This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Urban Built-Up Area Extraction From Log-Transformed NPP-VIIRS Nighttime Light Composite Data

Bailang Yu, *Senior Member, IEEE*, Min Tang, Qiusheng Wu, Chengshu Yang, Shunqiang Deng, Kaifang Shi, Chen Peng, Jianping Wu, and Zuoqi Chen, *Member, IEEE*

***Abstract*— Accurate information on urban areas at regional and global scales is required for various socioeconomic and environmental applications. The nighttime light (NTL) composite data have proven to be an effective data source for extracting urban areas. Various urban mapping methods have been proposed in the literature to extract urban built-up areas from the Defense Meteorological Satellite Program's Operational Linescan System NTL data with a variable accuracy. However, most of the previous methods cannot be directly applied to the NTL data derived from the Suomi National Polar-orbiting Partnership Satellite with the Visible Infrared Imaging Radiometer Suite (NPP-VIIRS) sensor onboard. In this letter, we introduced a logarithmic transformation to preprocess the NPP-VIIRS NTL composite data. Then, four popular methods for urban built-up area extraction were tested using the original and log-transformed NTL data, respectively. The selected methods included the thresholding technique, Sobel-based edge detection, neighborhood statistics analysis, and watershed segmentation. The accuracy of the results was evaluated through validating the urban areas derived using each method against the referenced urban areas obtained from the National Land Cover Database for the U.S.. The results indicated that logarithmic transformation is an effective procedure for enhancing the difference between urban built-up areas and nonurban areas. The selected methods for urban built-up area extraction were found to perform better on the log-transformed NTL data than the original NTL data.**

***Index Terms*— Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS), logarithmic transformation, National Polar-orbiting Partnership Visible Infrared Imaging Radiometer Suite (NPP-VIIRS), nighttime light (NTL), urban built-up area extraction.**

I. Introduction

ACCURATE information on urban areas at regional and global scales is required for various socioeconomic and environmental applications [1]–[5]. The nighttime light (NTL) composite data from the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) launched in the 1970s have been proven to be an effective data source for extracting urban areas. Various urban mapping methods, such as thresholding technique [6], neighborhood statistics analysis [7], and watershed segmentation [8], have been proposed in the literature to extract urban built-up areas from the DMSP-OLS NTL data. Nevertheless, the DMSP-OLS NTL data products have a set of shortcomings, such as coarse spatial resolution (~1 km), limited quantization and saturation within urban centers, and lack of onboard calibration [9]–[11]. As a successor to the DMSP-OLS, the Suomi National Polar-orbiting Partnership (NPP) satellite with the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor onboard was launched in October 2011. The NPP-VIIRS NTL composite data provide several improvements over the DMSP-OLS NTL data, including finer spatial resolution (~500 m), lower radiance detection limits, and wider dynamic range [12], [13]. Shi *et al.* [14] found that the NPP-VIIRS NTL composite data could also be a potentially effective urban mapping tool. Sharma *et al.* [15] proposed a new index by combining the Moderate Resolution Imaging Spectroradiometer-based multispectral data and the NPP-VIIRS NTL composite data for extracting global urban built-up areas. How to efficiently extract the urban built-up area from the NPP-VIIRS NTL composite data is a topic of growing interest. However, a few studies have quantitatively evaluated the applicability of urban area extraction methods designed for the DMSP-OLS NTL data to NPP-VIIRS NTL composite data [16].

The objective of this letter was to propose a preprocessing method by applying a logarithmic transformation to the original NPP-VIIRS NTL composite data for urban built-up area extraction. To demonstrate the effectiveness of the proposed method, we selected four popular methods (e.g., thresholding technique, Sobel-based edge detection, neighborhood statistics analysis, and watershed segmentation) for urban built-up area

Manuscript received September 24, 2017; revised January 11, 2018 and March 12, 2018; accepted April 24, 2018. This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFE0100700, in part by the National Natural Science Foundation of China under Grant 41471449, in part by the Innovation Program of Shanghai Municipal Education Commission under Grant 15ZZ026, and in part by the Fundamental Research Funds for the Central Universities of China. *(Corresponding author: Zuoqi Chen.)*

B. Yu, M. Tang, C. Yang, S. Deng, K. Shi, C. Peng, J. Wu, and Z. Chen are with the Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University, Shanghai 200241, China, and also with the School of Geographic Sciences, East China Normal University, Shanghai 200241, China (e-mail: blyu@geo.ecnu.edu.cn; tangmin0321@163.com; chengshu.y@qq.com; dengshunqiang@outlook.com; shikf1986@126.com; peng-chen1992@foxmail.com; jpwu@geo.ecnu.edu.cn; zqchen@geo.ecnu.edu.cn).

Q. Wu is with the Department of Geography, Binghamton University, State University of New York, Binghamton, NY 13902 USA (e-mail: wqs@binghamton.edu).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the author.

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2018.2830797

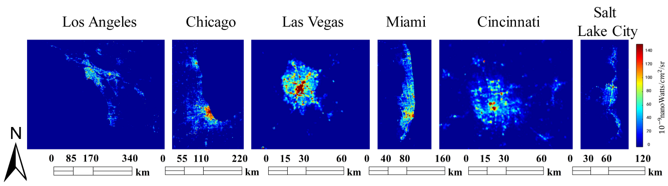

Fig. 1. NPP-VIIRS NTL images of six study areas.

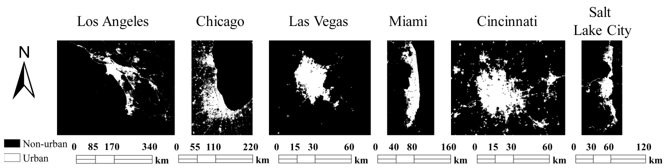

Fig. 2. Urban built-up areas extracted from NLCD 2011.

extraction and applied them to both the original and log-transformed NPP-VIIRS NTL composite data. The accuracy of the results was evaluated through validating the urban areas derived using each method against the referenced urban areas obtained from the National Land Cover Database (NLCD) for the United States.

II. Study Area and Data

A. Study Area

To test the performance of our proposed method, we selected 38 major U.S. cities as our study sites, which are distributed across the contiguous U.S. with a diverse urbanization level. Due to the journal page limit policy, we only showed six cities in Sections III and IV which were selected based on the same reason, including Los Angeles, CA, USA; Chicago, IL, USA; Las Vegas, NV, USA; Miami, FL, USA; Cincinnati, OH, USA; and Salt Lake City, UT, USA. The remaining 32 selected cities are listed in Table S1 in the Supplementary Material.

B. Data

The monthly NPP-VIIRS NTL composite data of April 2012 were acquired from the National Geophysical Data Center, National Oceanic and Atmospheric Administration (NOAA) at http://ngdc.noaa.gov/eog/download.html. The images were converted to the Albers Equal-Area Conic projection and resampled to 500 m (Fig. 1).

We applied the global raster water mask [17] to exclude water pixels for reducing the influence of water reflecting moonlight at night. Subsequently, we used the NLCD 2011 to extract urban pixels, where the percent developed imperviousness is higher than 20% [18], as the reference data (Fig. 2).

III. Methodology

A. Logarithmic Transformation of NPP-VIIRS NTL Composite Data

Fig. 3 shows a spatial distribution comparison of NTL values among the DMSP-OLS stable lights, original NPP-VIIRS composite data, and logarithmic transformed NPP-VIIRS composite data across Las Vegas, NV, USA.

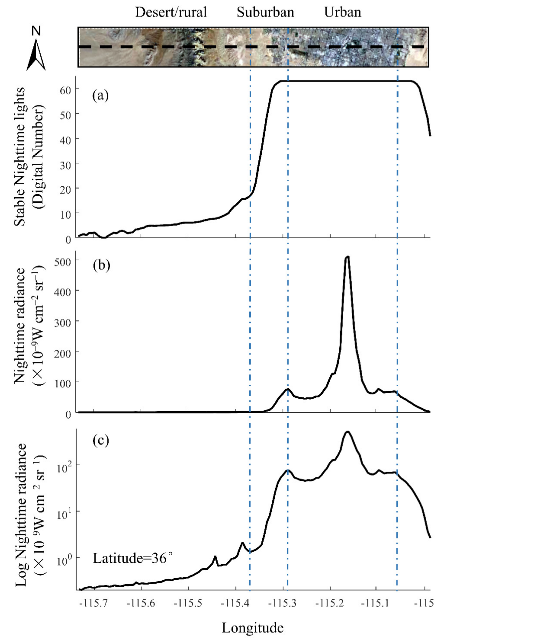

Fig. 3. Latitudinal transect of NTLs across Las Vegas, NV, USA. (a) DMSP-OLS stable NTL. (b) NPP-VIIRS NTL radiances. (c) Log-transformed NPP-VIIRS NTL radiances.

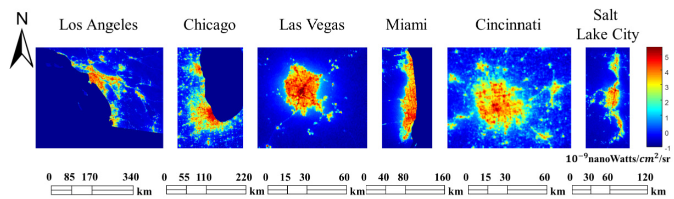

Fig. 4. Log-transformed NPP-VIIRS NTL images of six study areas.

Las Vegas was a city built in desert, which makes it an obvious gap between the urban and nonurban areas and to be a typical area to illustrate the effect of logarithmic transformation. For the DMSP-OLS stable NTL data, only minor radiance changes can be observed within urban and nonurban areas, respectively. The limited quantization results in a truncated digital number (DN) value distribution, which transforms all DN values larger than 63 as 63. The only dramatic DN value change of the truncated DMSP-OLS stable data appears around suburban location which allows it to be detected easily. The detected suburban area can be further used for urban extraction.

Compared with the DMSP-OLS instrument, the NPP-VIIRS offers a substantial number of improvements in terms of spatial resolution, dynamic range, quantization, and calibrations [13], [19]. With the improved spatial resolution, more radiance values of complicated landscapes can be observed. With the radiance value of the NPP-VIIRS NTL composite data in the floating-point format rather than integer format, it increases generally, which makes minor radiance changes hard to be detected. Also with the wider dynamic range, a radiance value of pixels within urban areas can be recorded without truncation. Owing to the extremely high development of the urban core, for the original NPP-VIIRS NTL composite data, radiance range within and around the urban core is much

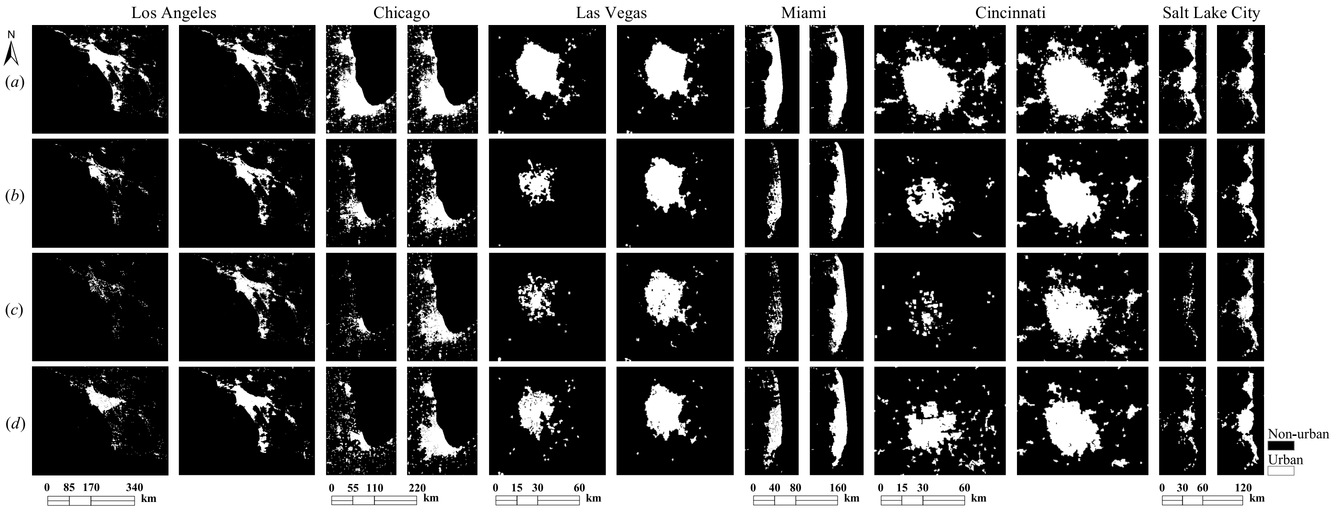

Fig. 5. Urban built-up areas extracted from (Left) original and (Right) log-transformed NPP-VIIRS NTL data using four methods. (a) Thresholding technique. (b) Sobel-based edge detection. (c) Neighborhood statistics analysis. (d) Watershed segmentation.

wider, and its radiance variance is much larger. Mostly, this means a bias for the entire radiance distribution and little spatial information to be observed within the nonurban and suburban areas. Therefore, although sharp radiance rises can be observed, the rise within the urban area is even sharper than that around the suburban zone [Fig. 3(b)]. Radiance variations within nonurban areas and suburban areas are suppressed compared to the DMSP-OLS stable NTL, which results in difficulty for classification.

The logarithmic transformation is an effective function for normalizing such skewed and biased data, which makes it easier to visualize, analyze, and interpret the data [20]. This logarithmic transformation can be expressed as

$$\text{Log\_VIIRS}_i = \ln(\text{VIIRS}_i) \tag{1}$$

where Log_VIIRS_i and VIIRS_i indicate the logarithmic and original NPP-VIIRS NTL radiance, respectively.

As can be seen in Fig. 3, it is as follows.

1) For the nonurban areas, the DMSP-OLS stable NTL image shows gentle DN value changes, while the radiance range of the original NPP-VIIRS NTL composite data is small. After the logarithmic transformation, the sharp radiance jump around the urban core is suppressed, and the radiance variance within the nonurban and suburban areas is strengthened.
2) For the suburban area, the DMSP-OLS DN values show a dramatic increase from dim lights to a saturated value. This is most likely owing to the overglow effect and the coarse spatial resolution of the DMSP-OLS NTL data [21]. In contrast, the radiance of the original and log-transformed NPP-VIIRS NTLs increases gradually. Besides, more detailed spectrum information appears with gentle fluctuations.
3) In the urbanized areas, DN values of DMSP-OLS NTLs are stable and reach the peak value of 63. However, the landscape of the NPP-VIIRS NTL data is uneven, which may be attributed to unbalanced urban development. Furthermore, although the radiance of NPP-VIIRS NTLs under logarithmic transformation similarly shows prominent variations, the variations tend to be milder. As can be seen in Fig. 3(c), the radiance jump around suburban is detectable after the logarithmic transformation.

B. Urban Built-Up Area Extraction From Log-Transformed NPP-VIIRS NTL Composite Data

After applying a logarithmic transformation to the NPP-VIIRS NTL composite data, we tested the applicability of four popular methods previously designed to extract urban built-up areas from the DMSP-OLS NTL data. The four selected methods include the thresholding technique based on the sudden jumps of urban perimeter [6], Sobel-based edge detection [22], neighborhood statistics analysis [7], and watershed segmentation [8]. For a detailed description of these four methods, readers are recommended to refer to the above-mentioned four relevant letters.

IV. RESULTS AND DISCUSSION

The log-transformed NPP-VIIRS NTL images of the six study areas are shown in Fig. 4. Compared with the original NPP-VIIRS NTL images shown in Fig. 1, the log-transformed NTL images clearly exhibit more spatial variations within cities. The rate of change of the log-transformed NPP-VIIRS NTL data decreases with increasing NTL value. To evaluate how the logarithmic transformation impacts the urban built-up area extraction, four popular methods were selected to extract urban built-up areas using the NPP-VIIRS NTL composite data without and with (Fig. 5) logarithmic transformation, respectively. When using thresholding technique, the comparison results indicate that the logarithmic transformation has fewer impacts on the urban built-up area extraction. However, for the other three methods, the logarithmic transformation process improves the quality of extracted urban built-up areas by comparing with the reference data (Fig. 2).

A quantitative analysis, including user's accuracy (Table I), producer's accuracy (Table II), and overall accuracy (Table III), of the four methods is also provided. The results

TABLE I
USER'S ACCURACY OF EACH METHOD

Cities	Thresholding technique		Sobel-based edge detection		Neighborhood statistics		Watershed segmentation	
	Original	Log	Original	Log	Original	Log	Original	Log
Los Angeles	0.93	0.93	0.97	0.90	0.92	0.87	0.82	0.85
Chicago	0.96	0.96	0.97	0.89	0.93	0.87	0.77	0.84
Las Vegas	0.99	0.99	0.94	0.84	0.87	0.82	0.83	0.84
Miami	0.99	0.98	0.99	0.93	0.96	0.91	0.92	0.94
Cincinnati	0.97	0.97	0.97	0.84	0.92	0.80	0.70	0.83
Salt Lake City	0.94	0.95	0.97	0.88	0.93	0.86	0.88	0.87

*Original and Log indicate the original and log-transformed NPP-VIIRS NTL data, respectively.

TABLE II
PRODUCER'S ACCURACY OF EACH METHOD

Cities	Thresholding technique		Sobel-based edge detection		Neighborhood statistics		Watershed segmentation	
	Original	Log	Original	Log	Original	Log	Original	Log
Los Angeles	0.78	0.78	0.50	0.81	0.17	0.83	0.52	0.82
Chicago	0.69	0.69	0.38	0.77	0.17	0.77	0.36	0.78
Las Vegas	0.67	0.67	0.47	0.94	0.40	0.92	0.69	0.95
Miami	0.77	0.82	0.51	0.90	0.27	0.91	0.62	0.90
Cincinnati	0.57	0.58	0.45	0.84	0.16	0.86	0.62	0.84
Salt Lake City	0.76	0.75	0.38	0.83	0.16	0.83	0.41	0.90

*Original and Log indicate the original and log-transformed NPP-VIIRS NTL data, respectively.

TABLE III
OVERALL ACCURACY OF EACH METHOD

Cities	Thresholding technique		Sobel-based edge detection		Neighborhood statistics		Watershed segmentation	
	Original	Log	Original	Log	Original	Log	Original	Log
Los Angeles	0.98	0.95	0.93	0.96	0.88	0.98	0.62	0.98
Chicago	0.92	0.86	0.82	0.90	0.76	0.94	0.41	0.94
Las Vegas	0.95	0.95	0.94	0.97	0.93	0.97	0.72	0.97
Miami	0.95	0.94	0.87	0.95	0.81	0.97	0.72	0.97
Cincinnati	0.86	0.86	0.89	0.94	0.84	0.93	0.58	0.94
Salt Lake City	0.95	0.95	0.92	0.96	0.89	0.96	0.52	0.96

*Original and Log indicate the original and log-transformed NPP-VIIRS NTL data, respectively.

show that, except for the thresholding technique, logarithmic transformation is good for the other three urban built-up mapping methods using the NPP-VIIRS NTL composite data. The accuracy of the thresholding technique has a negligible difference between using the original and log-transformed NPP-VIIRS NTL composite data, and either of them is acceptable. Their producer's accuracy of most cities is lower than 0.80. The reason could be that the fine spatial resolution and high-radiance detection ability of the NPP-VIIRS NTL composite data make the NPP-VIIRS NTL value in urban built-up areas heterogeneous and fluctuant, which causes that the estimation of thresholding value based on sudden jumps of the urban perimeter becomes difficult.

For the other three methods, when the NPP-VIIRS NTL composite data were logarithmic transformed, the user's accuracy and overall accuracy in these six study areas are all close to or greater than 80%, and are much better than using the original NPP-VIIRS NTL composite data. For example, in Las Vegas, the Sobel-based edge detection method with the log-transformed NPP-VIIRS NTL composite data has a user's accuracy of 94%, which is much higher than the accuracy (47%) of using the original NPP-VIIRS NTL composite data. However, the Sobel-based edge detection and neighborhood statistics analysis methods using log-transformed NPP-VIIRS NTL composite data have a lower user's accuracy than using the original NPP-VIRIS NTL composite data but are good (greater than 0.80) enough to accept the results.

To evaluate the reliability of the proposed method, we extended the examination to additional 32 U.S. cities (see Table S1 in the Supplementary Material), which have a diverse urbanization level. When using the Sobel-based edge detection method, neighborhood statistics analysis method, and watershed segmentation method, the accuracy assessment (Tables S2–S4 in the Supplementary Material) shows that most of these cities have a substantial improvement in producer's accuracy (from 47% to 83%, from 41% to 86%, and from 72% to 83%, respectively, on average) and overall accuracy (from 87% to 92%, from 70% to 91%, and from 62% to 93%, respectively, on average). However, the accuracy of the thresholding technique did not improve after logarithmic transformation. On average, its overall accuracy is 86% for the original NPP-VIIRS NTL data and 84% for the log-transformed NPP-VIIRS NTL data.

The result indicates that the logarithmic transformation of NPP-VIIRS NTL data could help to enlarge the NTL difference between urban built-up area and nonurban area and also improve the performance of urban built-up area extraction.

V. Conclusion

In this letter, we proposed a logarithmic function to transform the distribution of the NPP-VIIRS NTL composite data and evaluated the applicability of four selected methods which were originally designed for the DMSP-OLS NTL data. After transformation, the NPP-VIIRS NTL composite data have a narrow range and a low variance, which enlarges the gap between urban built-up areas and nonurban areas. By examining six major U.S. cities and extended 32 U.S. cities with four popular methods, we found that the thresholding technique has a relatively low producer's accuracy on both the original and log-transformed NPP-VIIRS NTL composite data. In contrast, the other three methods (i.e., Sobel-based edge detection, neighborhood statistics analysis, and watershed segmentation) have a significant improvement on the log-transformed NPP-VIIRS NTL composite data compared with the original NTL data. It is demonstrated that the logarithmic transformation can be used as an effective preprocessing method and enhances the performance of urban area extraction from the NPP-VIIRS NTL composite data.

References

[1] Y. Chen, X. Li, X. Liu, B. Ai, and S. Li, "Capturing the varying effects of driving forces over time for the simulation of urban growth by using survival analysis and cellular automata," *Landscape Urban Planning*, vol. 152, pp. 59–71, Aug. 2016.

[2] L. Zhang, J. Peng, Y. Liu, and J. Wu, "Coupling ecosystem services supply and human ecological demand to identify landscape ecological security pattern: A case study in Beijing–Tianjin–Hebei region, China," *Urban Ecosyst.*, vol. 20, no. 3, pp. 701–714, 2016.

[3] X. Liu *et al.*, "A future land use simulation model (FLUS) for simulating multiple land use scenarios by coupling human and natural effects," *Landscape Urban Planning*, vol. 168, pp. 94–116, Dec. 2017.

[4] Y. Chen, X. Li, X. Liu, and B. Ai, "Modeling urban land-use dynamics in a fast developing city using the modified logistic cellular automaton with a patch-based simulation strategy," *Int. J. Geograph. Inf. Sci.*, vol. 28, no. 2, pp. 234–255, 2014.

[5] J. N. Inkoom, B. K. Nyarko, and K. B. Antwi, "Explicit modeling of spatial growth patterns in Shama, Ghana: An agent-based approach," *J. Geovis. Spatial Anal.*, vol. 1, nos. 1–2, p. 7, Oct. 2017.

[6] M. L. Imhoff, W. T. Lawrence, D. C. Stutzer, and C. D. Elvidge, "A technique for using composite DMSP/OLS 'city lights' satellite data to map urban area," (in English), *Remote Sens. Environ.*, vol. 61, no. 3, pp. 361–370, Sep. 1997.

[7] Y. Su *et al.*, "A new method for extracting built-up urban areas using DMSP-OLS nighttime stable lights: A case study in the Pearl River Delta, southern China," *GISci. Remote Sens.*, vol. 52, no. 2, pp. 218–238, 2015.

[8] Y. Zhou, S. J. Smith, C. D. Elvidge, K. Zhao, A. Thomson, and M. Imhoff, "A cluster-based method to map urban area from DMSP/OLS nightlights," *Remote Sens. Environ.*, vol. 147, pp. 173–185, May 2014.

[9] P. C. Sutton, "A scale-adjusted measure of 'urban sprawl' using nighttime satellite imagery," *Remote Sens. Environ.*, vol. 86, no. 3, pp. 353–369, 2003.

[10] H. Letu, M. Hara, G. Tana, and F. Nishio, "A saturated light correction method for DMSP/OLS nighttime satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 389–396, Feb. 2012.

[11] H. Letu *et al.*, "Estimating energy consumption from night-time DMPS/OLS imagery after correcting for saturation effects," *Int. J. Remote Sens.*, vol. 31, no. 16, pp. 4443–4458, 2010.

[12] C. Cao, X. Xiong, S. Blonski, Q. Liu, B. Guenther, and F. Weng, "Suomi NPP VIIRS on-orbit performance, data quality, and new applications," *Proc. SPIE*, vol. 8528, p. 85280D, Nov. 2012.

[13] C. D. Elvidge, K. E. Baugh, M. Zhizhi, and F.-C. Hsu, "Why VIIRS data are superior to DMSP for mapping nighttime lights," in *Proc. Asia–Pacific Adv. Netw.*, vol. 35. 2013, pp. 19–62.

[14] K. Shi, C. Huang, B. Yu, B. Yin, Y. Huang, and J. Wu, "Evaluation of NPP-VIIRS night-time light composite data for extracting built-up urban areas," *Remote Sens. Lett.*, vol. 5, no. 4, pp. 358–366, 2014.

[15] R. C. Sharma, R. Tateishi, K. Hara, S. Gharechelou, and K. Iizuka, "Global mapping of urban built-up areas of year 2014 by combining MODIS multispectral data with VIIRS nighttime light data," *Int. J. Digit. Earth*, vol. 9, no. 10, pp. 1004–1020, Apr. 2016.

[16] Y. Dou, Z. Liu, C. He, and H. Yue, "Urban land extraction using VIIRS nighttime light data: An evaluation of three popular methods," *Remote Sens.*, vol. 9, no. 2, p. 175, 2017.

[17] M. L. Carroll, J. R. Townshend, C. M. DiMiceli, P. Noojipady, and R. A. Sohlberg, "A new global raster water mask at 250 m resolution," *Int. J. Digit. Earth*, vol. 2, no. 4, pp. 291–308, 2009.

[18] C. Homer *et al.*, "Completion of the 2011 National Land cover database for the conterminous United States—Representing a decade of land cover change information," *Photogramm. Eng. Remote Sens.*, vol. 81, no. 5, pp. 346–354, 2015.

[19] K. Shi *et al.*, "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.*, vol. 6, no. 2, pp. 1705–1724, Feb. 2014.

[20] C. Feng *et al.*, "Log-transformation and its implications for data analysis," *Shanghai Arch. Psychiatry*, vol. 26, no. 2, pp. 105–109, 2014.

[21] T. Ma, C. Zhou, T. Pei, S. Haynie, and J. Fan, "Responses of Suomi-NPP VIIRS-derived nighttime lights to socioeconomic activity in China's cities," *Remote Sens. Lett.*, vol. 5, no. 2, pp. 165–174, 2014.

[22] Y. Chen, "Edge detection in gray image," M.S. thesis, School Electron. Eng., Univ. Electron. Sci. Technol. Chengdu, China, 2009.