J. Geogr. Sci. 2016, 26(3): 325-338
DOI: 10.1007/s11442-016-1271-6

Extracting urban areas in China using DMSP/OLS nighttime light data integrated with biophysical composition information

CHENG Yang[1,2,3], *ZHAO Limin[1,2], WAN Wei[4], LI Lingling[1,2], YU Tao[1,2], GU Xingfa[1,2]

1. State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, CAS, Beijing 100101, China;
2. The Center for National Spaceborne Demonstration, Beijing 100101, China;
3. University of Chinese Academy of Sciences, Beijing 100049, China;
4. Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China

Abstract: DMSP/OLS nighttime light (NTL) image is a widely used data source for urbanization studies. Although OLS NTL data are able to map nighttime luminosity, the identification accuracy of distribution of urban areas (UAD) is limited by the overestimation of the lit areas resulting from the coarse spatial resolution. In view of geographical condition, we integrate NTL with Biophysical Composition Index (BCI) and propose a new spectral index, the BCI Assisted NTL Index (BANI) to capture UAD. Comparisons between BANI approach and NDVI-assisted SVM classification are carried out using UAD extracted from Landsat TM/ETM+ data as reference. Results show that BANI is capable of improving the accuracy of UAD extraction using NTL data. The average overall accuracy (OA) and Kappa coefficient of sample cities increased from 88.53% to 95.10% and from 0.56 to 0.84, respectively. Moreover, with regard to cities with more mixed land covers, the accuracy of extraction results is high and the improvement is obvious. For other cities, the accuracy also increased to varying degrees. Hence, BANI approach could achieve better UAD extraction results compared with NDVI-assisted SVM method, suggesting that the proposed method is a reliable alternative method for a large-scale urbanization study in China's mainland.

Keywords: urban area distribution; DMSP/OLS; biophysical composition index; BANI; China

1 Introduction

Human activities are predominantly concentrated in the urban areas. Accurate understanding of the spatiotemporal distribution of urban areas (UAD) is an effective way to unveil the mechanisms of interaction between land-use systems and terrestrial ecosystems. It also serves the basic needs for urban–rural planning, urban resource management, environmental

Received: 2015-04-30 **Accepted:** 2015-09-07
Foundation: National Civil Aerospace Pre-research Project (non-disclosure)
Author: Cheng Yang, PhD Candidate, specialized in land use. E-mail: minnie880508@163.com
***Corresponding author:** Zhao Limin, PhD, specialized in thermal infrared remote sensing. E-mail: zhaolm@radi.ac.cn

assessment, and global change research (Weng, 2012). Remote sensing technique can capture land use and land cover conveniently, objectively, and continuously. It is an effective approach for current UAD extraction when combined with urban administrative unit statistical data (Schneider, 2012; Schneider *et al.*, 2010; Weng, 2012). Nighttime light (NTL) remote sensing data has become a new approach for large-scale urbanization study and has attracted widespread attention because of its macroscopic perspective (Doll, 2008; Elvidge *et al.*, 2007; Lu *et al.*, 2014; Ma *et al.*, 2012; Potere *et al.*, 2009).

NTL signal is captured by Operational Linescan System (OLS) sensor on Defense Meteorological Satellite Program (DMSP). It is found that the night lighted areas on Earth are coincident with the distribution of population and energy consumption. Thus, NTL information provides an accurate, economic, and direct way to describe the global distribution and development of urban areas, making it a powerful tool for human activity study (Forbes, 2013; He *et al.*, 2013; Small and Elvidge, 2013; Wu *et al.*, 2013). The method for large-scale UAD extraction using DMSP/OLS data can be categorized as: 1) NTL data thresholding, and 2) combining multi-sensor remote sensing data and auxiliary products. The threshold method mainly includes empirical threshold (Elvidge *et al.*, 1997a; Elvidge *et al.*, 1997b), abruptly changing detection (Imhoff *et al.*, 1997), high resolution data comparison (Henderson *et al.*, 2003), and statistical data comparison (He *et al.*, 2006) methods. DMSP/OLS data are recorded as 6-bit digital numbers (DNs) that are often saturated in the core of the cities. This is why DNs detected from sensors are consistently less than the exact values (Zhang *et al.*, 2013). Meanwhile, distribution of light areas in urban fringe areas, small towns as well as connected regions between cities detected by the OLS is consistently larger compared to the spatial distribution of the associated settlements. This is due to the coarse spatial resolution of the OLS sensor and the disturbance in the signal. Since urban expansions of different economic levels and different periods in China are significantly dissimilar, the determination of threshold tends to be empirical, regional, and temporal. It remains difficulty to extract the UAD accurately for large-scale researches (Small *et al.*, 2011; Small *et al.*, 2005). A cluster-based threshold method was developed to delineate the urban extent, which the optimal threshold for each potential urban cluster is evaluated relying on urban cluster size and overall NTL magnitude (Zhou *et al.*, 2014).

The core of multi-sensor remote sensing assisted method is to provide the underlying surface properties of urban landscape, which could subsequently correct or eliminate the saturation of NTL data. Previous research shows that vegetation and impervious surfaces have a strong negative linear relationship (Weng, 2012). Thus, some researchers tried to use NTL data combined with Normalized Difference Vegetation Index (NDVI) data to carry out urbanization studies, including urban land cover pattern intensification (Zhang *et al.*, 2013), urban energy consumption evaluation (He *et al.*, 2013), urban population estimation (Zhuo *et al.*, 2009) and urban spatial distribution extraction (Cao *et al.*, 2009; Lu *et al.*, 2008; Pandey *et al.*, 2013). He *et al.* (2014) estimated natural habitat loss caused by urban sprawl in China over the period 1992–2012, integrating NTL data, NDVI and land surface temperature (LST). Although the NDVI data is a favorable choice to provide supporting information for extracting the UAD, there are still a few drawbacks: 1) NDVI cannot differentiate impervious surface and bare soil effectively. It is difficult to describe the underlying surface of low vegetated areas; 2) it is not suitable for the identification of fast-growing cities. In

most developed countries, the cities have been well planned and some of them even have a history of hundreds of years. The distributions of green land, residential areas and commercial zones would not change frequently. However, in developing countries, due to the rapid expansion and its resulting demolition and retrofitting, the distribution of vegetation tend to be irregular, (e.g., the bare soil and impervious surface are often mixed with vegetation). As a developing country with a vast territory, large population, and rapid economic growth (Wang *et al.*, 2012), China has been experiencing a swift process of urbanization. The rapid urbanization of China has led to many complex problems, such as environmental problems and resource shortages. Most existing methods of urbanization study are applicable to developed countries, which would not suit the conditions of developing countries. Consequently, it is necessary to search a more suitable UAD extraction method to carry out a large-scale NTL study of developing countries like China.

Biophysical Composition Index (BCI), firstly proposed by Deng and Wu (2012), is a simple and convenient spectral enhancement technique, which is designed to successfully discriminate three urban land cover compositions, vegetation, impervious surfaces, and soil (Scott *et al.*, 2014).It follows Ridd's V-I-S conceptual model (Zhang *et al.*, 2014) and was employed to quantitatively represent urban land cover principal materials for urban environment and landscape (Wu *et al.*, 2014). Without extra shortwave infrared or thermal infrared information, BCI can be applied to images of multiple spectral resolution and spatial resolution. BCI shows a strong positive correlation with urban impervious surface and a high negative correlation with vegetation. Moreover, BCI is capable of differentiating between bare soil and impervious surface to compensate for the insufficiency of NDVI data. NTL data combined with BCI will improve the reliability of UAD extraction results, because of the ability of V-I-S enhancement and the detection of human activities. Hence, we propose the BCI Assisted NTL Index (BANI) combining BCI with NTL. This study maps China's UAD by combining nighttime light and BCI through BANI index. We also perform accuracy assessments to quantify the efficiency of this method using the results of NDVI-assisted SVM classifier and Landsat TM/ETM+ data.

2 Study area and data

In this study, we mainly focus on the UAD extraction over China's mainland. According to *China Urban Statistical Yearbook 2013* complied by the National Bureau of Statistics of China at the end of 2012, the sum of provincial level administrative units of China's mainland was 31, including 23 provinces, 4 municipalities, and 5 autonomous regions. The total number of cities was 657, including 15 sub-province cities, 270 prefecture-level cities, and 368 county-level cities. Land area under prefecture-level cities was 4.76 million square kilometers and the total population was 1.26 billion.

2.1 DMSP/OLS stable nighttime light data

Version 4 DMSP/OLS NTL Time Series datasets were taken from the website of the National Geophysical Data Center at National Oceanic and Atmospheric Administration (http://ngdc.noaa.gov/eog/download.html), which have a swath width of 3000 km and are aggregated and composited to 30 arc second grids. This data was the average of digital

number values from annual VNIR channel stable nighttime light data, ranging from 0 to 63, which filters light pixels generated by accidental factors, such as gas flares and fires. Liu (2012) analyzed DMSP/OLS NTL statistical data from 1992 to 2010 in China and pointed out that DMSP/OLS NTL data in 2007 from satellite F16 could be used as the reference dataset, because it had the highest accumulated DN value. Therefore, the data employed in the research were DMSP/OLS stable NTL data from 2007, which were resampled with a 1 km spatial resolution and projected to Lambert Azimuthal Equal Area, and then clipped in accordance with the scale of China Vector Data.

2.2 MODIS data

The MODIS/Terra Surface Reflectance 8-Day L3 Global 500-m SIN Grid version 4 (MOD09A1) products were the primary source for the composition of BCI. The data were acquired from the Goddard Space Flight Center of National Aeronautics and Space Administration (NASA) LAADS Web (http://ladsweb.nascom.nasa.gov/data/search.html), which had been processed by radiometric calibration, atmospheric and aerosol correction, and edge distortion correction. To be consistent with DMSP/OLS NTL data, we selected 18 MOD09A1 datasets of good quality in September (the growing season) 2007 as the data source. MOD09A1 products contained quality assessment (QA) band. This band marked cloud state and water state, with which labeled data could be extracted through QA decoding for cloud mask and water mask. Besides that, re-projection and resampling were also conducted to correspond with DMSP/OLS NTL data.

2.3 Landsat TM/ETM+ data

Landsat TM/ETM+ images with a spatial resolution of 30 m were used for accuracy assessment. The data set was provided by the International Scientific & Technical Data Mirror Site, Computer Network Information Center, Chinese Academy of Sciences (http://www.gscloud.cn), on which radiometric and geometric corrections were processed. ETM+ data gap-fill process was accomplished by multi-images local adaptive regression analysis model (Liu *et al.*, 2010).

2.4 Auxiliary data

Data collection included a map of 1:4,000,000 scale offered by the National Geomatics Center of China and the *China Urban Statistical Yearbook 2007* complied by the National Bureau of Statistics of China.

3 Methodology

3.1 Division of study area

In the study of large-scale remote sensing classification and information extraction, if there were great differences of entropy between land information, classifying the study area on the basis of some strategies will contribute to the improvement of accuracy (Schneider *et al.*, 2010). China's mainland is so extensive that there were appreciable differences in population size and economic development level at different regions. It is necessary to divide the

study area into sub-regions according to provincial characteristics. Yang *et al.* (2013) carried out a study of China's mainland UAD extraction based on division using NTL data, with an average Kappa coefficient of 0.69. The definition of "city" in this paper is by virtue of population and economy. Thus, according to the differences of population and economy, China can be divided into eight economic zones according to the Strategy and Policy of Regional Coordinated Development Report by the Development Research Center (Liu *et al.*, 2002). The economic zones are Northeast China (NEC), Northern Coastal China (NCC), Eastern Coastal China (ECC), Southern Coastal China (SCC), Middle Reaches of the Yellow River (MRYLR), Middle Reaches of the Yangtze River (MRYTR), Southwest China (SWC), and Northwest China (NWC). Moreover, one sample city from each urban development level was chosen from each economic zone for accuracy assessment. The eight selected cities are Beijing, Chengdu, Harbin, Huainan, Quanzhou, Urumqi, Wuxi, and Xi'an, as illustrated in Figure 1. Table 1 reports the total population and GDP of each economic zone and sample city.

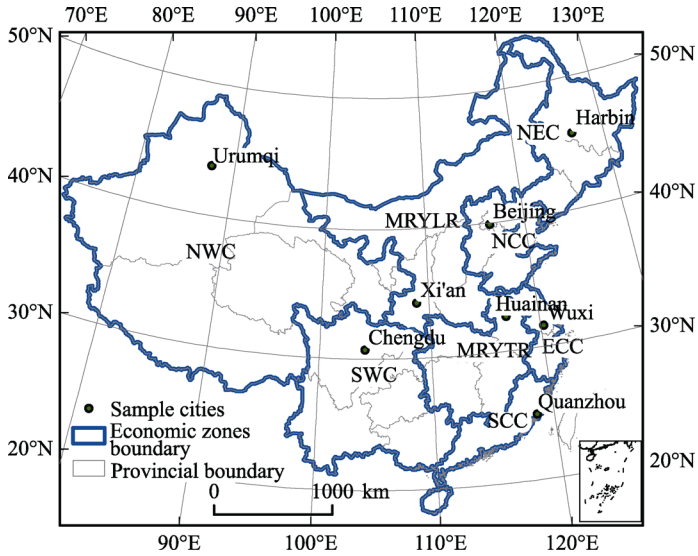

Figure 1 Division of economic zones and location of sample cities

Table 1 Socio-economic statistics of each economic zone in 2007

Economic zones/ Sample cities	Average GDP (billion RMB)	GDP per capita (RMB)	Population (million)
NEC / Harbin	779.11 / 175.67	21,197 / 37,052	108.52 / 4.75
NCC / Beijing	1351.98 / 920.76	38,003 / 60,045	190.58 / 11.42
ECC / Wuxi	1890.35 / 216.29	45,902 / 92,385	145.43 / 2.36
SCC / Quanzhou	1385.23 / 48.80	24,538 / 37,556	138.75 / 1.02
MRYLR / Xi'an	807.57 / 132.95	18,239 / 20,818	189.06 / 5.49
MRYTR / Huainan	782.38 / 24.77	13,844 / 15,851	225.40 / 1.66
SWC / Chengdu	5613.33 / 209.19	11,513 / 32,722	239.87 / 5.03
NWC / Urumqi	1648.11 / 80.97	13,672 / 31,806	61.58 / 2.22

3.2 Biophysical Composition Index (BCI) and NDVI calculations

(1) Processing of BCI

Two steps were performed in the pre-processing stage of the BCI calculation using MODIS surface reflectance data. Firstly, water pixels were masked out using QA of MOD09A1 products. Secondly, Tasseled Cap (TC) transformation was conducted. Transformation parameters proposed by Zhang *et al.* (2002) were adopted in the TC transformation of MOD09A1 data.

$$TC_i = C_{ij} * Band_j \tag{1}$$

where TC_i (i=1, 2, 3) are the first three TC components, namely brightness (TC1), greenness (TC2), and wetness (TC3); C_{ij} are the specific parameters of TC transformation listed in Table 2; $Band_j$ are the band numbers of MOD09A1 data.

Table 2 Tasseled Cap coefficients for MODIS

	Band 1 (Red)	Band 2 (Near-IR)	Band 3 (Blue)	Band 4 (Green)	Band 5 (M-IR)	Band 6 (M-IR)	Band 7 (M-IR)
TC1	0.3956	0.4718	0.3354	0.3834	0.3946	0.3434	0.2964
TC2	–0.3399	0.5952	–0.2129	–0.2222	0.4617	–0.1037	–0.4600
TC3	0.1084	0.0912	0.5065	0.4040	–0.2410	–0.4658	–0.5306

After the TC transformation, each derived TC component was linearly normalized within the range 0 to 1. Following the algorithm developed by Deng (2012), *BCI* was derived using Eq. (2) after pre-processing.

$$BCI = \frac{(H + L)/2 - V}{(H + L)/2 + V} \tag{2}$$

where H is "high albedo", the normalized TC1; V is "vegetation", the normalized TC2; and L is "low albedo", the normalized TC3. These three components can be given by the following formula:

$$H = \frac{TC1 - TC1_{\min}}{TC1_{\max} - TC1_{\min}} \tag{3}$$

$$V = \frac{TC2 - TC2_{\min}}{TC2_{\max} - TC2_{\min}} \tag{4}$$

$$L = \frac{TC3 - TC3_{\min}}{TC3_{\max} - TC3_{\min}} \tag{5}$$

where $TC_{\max}$ and $TC_{\min}$ are the maximum and minimum values of the ith TC component, respectively.

Figure 2 compares the distribution of NTL and BCI at each economic zone in 2007. Urban areas, indicated by a white tone, have the highest BCI values (positive). Soil and mixed land cover have a BCI value close to zero, and are displayed with a tone of medium gray. Vegetation has the lowest BCI value (negative). Note that the west of MRYLR and the northwest of NWC are desert areas, which also have high BCI values and displayed with a tone of white.

For further analysis, the statistical characteristics of BCI in each economic zone are necessary. Arithmetic mean $\overline{X}_{BCI}$ and standard deviation S_{BCI} of MODIS BCI data in each

economic zone were calculated for the preparation of training samples selection for the next step. Statistical results of MODIS BCI data in 2007 are shown in Table 3.

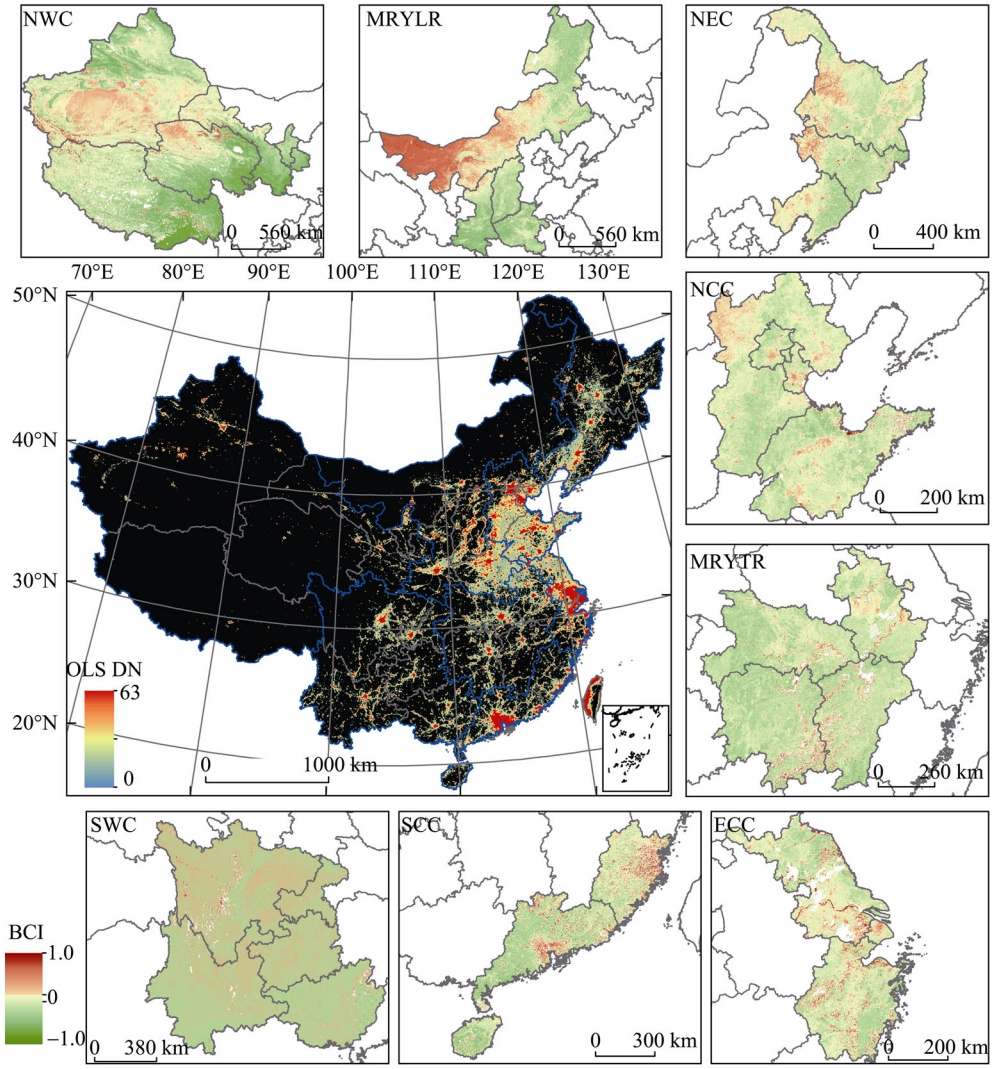

Figure 2 Contrast between NTL and BCI of the eight economic zones in 2007

Table 3 Discrepancies of MODIS BCI among eight economic zones in 2007

Economic zones	$\overline{X}_{BCI}$	S_{BCI}	$\overline{X}_{BCI} + S_{BCI}$
NEC	–0.1264	0.1002	–0.0262
NCC	–0.0572	0.1273	0.0701
ECC	–0.2808	0.1233	–0.1575
SCC	–0.2342	0.1521	–0.0821
MRYLR	0.0681	0.1968	0.2649
MRYTR	–0.3279	0.1041	–0.2238
SWC	–0.5054	0.0400	–0.4654
NWC	–0.4083	0.0979	–0.3104

(2) Processing of NDVI

To evaluate the performance of BCI-assisted classification for delineating UAD from NTL data, we employed NDVI-assisted classification for comparisons. NDVI data were calculated according to Eq. (6) using MOD09A1 datasets. Figure 3 shows the calculated NDVI data in China.

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{6}$$

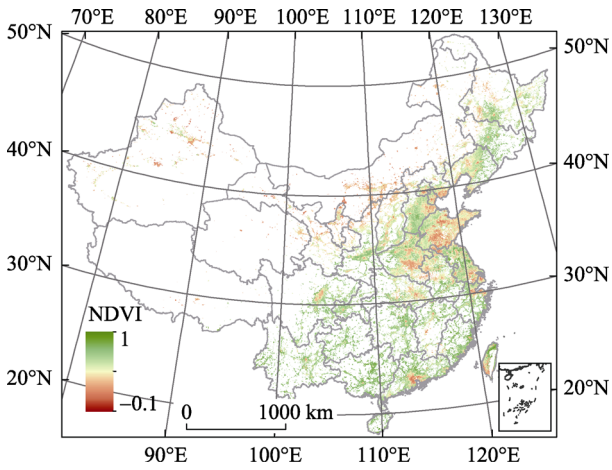

Figure 3 NDVI data calculated through MOD09A1 datasets of China in September 2007 and NDVI values show only light areas identified in DMSP/OLS NTL data

3.3 BANI: The BCI-Assisted NTL Index

We propose the BCI Assisted NTL Urban Index (BANI), which is based on the correlations between BCI, NTL, and urban surfaces. As illustrated in DMSP/OLS NTL data, urban areas are presented as light patches. Furthermore, closer towards the urban patch center, the pixels get brighter, as those areas are more developed. These areas are often the location with the higher density of impervious surfaces and the BCI values are between 0 and 1. BCI values of the rural surroundings with low presence of impervious surfaces are between –1 and 0. The BANI is to develop a robust index, which uses an urban impervious surface signal to increase inter-urban variability with NTL. The relationship between BCI and NTL is positive. While NTL values gradually increase towards the urban core, BCI values also get higher.We define BANI as:

$$BANI = NTL * (1 + BCI)^2 \tag{7}$$

where *BCI* is derived from MODIS, with the range between –1 and 1. The BANI calculation results in highlight values of urban core areas. To make BANI values in different economic zones consistent, we normalize NTL values to the range of 0 to 1. Figure 4 shows the calculated BANI data through NTL data and BCI data of China. A BANI value of 0 means the pixel cover with a corresponding NTL value of 0 (dark area). A higher BANI value indicates this area is closer to urban core.

After BANI calculation, in terms of BANI value and BCI value, we set thresholds according to Tables 3 and 4. Then label pixels with values less than the threshold as non-urban

pixels.

Table 4 Discrepancies of BANI among eight economic zones in 2007

Economic zones	$\overline{X}_{BANI}$	S_{BANI}	$\overline{X}_{BANI} + S_{BANI}$
NEC	0.0336	0.1094	0.1430
NCC	0.1430	0.2500	0.3930
ECC	0.1194	0.1886	0.3080
SCC	0.0712	0.1877	0.2589
MRYLR	0.0271	0.0995	0.1266
MRYTR	0.0230	0.0680	0.0910
SWC	0.0061	0.0220	0.0281
NWC	0.0013	0.0143	0.0156

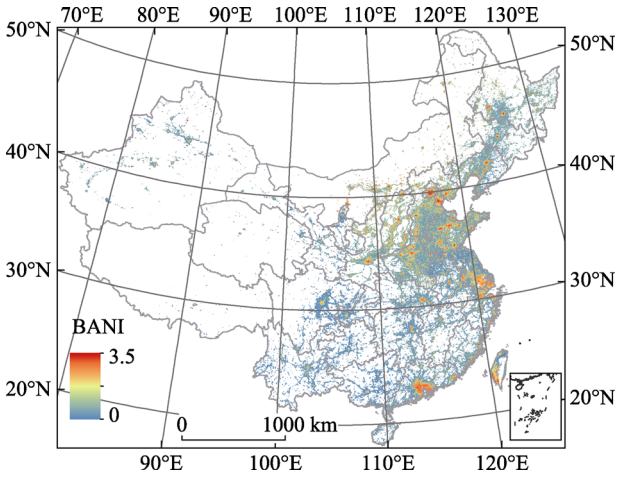

Figure 4 BANI data calculated through NTL data and BCI data of China in September 2007 and BANI values show only light areas identified in DMSP/OLS NTL data

3.4 NDVI-assisted SVM classification

The NDVI and DMSP/OLS NTL data were taken as the inputs of an iterative classification to classify urban and non-urban pixels, after initial training sets were built. Pixels with an OLS value of more than 30 were selected as potential pixels indicating an urbanized locality. Pixels with OLS values less than 30 and NDVI greater than 0.4 were set as non-urban samples. Then we perform SVM classification based on region-growing iterative method here. Post-classification procedure eliminated pseudo-urban pixels with NDVI values greater than 0.4 (Cao *et al.*, 2009).

4 Results and Accuracy Assessment

Landsat TM/ETM+ images were most commonly available for detecting urban areas and mapping their changes (Forsythe, 2004; Mundia and Aniya, 2005; Zhou *et al.*, 2008). The classification results using Landsat TM/ETM+ data were accurate enough to be used as reference maps for accuracy assessment since the fine spatial resolution (30 m) and spectrum

information (Henderson *et al.*, 2003; Small *et al.*, 2005). In this study, the UAD results extracted from TM/ETM+ multi-spectral data were captured based on SVM classification and the selection of training samples were through visual interpretation. Selection of training samples is a decisive factor for classification results, thus it would affect the accuracy of comparisons. The selection of these cities took a full consideration of the scale of the city and the level of development, so that they could reflect the applicability of this method comprehensively. We used Landsat TM/ETM+ data of eight sample cities for qualitative and quantitative UAD extraction results after comparing BANI approach and NDVI-assisted classification. In addition, these results were resampled to 1 km spatial resolution.

First, qualitative analysis of UAD extraction results was carried out (Figures 5c, 5d, and 5e). Compared to the urban land information extracted from Landsat TM/ETM+ data, both NDVI-assisted and BANI method were capable of acquiring urban information effectively. However, for Chengdu, Huainan, Wuxi, Urumqi, and Xi'an, the latter worked better than NDVI-assisted SVM algorithm in subtle features. The results were closer to that of reference data and the outline was more detailed.

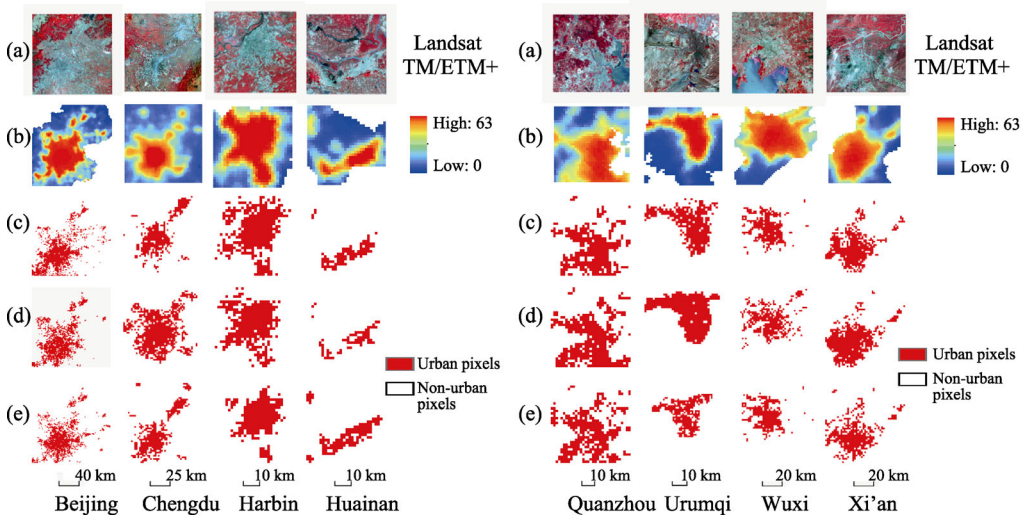

Figure 5 Landsat TM/ETM+ images (a); DMSP/OLS NTL images (b); Urban extent from Landsat TM/ETM+ classification (1 km) (c); Results of NDVI-assisted SVM method (d); Results of BANI approach (e)

We further analyzed UAD extraction results quantitatively. Accuracy assessments of UAD extraction results were performed (Table 5). The standards of these assessments include the number of urban pixels and two accuracy assessment indices of overall accuracy (OA) and Kappa coefficient. The OA and Kappa coefficients were calculated based on the error matrix from DMSP/OLS and Landsat TM/ETM+ results of each city. The OA of BANI algorithm ranged from 93.77% to 96.51%, and Kappa coefficient of that ranged from 0.79 to 0.88; while the OA and Kappa coefficient of NDVI-assisted SVM algorithm ranged from 86.11% to 92.18% and 0.47 to 0.63, respectively. The average OA and Kappa coefficient of BANI approach improved from 88.53% to 95.10% and from 0.56 to 0.84, respectively. It is possible to conclude that the accuracy of the former was better and that it had better coherency with TM/ETM+.

Table 5 Accuracy assessment of urban areas characteristic

Sample cities	BANI algorithm		NDVI-assisted SVM		Comparisons	
	OA (%)	Kappa	OA (%)	Kappa	OA (%)	Kappa
Beijing	94.18	0.88	90.27	0.57	3.91	0.31
Chengdu	96.51	0.87	86.74	0.47	9.77	0.40
Harbin	93.77	0.79	87.10	0.60	6.67	0.19
Huainan	96.33	0.87	92.18	0.48	4.15	0.39
Quanzhou	94.16	0.81	86.76	0.57	7.40	0.24
Wuxi	95.81	0.85	88.83	0.55	6.98	0.30
Urumqi	94.42	0.82	85.11	0.59	9.31	0.23
Xi'an	95.61	0.83	91.25	0.63	4.36	0.20
Average	95.10	0.84	88.53	0.56	6.57	0.28

Different cities present various land covers. Vegetation fraction is key to score the vegetation coverage of the land surface. Based on NDVI index, we adopted dimidiate pixel model to estimate the vegetation fraction of eight selected cities using Landsat TM/ETM data. Dimidiate pixel model assumed that spectral information observed by sensors consists of two parts, information contributed by the vegetation and the soil. Vegetation fraction is proposed by Eq. (8), in which NDVI is set as the input (Li *et al.*, 2004).

$$f_c = \frac{NDVI - NDVI_{soil}}{NDVI_{veg} - NDVI_{soil}} \tag{8}$$

where $NDVI_{soil}$ is NDVI value of no vegetation pixel, $NDVI_{veg}$ is NDVI value of pure vegetation pixel. We did statistical analysis of NDVI values including accumulative percent of each economic zone. Accumulative percent means the percentage of the total NDVI value which is less than or equal to a certain value. Based on the NDVI accumulative percent of each zone (Li *et al.*, 2004), the corresponding values of 5% confidence intervals were chosen as $NDVI_{soil}$. According to the results, we divided them into five levels, 0–0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8 and 0.8–1.0 (Table 6). The degree of mixing of the impervious surface, bare soil and vegetation was great when the cumulative values were between 0.2 and 0.8. During this interval, land cover is made from a mix of impervious surfaces, vegetation, and soil, which is called hybrid interval. The accuracy of results for cities, particularly, with hybrid interval values above 56%, such as Chengdu, Wuxi and Huainan had a high accuracy and improved considerably. OA improved from 86.74%, 88.83% and 92.18% to 96.51%, 95.81% and 96.33%, respectively. And Kappa coefficient improved from 0.48, 0.55 and 0.48 to 0.87, 0.85 and 0.87, respectively. The accuracies of results for cities with hybrid interval values between 50% and 56% (Urumqi in NWC, Beijing in NCC and Xi'an in SWC) also reached a high accuracy and showed significant improvements. OA improved from 85.11%, 90.27% and 91.25% to 94.42%, 94.18% and 95.61%, respectively. Kappa coefficient improved from 0.59, 0.57 and 0.63 to 0.82, 0.88 and 0.83, respectively.

Overall, UAD as well as area statistics of the extraction results of the BANI approach were consistent with that of finer spatial resolution. The accuracy was better than that of NDVI-assisted SVM classification. Thus, the results of BANI method are capable of reflecting the exact urban land-use information in China's mainland and are highly reliable.

Table 6 The vegetation fraction percentage of each division in eight selected cities in 2007 (%)

Vegetation fraction	0–0.2	0.2–0.4	0.4–0.6	0.6–0.8	0.8–1.0
Beijing	29.72	19.78	15.78	16.80	17.92
Chengdu	18.16	17.69	22.52	22.78	18.86
Harbin	20.33	13.15	14.17	18.52	33.84
Huainan	17.31	12.52	20.63	23.15	26.39
Quanzhou	25.93	12.71	13.31	19.95	28.09
Wuxi	24.02	20.61	20.10	18.03	17.25
Urumqi	38.45	29.14	15.13	7.98	9.30
Xi'an	16.81	13.42	15.90	22.81	31.06

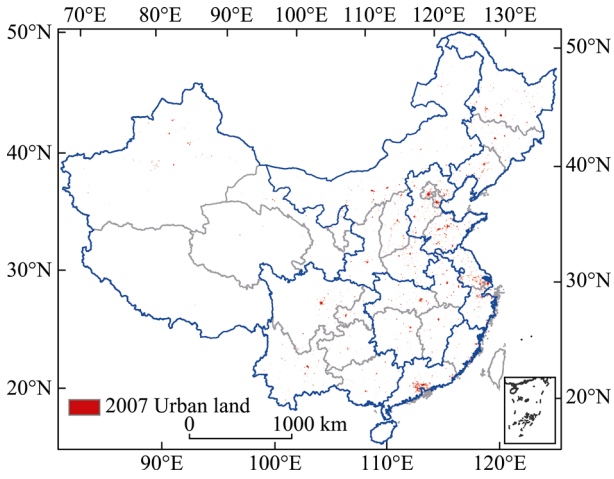

Figure 6 China's mainland urban land classified by BANI approach in 2007

Figure 6 presents the urban land map of China in 2007 using BANI procedures. In 2007, there were 51,617 km^2 of total urban land in China's mainland. Among them, the area of urban land situated in the coastal areas of China, such as NCC (11,427 km^2), ECC (7015 km^2), and SCC (7391 km^2) were much larger than those in land- locked regions of China, such as NWC (3331 km^2) and SWC (4973 km^2). The areas of urban land in central regions such as MRYLR, MRYTR and NEC were 6285 km^2, 5179 km^2 and 5998 km^2.

5 Discussion and conclusion

This study proposed an index BANI, which combined with DMSP/OLS NTL data and BCI to map the urban areas of China's mainland. This method effectively captured UAD and the average values of OA and Kappa coefficient reached 95.10% and 0.84, respectively.

We made a comparison of BANI approach and NDVI-assisted SVM algorithm through Landsat TM/ETM+ classification results of eight selected cities with different development level. From a qualitative point of view, the results of BANI approach showed a better result and a more detailed outline, similar to that of the reference data. Area statistical data of extraction results using BANI method were closer to the reference data. The accuracy of cities with hybrid interval values above 56% improved drastically. Hence, BANI algorithm is a reliable alternative method for extracting urban land data.

BCI has an advantage of enhancing vegetation, impervious surface, and soil to reflect the cover attributes of urban underlying surface. Combined with DMSP/OLS NTL data, BCI can improve UAD extraction accuracy of middle and small cities or developing cities. However, in terms of quantitative analysis and application, further investigations are required in the following aspects:

(1) Spatial visualization of the development level. Urbanization level index acquired based on NTL data, socio-economic statistic data, land use data and population grid data are combined to establish the relevant model to achieve spatialization of macro-scale economic development index (such as GDP, population, energy consumption, carbon emissions and primary productivity). Thus, spatialized socio-economic information is available for macro-economic and overall regional policy development.

(2) Application of new NTL data. The new NTL remote sensing data can detect nighttime light with finer spatial resolution and has a radiation-detecting performance, such as NPP/VIIRS data, EROS-B data, etc. Combining BCI information obtained from high spatial resolution data with these data, it is capable of improving NTL data details and deepening application fields.

Acknowledgements

The National Civil Aerospace Pre-research Project supported this study for the 12th Five-Year Plan of State Administration for Science, Technology and Industry for National Defence. We greatly appreciate the editor and reviewers for the valuable suggestions and comments for the completion of this manuscript.

References

Cao X, Chen J, Imura H *et al.*, 2009. A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data. *Remote Sensing of Environment*, 113(10): 2205–2209.

Deng C B, Wu C S, 2012. BCI: A biophysical composition index for remote sensing of urban environments. *Remote Sensing of Environment*, 127: 247–259.

Doll C N, 2008. CIESIN thematic guide to night-time light remote sensing and its applications. Center for International Earth Science Information Network of Columbia University, Palisades, NY.

Elvidge C, Baugh K, Hobson V *et al.*, 1997a. Satellite inventory of human settlements using nocturnal radiation emissions: A contribution for the global toolchest. *Global Change Biology*, 3(5): 387–395.

Elvidge C D, Baugh K E, Kihn E A *et al.*, 1997b. Mapping city lights with nighttime data from the DMSP Operational Linescan System. *Photogrammetric Engineering and Remote Sensing*, 63(6): 727–734.

Elvidge C D, Cinzano P, Pettit D R *et al.*, 2007. The Nightsat mission concept. *International Journal of Remote Sensing*, 28(12): 2645–2670.

Forbes D J, 2013. Multi-scale analysis of the relationship between economic statistics and DMSP-OLS night light images. *Giscience & Remote Sensing*, 50(5): 483–499.

Forsythe K W, 2004. Pansharpened Landsat 7 imagery for improved urban area classification. *Geomatica*, 58(1): 23–31.

He C, Ma Q, Liu Z *et al.*, 2013. Modeling the spatiotemporal dynamics of electric power consumption in Mainland China using saturation-corrected DMSP/OLS nighttime stable light data. *International Journal of Digital Earth*, 7(12): 1–22.

He C Y, Liu Z F, Tian J *et al.*, 2014. Urban expansion dynamics and natural habitat loss in China: A multiscale landscape perspective. *Global Change Biology*, 20: 2886–2902.

He C Y, Shi P J, Li J G *et al.*, 2006. Restoring urbanization process in China in the 1990s by using non-radiance calibrated DMSP/OLS nighttime light imagery and statistical data. *Chinese Science Bulletin*, 51(13): 1614–1620.

Henderson M, Yeh E T, Gong P *et al.*, 2003. Validation of urban boundaries derived from global night-time satellite imagery. *International Journal of Remote Sensing*, 24(3): 595–609.

Imhoff M L, Lawrence W T, Stutzer D C *et al.*, 1997. A technique for using composite DMSP/OLS "city lights" satellite data to map urban area. *Remote Sensing of Environment*, 61(3): 361–370.

Li M, Wu B, Yan C *et al.*, 2004. Estimation of vegetation fraction in the upper basin of Miyun Reservoir by remote sensing. *Resources Science*, 26(4): 153–159. (in Chinese)

Liu J Y, Deng X Z, Liu M L *et al.*, 2002. Study on the spatial patterns of land-use change and analyses of driving

forces in northeastern China during 1990–2000. *Chinese Geographical Science*, 12(4): 299–308.
Liu J Y, Zhang Z X, Xu X L *et al.*, 2010. Spatial patterns and driving forces of land use change in China during the early 21st century. *Journal of Geographical Sciences*, 20(4): 483–494.
Liu Z F, He C Y, Zhang Q F *et al.*, 2012. Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008. *Landscape and Urban Planning*, 106(1): 62–72.
Lu D S, Li G Y, Kuang W H *et al.*, 2014. Methods to extract impervious surface areas from satellite images. *International Journal of Digital Earth*, 7(2): 93–112.
Lu D S, Tian H Q, Zhou G M *et al.*, 2008. Regional mapping of human settlements in southeastern China with multisensor remotely sensed data. *Remote Sensing of Environment*, 112(9): 3668–3679.
Ma T, Zhou C H, Pei T *et al.*, 2012. Quantitative estimation of urbanization dynamics using time series of DMSP/OLS nighttime light data: A comparative case study from China's cities. *Remote Sensing of Environment*, 124: 99–107.
Mundia C, Aniya M, 2005. Analysis of land use/cover changes and urban expansion of Nairobi city using remote sensing and GIS. *International Journal of Remote Sensing*, 26(13): 2831–2849.
Pandey B, Joshi P K, Seto K C, 2013. Monitoring urbanization dynamics in India using DMSP/OLS night time lights and SPOT-VGT data. *International Journal of Applied Earth Observation and Geoinformation*, 23: 49–61.
Potere D, Schneider A, Angel S *et al.*, 2009. Mapping urban areas on a global scale: which of the eight maps now available is more accurate? *International Journal of Remote Sensing*, 30(24): 6531–6558.
Schneider A, 2012. Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach. *Remote Sensing of Environment*, 124: 689–704.
Schneider A, Friedl M A, Potere D, 2010. Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions'. *Remote Sensing of Environment*, 114(8): 1733–1746.
Scott D, Petropoulos G, Moxley J *et al.*, 2014. Quantifying the physical composition of urban morphology throughout wales based on the time series (1989–2011) analysis of Landsat TM/ETM+ images and supporting GIS data. *Remote Sensing*, 6(12): 11731–11752.
Small C, Elvidge C D, 2013. Night on Earth: Mapping decadal changes of anthropogenic night light in Asia. *International Journal of Applied Earth Observation and Geoinformation*, 22: 40–52.
Small C, Elvidge C D, Balk D. *et al.*, 2011. Spatial scaling of stable night lights. *Remote Sensing of Environment*, 115(2): 269–280.
Small C, Pozzi F, Elvidge C D, 2005. Spatial analysis of global urban extent from DMSP-OLS night lights. *Remote Sensing of Environment*, 96(3/4): 277–291.
Wang W, Cheng H, Zhang L, 2012. Poverty assessment using DMSP/OLS night-time light satellite imagery at a provincial scale in China. *Advances in Space Research*, 49(8): 1253–1264.
Weng Q H, 2012. Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sensing of Environment*, 117: 34–49.
Wu C, Deng C, Jia X, 2014. Spatially constrained multiple endmember spectral mixture analysis for quantifying subpixel urban impervious surfaces. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 7(6): 1976–1984.
Wu J S, Wang Z, Li W F *et al.*, 2013. Exploring factors affecting the relationship between light consumption and GDP based on DMSP/OLS nighttime satellite imagery. *Remote Sensing of Environment*, 134: 111–119.
Yang Y, He C Y, Zhang Q F *et al.*, 2013. Timely and accurate national-scale mapping of urban land in China using Defense Meteorological Satellite Program's Operational Linescan System nighttime stable light data. *Journal of Applied Remote Sensing*, 7(1) 073535: 1–18.
Zhang Q L, Schaaf C, Seto K C, 2013. The vegetation adjusted NTL urban index: A new approach to reduce saturation and increase variation in nighttime luminosity. *Remote Sensing of Environment*, 129: 32–41.
Zhang X Y, Schaaf C B, Friedl M A *et al.*, 2002. MODIS tasseled cap transformation and its utility. *IGARSS 2002: IEEE International Geoscience and Remote Sensing Symposium and 24th Canadian Symposium on Remote Sensing*, *Vols I-VI, Proceedings*, 1063–1065.
Zhang Y, Zhang H, Lin H, 2014. Improving the impervious surface estimation with combined use of optical and SAR remote sensing images. *Remote Sensing of Environment*, 141(2): 155–167.
Zhou Q, Li B, Kurban A, 2008. Trajectory analysis of land cover change in arid environment of China. *International Journal of Remote Sensing*, 29(4): 1093–1107.
Zhou Y, Smith S J, Elvidge C D *et al.*, 2014. A cluster-based method to map urban area from DMSP/OLS nightlights. *Remote Sensing of Environment*, 147(18): 173–185.
Zhuo L, Ichinose T, Zheng J *et al.*, 2009. Modelling the population density of China at the pixel level based on DMSP/OLS non-radiance-calibrated night-time light images. *International Journal of Remote Sensing*, 30(4): 1003–1018.