}, \quad (15)$$

where OA_j^m , CE_j^m and OE_j^m are the OA, CE, and OE for the m th method in the j th evaluation area, respectively. The AD and QD were calculated using the following equations [51]:

$$AD_j^m = 2 \cdot \min\left(\frac{L_j^m}{TA_j}, \frac{V_j^m}{TA_j}\right) \quad (16)$$

$$QD_j^m = 1 - OA_j^m - AD_j^m \quad (17)$$

where AD_j^m and QD_j^m are the AD and QD for the m th method in the j th evaluation area, respectively.

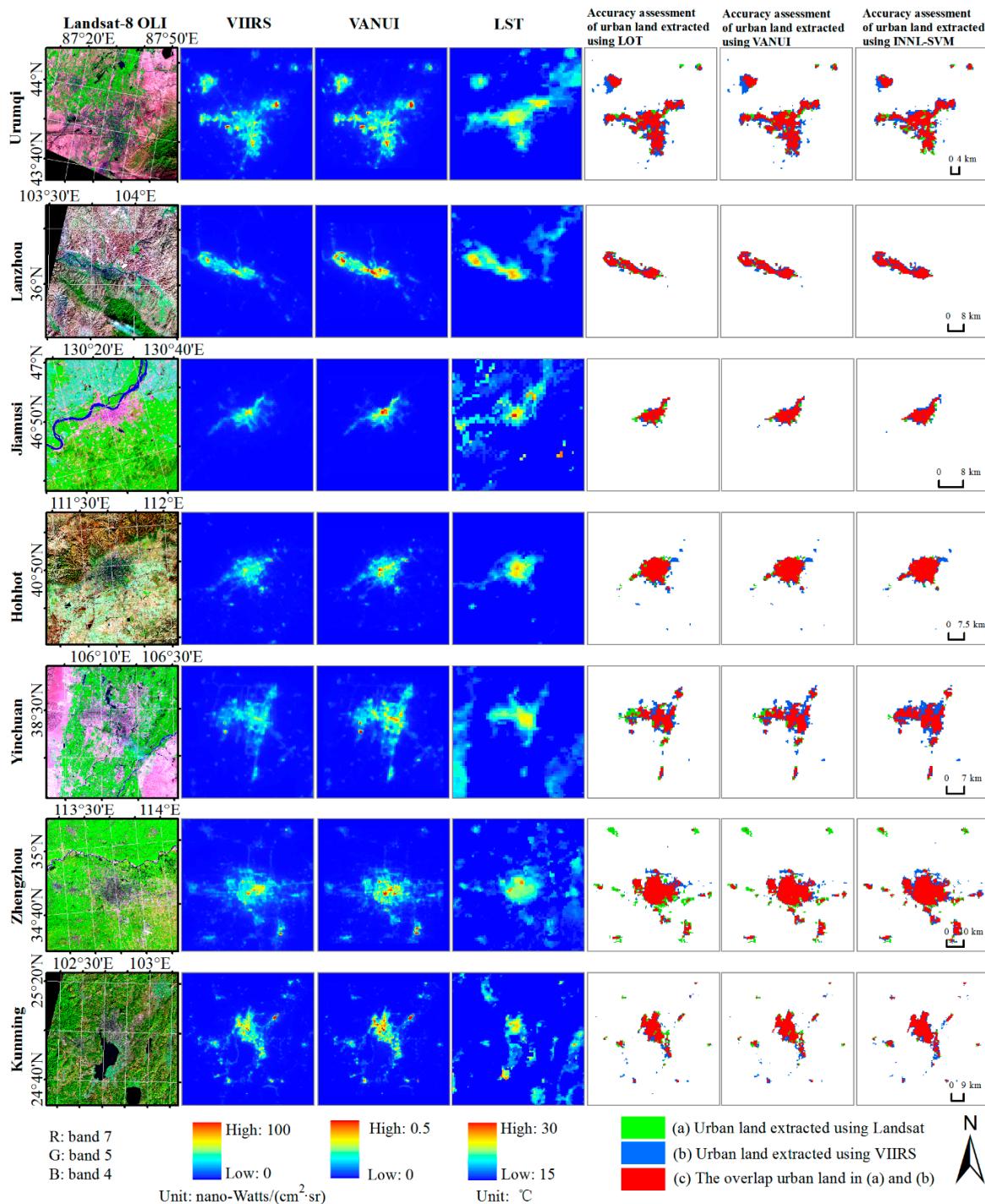
Specifically, the LOT, VANUI, and INNL-SVM were implemented using the ENVI 4.8/IDL8.0 software [52], and the accuracy assessment was performed using the ArcMap 10.0 software [53].

4. Results

Among the three methods, INNL-SVM had the best performance with an average kappa of 0.80, an average OA of 97.33%, an average QD of 1.49%, an average AD of 1.18%, an average CE of 26.89%, and an average OE of 8.44% (Figures 3–5). The LOT method had an average kappa of 0.75, an average OA of 96.92%, an average QD of 1.05%, an average AD of 2.03%, an average CE of 26.03%, and an average OE of 19.65% (Figures 3–5). The VANUI method had an average kappa of 0.78, an average OA of 97.20%, an average QD of 1.16%, and an average AD of 1.64% (Figures 3–5). INNL-SVM had the highest average kappa (2.56% higher than that of VANUI and 6.67% higher than that of LOT), the highest average OA (0.13% higher than VANUI and 0.41% higher than LOT), the lowest AD (0.46% lower than VANUI and 0.85% lower than LOT), and the lowest OE (5.25% lower than VANUI and 11.21% lower than LOT) (Figures 3–5).

In the Hohhot evaluation area, INNL-SVM showed the highest accuracy with a kappa of 0.86, an OA of 98.51%, a QD of 0.82%, an AQ of 0.67%, a CE of 24.58%, and an OE of 9.40% (Figures 3–5). In this area, the INNL-SVM accuracy was much higher than that of both the LOT and VANUI methods in terms of kappa (7.50% higher than LOT and 3.61% higher than VANUI), OA (0.60% higher than LOT and 0.26% higher than VANUI), AD (0.81% lower than LOT and 0.47% lower than VANUI), CE (3.89% lower than LOT and 1.10% lower than VANUI), and OE (7.71% lower than LOT and 4.47% lower than VANUI) (Figures 3–5).

In the Yinchuan evaluation area, INNL-SVM showed a relatively low accuracy, with a kappa of 0.72, an OA of 96.11%, a QD of 3.36%, an AD of 0.53%, a CE of 39.43%, and an OE of 4.51% (Figures 3–5). Nonetheless, the INNL-SVM accuracy was still higher than the accuracy of the LOT and VANUI methods, with much higher kappa values (7.46% and 4.35% higher, respectively) and OA (0.03% and 0.17% higher, respectively), as well as much lower AD (2.38% and 1.35% lower, respectively) and OE (20.39% and 11.57% lower, respectively) (Figures 3–5).

**Figure 3. Accuracy assessment.**

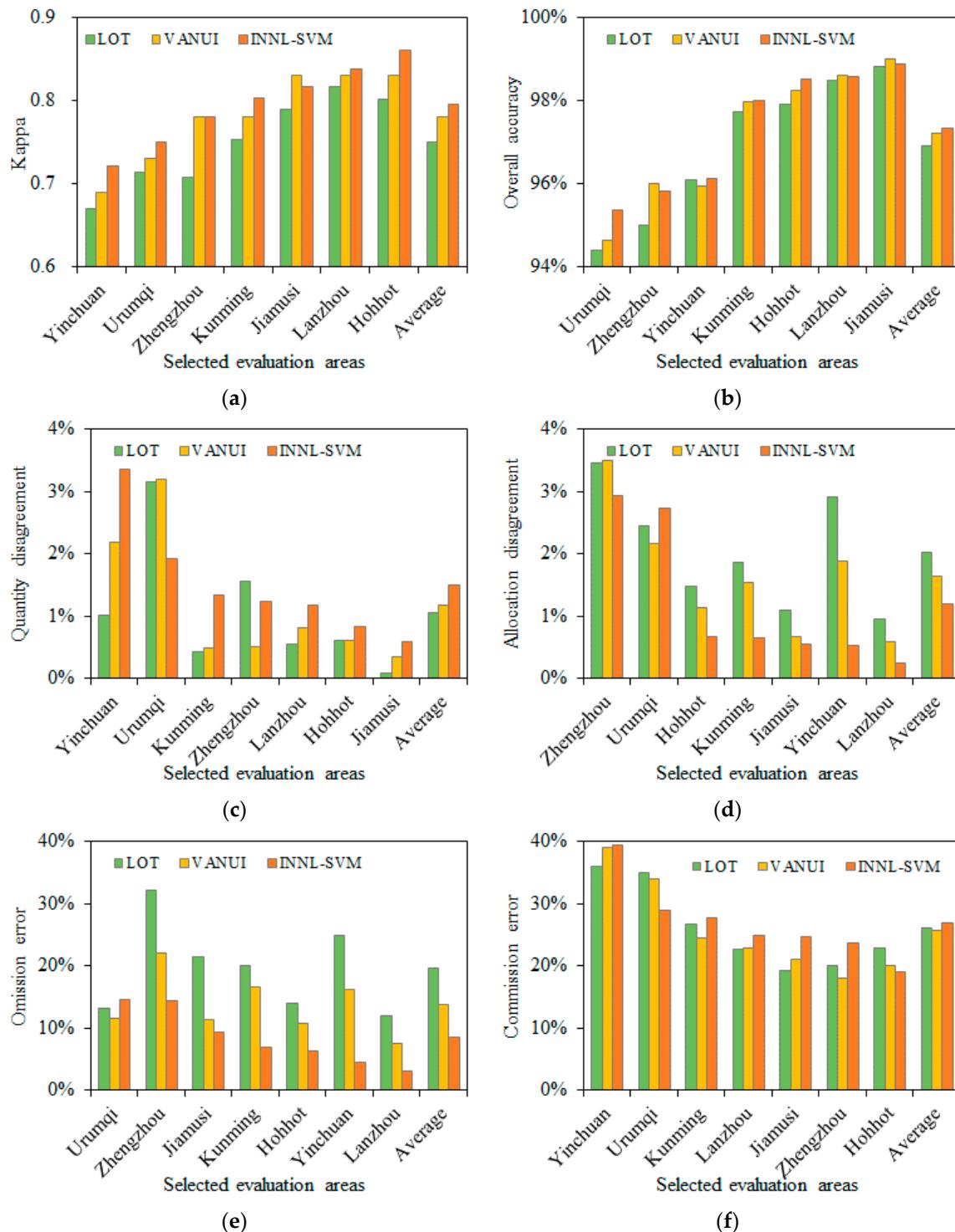


Figure 4. Accuracy evaluation of the different methods: (a) kappa; (b) overall accuracy; (c) quantity disagreement; (d) allocation disagreement; (e) omission error; and (f) commission error.

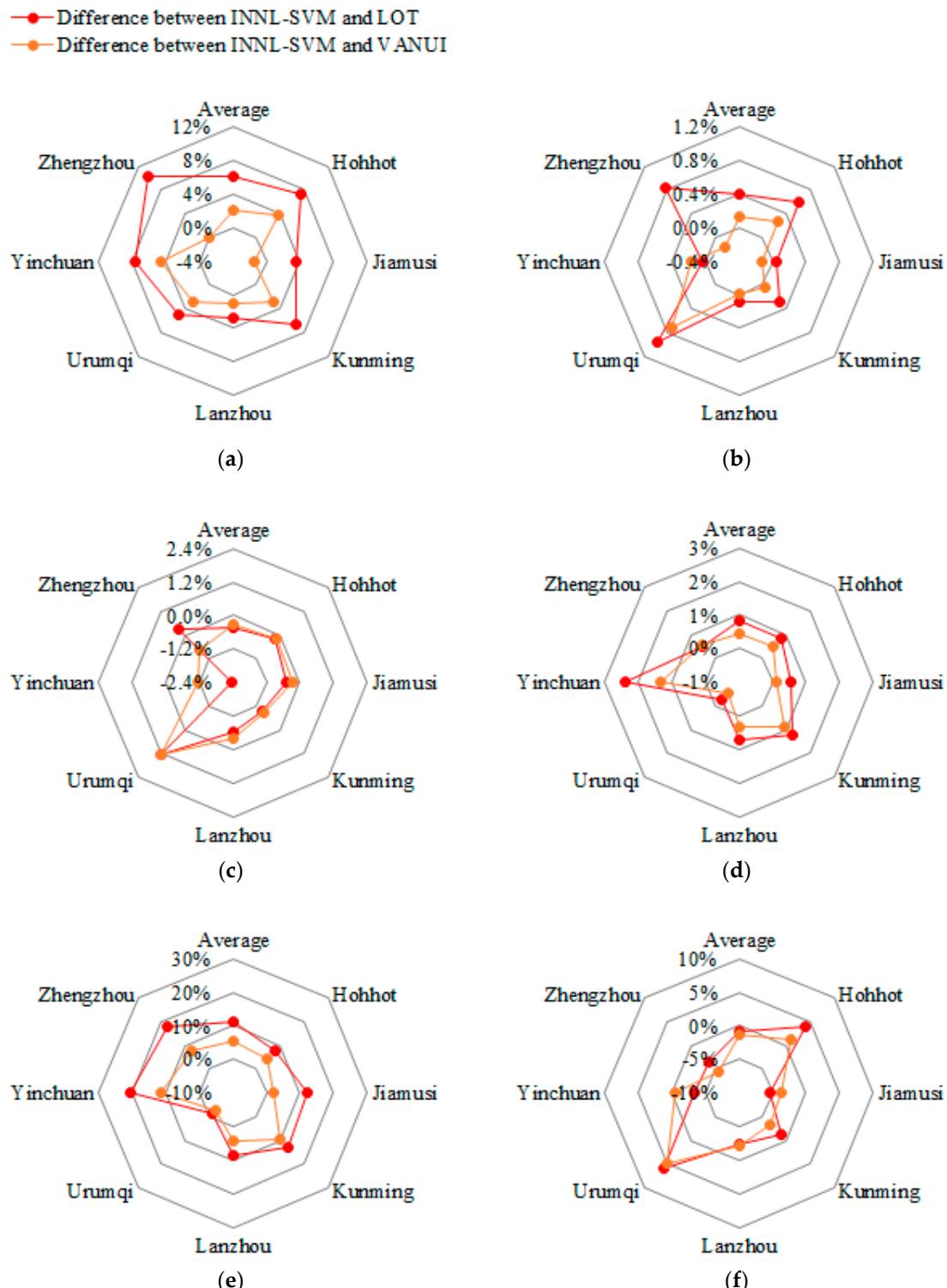


Figure 5. Accuracy comparison of the different methods. (a) Increase of kappa; (b) Increase of overall accuracy; (c) Decrease of quantity disagreement; (d) Decrease of allocation disagreement; (e) Decrease of omission error; (f) Decrease of commission error.

5. Discussion

5.1. INNL-SVM Provides a Reliable Approach for Rapidly and Accurately Extracting Urban Land Area Using VIIRS Nighttime Light Data

First, INNL-SVM can rapidly extract urban land information independently without ancillary data for facilitating sample selection. Among three popular methods, both the LOT and VANUI approaches require ancillary data to determine optimal thresholds [23,32,33]. However, numerous human and computational resources are needed to obtain one-time accurate ancillary data for large-scale studies. Therefore, it is difficult to extract urban land information for multiple years at large spatial scales using either the LOT or VANUI method [23,32,33]. By contrast, INNL-SVM can automatically determine thresholds for training sample selection and SVM classification according to regional characteristics of nighttime light, vegetation coverage, and LST, increasing its efficiency for extracting urban land area at large scales and over long time series [5,8].

Second, INNL-SVM clearly reduces errors arising from the overflow effect and low brightness of some urban areas in the VIIRS nighttime light data. Because the overflow effect exists in the VIIRS nighttime light data [54], extracting urban land with the LOT method based solely on nighttime light data resulted in obvious overestimation (Area A in Figure 6). In addition, the VIIRS overpass time is near 1:30 a.m., when some urban areas exhibit a decline in nighttime brightness [17], resulting in urban land area underestimation for the LOT method (Area B in Figure 6). The VANUI, which combines nighttime light and NDVI data, benefits from the characteristic that urban land vegetation coverage is generally lower than non-urban land [35], which reduces urban land CE and OE by mitigating the overflow effect and impacts of low brightness (Areas A and B in Figure 6). However, the difference between urban land and barren land vegetation coverage is not obvious. Therefore, barren land near urban areas results in obvious CE using the VANUI, specifically in arid and semiarid regions such as the Hohhot evaluation area (Area C in Figure 6). Moreover, in arid and semiarid regions, some urban areas have relatively higher vegetation coverage than non-urban areas, and these urban areas would be omitted using the VANUI (Area D in Figure 6). The LST in urban areas is usually higher than the LST in surrounding non-urban areas because of the combined effects of impervious urban surfaces, buildings, and atmospheric pollutants [5,55]. In arid/semitropic and pre-urban regions without obvious differences in vegetation coverage between urban land and non-urban land, INNL-SVM integrating LST data can be used to better distinguish urban land and non-urban land and thus improve the accuracy of extracted urban land area (Areas C and D in Figure 6).

Third, existing methods for improving LST data will enable increasing INNL-SVM accuracy in the near future. Recently, some researchers have successfully used geographically weighted regression (GWR) and temperature cycle models (TCMs) to improve the spatial resolution of the MODIS LST data and reduce the effects of meteorological factors (e.g., clouds, wind, and snow) [56,57]. Such approaches provide reliable means for increasing INNL-SVM accuracy, which is limited by spatial resolution and the LST data quality (Figure 4 and Areas E and F in Figure 6). Therefore, INNL-SVM has great potential for effectively extracting urban land areas using the VIIRS nighttime light data over multiple years at a large scale.

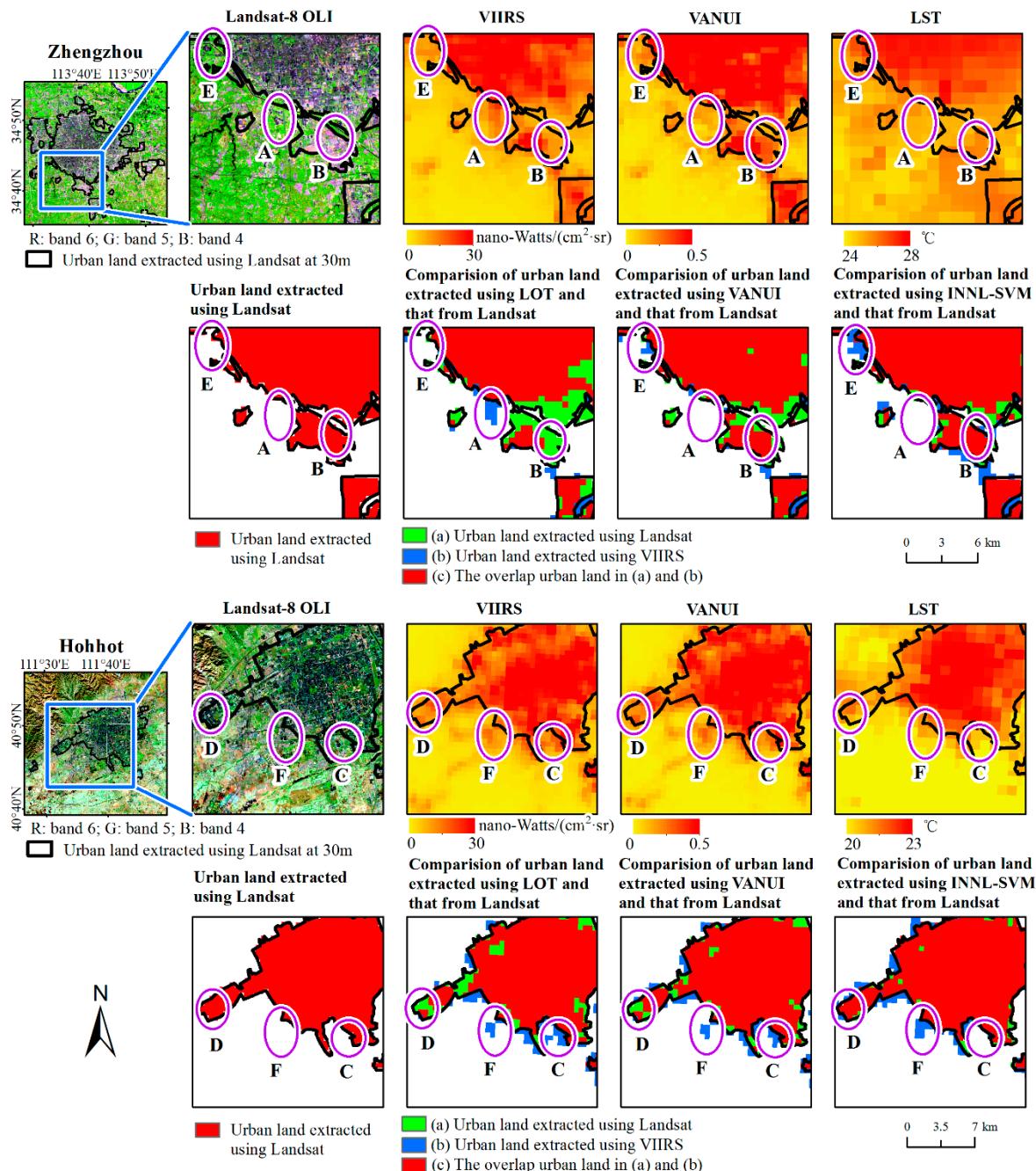


Figure 6. Spatial comparison of the results of different methods (in columns) for two regions (Zhengzhou as the 1st example, and Hohhot as the 2nd example).

6. Conclusions

INNL-SVM had the highest accuracy for extracting urban land area using the VIIRS nighttime light data. The average NNL-SVM kappa for the seven evaluation areas was 0.80, which was 2.56% higher than VANUI and 6.67% higher than LOT. The superior performance of INNL-SVM is mainly attributed to the integration of nighttime light, NDVI, and LST data. Through integration, INNL-SVM utilizes both the social (intensity of human activities) and physical (vegetation cover conditions and surface temperature) indicators of urban development. This effectively resolves urban land overestimation and underestimation arising from the overflow effect and low light brightness in the VIIRS nighttime light data and similar vegetation cover between urban land and non-urban land in arid/semi-arid regions.

Additionally, INNL-SVM can rapidly extract urban land area by automatically selecting training samples. Thus, INNL-SVM provides an effective approach for the timely and accurate extraction of urban land area using VIIRS nighttime light data and has the potential for wide application.

Acknowledgments: We would like to express our respect and gratitude to the anonymous reviewers and editors for their professional comments and suggestions. This work was supported in part by the National Natural Science Foundation of China (Grant Nos. 41621061 & 41501195).

Author Contributions: Yinyin Dou and Zhifeng Liu collected and processed the data, performed the analysis, and wrote the paper. Chunyang He conceived and designed the study and wrote the paper. Huanbi Yue contributed to the analysis of the data. All authors reviewed and edited the draft, approved the submitted manuscript, and agreed to be listed and accepted the version for publication.

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

References

1. Grimm, N.B.; Faeth, S.H.; Golubiewski, N.E.; Redman, C.L.; Wu, J.G.; Bai, X.M.; Briggs, J.M. Global change and the ecology of cities. *Science* **2008**, *319*, 756–760. [[CrossRef](#)] [[PubMed](#)]
2. Seto, K.C.; Fragkias, M.; Guneralp, B.; Reilly, M.K. A meta-analysis of global urban land expansion. *PLoS ONE* **2011**, *6*, e23777. [[CrossRef](#)] [[PubMed](#)]
3. Wu, J. Urban ecology and sustainability: The state-of-the-science and future directions. *Landscape. Urban Plan.* **2014**, *125*, 209–221. [[CrossRef](#)]
4. Angel, S.; Parent, J.; Civco, D.L.; Blei, A.; Potere, D. The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050. *Prog. Plan.* **2011**, *75*, 53–107. [[CrossRef](#)]
5. He, C.; Liu, Z.; Tian, J.; Ma, Q. Urban expansion dynamics and natural habitat loss in China: A multiscale landscape perspective. *Glob. Chang. Biol.* **2014**, *20*, 2886–2902. [[CrossRef](#)] [[PubMed](#)]
6. Liu, Z.; He, C.; Wu, J. General spatiotemporal patterns of urbanization: An examination of 16 world cities. *Sustainability* **2016**, *8*, 41. [[CrossRef](#)]
7. Liu, Z.; He, C.; Wu, J. The relationship between habitat loss and fragmentation during urbanization: An empirical evaluation from 16 world cities. *PLoS ONE* **2016**, *11*, e0154613. [[CrossRef](#)] [[PubMed](#)]
8. Xu, M.; He, C.; Liu, Z.; Dou, Y. How did urban land expand in China between 1992 and 2015? A multi-scale landscape analysis. *PLoS ONE* **2016**, *11*, e0154839. [[CrossRef](#)] [[PubMed](#)]
9. Seto, K.C.; Guneralp, B.; Hutyra, L.R. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 16083–16088. [[CrossRef](#)] [[PubMed](#)]
10. Liu, Z.; He, C.; Zhou, Y.; Wu, J. How much of the world's land has been urbanized, really? A hierarchical framework for avoiding confusion. *Landscape. Ecol.* **2014**, *29*, 763–771. [[CrossRef](#)]
11. Shi, K.; Huang, C.; Yu, B.; Yin, B.; Huang, Y.; Wu, J. Evaluation of NPP-VIIRS night-time light composite data for extracting built-up urban areas. *Remote Sens. Lett.* **2014**, *5*, 358–366. [[CrossRef](#)]
12. Shi, K.; Yu, B.; Huang, Y.; Hu, Y.; Yin, B.; Chen, Z.; Chen, L.; Wu, J. Evaluating the ability of NPP-VIIRS nighttime light data to estimate the gross domestic product and the electric power consumption of China at multiple scales: A comparison with DMSP-OLS data. *Remote Sens.* **2014**, *6*, 1705–1724. [[CrossRef](#)]
13. Liao, L.B.; Weiss, S.; Mills, S.; Hauss, B. Suomi NPP VIIRS day-night band on-orbit performance. *J. Geophys. Res. Atmos.* **2013**, *118*, 12705–12718. [[CrossRef](#)]
14. Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.; et al. Global land cover mapping at 30 m resolution: A pok-based operational approach. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 7–27. [[CrossRef](#)]
15. Shi, K.F.; Yu, B.L.; Hu, Y.J.; Huang, C.; Chen, Y.; Huang, Y.X.; Chen, Z.Q.; Wu, J.P. Modeling and mapping total freight traffic in China using NPP-VIIRS nighttime light composite data. *GISci. Remote Sens.* **2015**, *52*, 274–289. [[CrossRef](#)]
16. Yu, B.; Shi, K.; Hu, Y.; Huang, C.; Chen, Z.; Wu, J. Poverty evaluation using NPP-VIIRS nighttime light composite data at the county level in China. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 1217–1229. [[CrossRef](#)]
17. Elvidge, C.D.; Baugh, K.E.; Zhizhin, M.; Hsu, F.-C. Why VIIRS data are superior to DMSP for mapping nighttime lights. *Proc. Asia-Pac. Adv. Netw.* **2013**, *35*, 62–69. [[CrossRef](#)]

18. Schueler, C.F.; Lee, T.F.; Miller, S.D. VIIRS constant spatial-resolution advantages. *Int. J. Remote Sens.* **2013**, *34*, 5761–5777. [[CrossRef](#)]
19. Sharma, R.C.; Tateishi, R.; Hara, K.; Gharechelou, S.; Iizuka, K. Global mapping of urban built-up areas of year 2014 by combining MODIS multispectral data with VIIRS nighttime light data. *Int. J. Digit. Earth* **2016**, *9*, 1004–1020. [[CrossRef](#)]
20. Townsend, A.C.; Bruce, D.A. The use of night-time lights satellite imagery as a measure of Australia’s regional electricity consumption and population distribution. *Int. J. Remote Sens.* **2010**, *31*, 4459–4480. [[CrossRef](#)]
21. Henderson, M.; Yeh, E.T.; Gong, P.; Elvidge, C.; Baugh, K. Validation of urban boundaries derived from global night-time satellite imagery. *Int. J. Remote Sens.* **2003**, *24*, 595–609. [[CrossRef](#)]
22. Xie, Y.; Weng, Q. Updating urban extents with nighttime light imagery by using an object-based thresholding method. *Remote Sens. Environ.* **2016**, *187*, 1–13. [[CrossRef](#)]
23. Li, Q.; Lu, L.; Weng, Q.; Xie, Y.; Guo, H. Monitoring urban dynamics in the southeast U.S.A. Using time-series DMSP/OLS nightlight imagery. *Remote Sens.* **2016**, *8*, 578. [[CrossRef](#)]
24. Zhang, X.; Li, P.; Cai, C. Regional urban extent extraction using multi-sensor data and one-class classification. *Remote Sens.* **2015**, *7*, 7671–7694. [[CrossRef](#)]
25. Elvidge, C.D.; Baugh, K.E.; Kihn, E.A.; Kroehl, H.W.; Davis, E.R. Mapping city lights with nighttime data from the DMSP operational linescan system, photogramm. *Photogramm. Eng. Remote Sens.* **1997**, *63*, 727–734.
26. Sutton, P. Modeling population density with night-time satellite imagery and GIS. *Comput. Environ. Urban* **1997**, *21*, 227–244. [[CrossRef](#)]
27. Imhoff, M.L.; Lawrence, W.T.; Stutzer, D.C.; Elvidge, C.D. A technique for using composite DMSP/OLS “city lights” satellite data to map urban area. *Remote Sens. Environ.* **1997**, *61*, 361–370. [[CrossRef](#)]
28. Ma, T.; Zhou, Y.; Zhou, C.; Haynie, S.; Pei, T.; Xu, T. Night-time light derived estimation of spatio-temporal characteristics of urbanization dynamics using DMSP/OLS satellite data. *Remote Sens. Environ.* **2015**, *158*, 453–464. [[CrossRef](#)]
29. Milesi, C.; Elvidge, C.D.; Nemani, R.R.; Running, S.W. Assessing the impact of urban land development on net primary productivity in the Southeastern United States. *Remote Sens. Environ.* **2003**, *86*, 401–410. [[CrossRef](#)]
30. He, C.; Shi, P.; Li, J.; Chen, J.; Pan, Y.; Li, J.; Zhuo, L.; Ichinose, T. Restoring urbanization process in China in the 1990s by using non-radiance-calibrated DMSP/OLS nighttime light imagery and statistical data. *Chin. Sci. Bull.* **2006**, *51*, 1614–1620. [[CrossRef](#)]
31. Liu, Z.; He, C.; Zhang, Q.; Huang, Q.; Yang, Y. Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008. *Landsc. Urban Plan.* **2012**, *106*, 62–72. [[CrossRef](#)]
32. Zhou, Y.; Smith, S.J.; Zhao, K.; Imhoff, M.; Thomson, A.; Bond-Lamberty, B.; Asrar, G.R.; Zhang, X.; He, C.; Elvidge, C.D. A global map of urban extent from nightlights. *Environ. Res. Lett.* **2015**, *10*, 054011. [[CrossRef](#)]
33. Zhang, Q.; Schaaf, C.; Seto, K.C. The vegetation adjusted NTL urban index: A new approach to reduce saturation and increase variation in nighttime luminosity. *Remote Sens. Environ.* **2013**, *129*, 32–41. [[CrossRef](#)]
34. Jing, W.; Yang, Y.; Yue, X.; Zhao, X. Mapping urban areas with integration of DMSP/OLS nighttime light and MODIS data using machine learning techniques. *Remote Sens.* **2015**, *7*, 12419–12439. [[CrossRef](#)]
35. Cao, X.; Chen, J.; Imura, H.; Higashi, O. A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data. *Remote Sens. Environ.* **2009**, *113*, 2205–2209. [[CrossRef](#)]
36. Yang, Y.; He, C.Y.; Zhang, Q.F.; Han, L.J.; Du, S.Q. Timely and accurate national-scale mapping of urban land in China using defense meteorological satellite program’s operational linescan system nighttime stable light data. *J. Appl. Remote Sens.* **2013**, *7*, 073535. [[CrossRef](#)]
37. China Statistical Yearbooks Database. Available online: <http://tongji.cnki.net/overseas/engnavi/navidefault.aspx> (accessed on 18 August 2016).
38. NOAA National Center for Environmental Information. Available online: <http://ngdc.noaa.gov> (accessed on 18 August 2016).
39. Mills, S.; Weiss, S.; Liang, C. VIIRS day/night band (DNB) stray light characterization and correction. *Proc. SPIE* **2013**, *8866*, 350–354.
40. National Aeronautics and Space Administration (NASA). Available online: <http://ladsweb.nascom.nasa.gov> (accessed on 18 August 2016).

41. Lu, D.; Tian, H.; Zhou, G.; Ge, H. Regional mapping of human settlements in southeastern China with multisensor remotely sensed data. *Remote Sens. Environ.* **2008**, *112*, 3668–3679. [[CrossRef](#)]
42. Holben, B.N. Characteristics of maximum-value composite images from temporal AVHRR data. *Int. J. Remote Sens.* **1986**, *7*, 1417–1434. [[CrossRef](#)]
43. Wan, Z. *Collection-5 MODIS Land Surface Temperature Products Users' Guide*; University of California: Santa Barbara, CA, USA, 2009.
44. Zakšek, K.; Oštir, K. Downscaling land surface temperature for urban heat island diurnal cycle analysis. *Remote Sens. Environ.* **2012**, *117*, 114–124. [[CrossRef](#)]
45. Mildrexler, D.J.; Zhao, M.; Running, S.W. Testing a MODIS global disturbance index across North America. *Remote Sens. Environ.* **2009**, *113*, 2103–2117. [[CrossRef](#)]
46. United States Geological Survey (USGS). Available online: <http://glovis.usgs.gov.com> (accessed on 18 August 2016).
47. World Wildlife Fund for Nature (WWF). Available online: <http://worldwildlife.org/publications/terrestrial-ecoregions-of-the-world> (accessed on 18 February 2017).
48. National Geomatics Center of China. Available online: <http://ngcc.sbsm.gov.cn> (accessed on 18 February 2017).
49. Bhatti, S.S.; Tripathi, N.K. Built-up area extraction using Landsat 8 OLI imagery. *GISci. Remote Sens.* **2014**, *51*, 445–467. [[CrossRef](#)]
50. Congalton, R.G.; Oderwald, R.G.; Mead, R.A. Assessing Landsat classification accuracy using discrete multivariate-analysis statistical techniques. *Photogramm. Eng. Remote Sens.* **1983**, *49*, 1671–1678.
51. Pontius, R.G.; Millones, M. Death to kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *Int. J. Remote Sens.* **2011**, *32*, 4407–4429. [[CrossRef](#)]
52. ITT Visual Information Solutions (ITT). The ENVI 4.8. 2010. Available online: <http://www.itvis.com> (accessed on 17 February 2017).
53. Environmental Systems Research Institute (ESRI). The ArcGIS 10.0. 2010. Available online: <http://www.esri.com> (accessed on 17 February 2017).
54. Baugh, K.; Hsu, F.C.; Elvidge, C.D.; Zhizhin, M. Nighttime lights compositing using the VIIRS day-night band: Preliminary results. *Proc. Asia-Pac. Adv. Netw.* **2013**, *35*, 70. [[CrossRef](#)]
55. Buyantuyev, A.; Wu, J. Urban heat islands and landscape heterogeneity: Linking spatiotemporal variations in surface temperatures to land-cover and socioeconomic patterns. *Landsc. Ecol.* **2010**, *25*, 17–33. [[CrossRef](#)]
56. Duan, S.-B.; Li, Z.-L. Spatial downscaling of MODIS land surface temperatures using geographically weighted regression: Case study in Northern China. *IEEE Trans. Geosci. Remote* **2016**, *54*, 6458–6469. [[CrossRef](#)]
57. Zhan, W.; Huang, F.; Quan, J.; Zhu, X.; Gao, L.; Zhou, J.; Ju, W. Disaggregation of remotely sensed land surface temperature: A new dynamic methodology. *J. Geophys. Res. Atmos.* **2016**, *121*, 10538–10554. [[CrossRef](#)]



© 2017 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 (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).